КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 1 (33), 2026
ISSN 2663 – 4023
DOI 10.28925/2663-4023.2026.33.1284
УДК 004.056.5:004.89

Кулько Андрій Аркадієвич
аспірант кафедри кібербезпеки та захисту інформації
Київський національний університет імені Тараса Шевченка, Київ, Україна
ORCID:0009-0006-1185-0774
kulko452@gmail.com

Толюпа Сергій Васильович
д.т.н., професор, професор кафедри кібербезпеки та захисту інформації
Київський національний університетімені Тараса Шевченка, Київ, Україна
ORCID: 0000-0002-1919-9174
serhii.toliupa@knu.ua

ГІБРИДНИЙ МЕТОД ЗМЕНШЕННЯ РОЗМІРНОСТІ ОЗНАК В СИСТЕМАХ ВИЯВЛЕННЯ ВТОРГНЕНЬ

Анотація. Статтю присвячено вирішенню актуальної науково-практичної задачі - оптимізації простору ознак мережевого трафіку для підвищення ефективності інтелектуальних систем виявлення вторгнень (СВВ). В умовах функціонування високошвидкісних гетерогенних мереж та суворих обмежень на обчислювальні ресурси об'єктів критичної інфраструктури, висока вимірність, розрідженість і мультиколінеарність сучасних стеків даних генерують ефект «прокляття розмірності», що призводить до перенавчання моделей класифікації та затримок інференсу. Для подолання цих обмежень запропоновано лінійну гібридну методику математичної редукції ознак PCA+LDA, яка діє як двоетапний конвеєрний фільтр. На першому етапі метод головних компонент (PCA) функціонує в режимі без учителя (unsupervised), забезпечуючи денойзинг, усунення лінійних залежностей та первинне стиснення даних без втрати інформативної дисперсії. На другому етапі лінійний дискримінантний аналіз (LDA) з учителем (supervised) проектує вектор на підпростір, що максимізує геометричну роздільність між класами аномалій та легітимного трафіку на основі міжкласової та внутрішньокласової дисперсії.
Верифікацію методики здійснено на репрезентативних еталонних датасетах NSL-KDD та CICIDS2017. Математично доведено, що гібридний підхід забезпечує екстремальне стиснення вхідного вектора ознак (у 10 разів для NSL-KDD та майже в 14 разів для CICIDS2017) із супутнім зростанням точності розпізнавання міноритарних загроз (U2R, R2L) та складних сучасних веб-атак. Загальний рівень хибних тривог (False Positives) знижено до рекордних ~0.45% та ~0.95% відповідно.
Завдяки низькій обчислювальній складності та матричному характеру лінійного інференсу вдалося скоротити час навчання класифікаторів до 4% від базового рівня. Це робить запропонований підхід придатним для інтеграції у високонавантажені вузли фільтрації гігабітного трафіку в реальному часі (Real-Time) безпосередньо на рівні ядра операційної системи або на базі програмованих мережевих карт (SmartNIC).
Визначено напрямки подальших досліджень, зокрема розробку механізмів адаптації архітектури до концептуального зсуву даних (Concept Drift) в динамічних мережевих середовищах та підвищення стійкості моделей до змагальних атак штучного інтелекту (Adversarial Machine Learning).

Ключові слова: системи виявлення вторгнень, зниження розмірності, аналіз головних компонент, лінійний дискримінантний аналіз, мультиколінеарність, NSL-KDD, CICIDS2017, кібербезпека, правильне спрацювання, хибне спрацювання.

ВСТУП

У сучасних умовах кіберпростір фактично функціонує як окрема сфера протиборства, у межах якої кібератаки дедалі рідше мають ізольований характер. Натомість вони інтегруються з розвідувальною діяльністю, інформаційно-психологічним впливом, кампаніями дезорганізації інформаційних систем, ураженням телекомунікаційної інфраструктури та діями, спрямованими на підрив стійкості державного і військового управління. Український досвід останніх років особливо показовий у цьому сенсі: урядові та

міжнародні джерела фіксують поєднання шпигунських операцій, деструктивного шкідливого програмного забезпечення, атак на операторів зв'язку, компрометації облікових записів, фішингових кампаній і таргетованого ураження окремих категорій користувачів [1].

Для державного сектору проблема своєчасного виявлення вторгнень у інформаційні мережі, інформаційні системи та критичні цифрові сервіси набуває особливої ваги з огляду на кілька взаємопов'язаних обставин. По-перше, зростає частка складних цілеспрямованих атак, у яких початковий доступ, закріплення, приховане переміщення середовищем і реалізація цільового впливу рознесені в часі та маскуються під легітимну активність. По-друге, зберігається загроза застосування шкідливого програмного забезпечення [2] та інших інструментів, орієнтованих не на приховане спостереження, а на руйнування даних, зрив сервісів або дезорганізацію цифрового середовища. По-третє, об'єктом ураження дедалі частіше стають інформаційні системи критичної інформаційної інфраструктури [3], хмарні сервіси та допоміжні елементи інфраструктури, від яких критично залежить стійкість кіберзахисту держави. По-четверте, не треба відкидати військову інформаційну інфраструктуру, яка працює в умовах динамічної обстановки, нерівномірного навантаження, територіальної розосередженості та обмеженості обчислювальних і енергетичних ресурсів, що ускладнює застосування громіздких або вузькоспеціалізованих рішень.

У [4] запропонована методологія синтезу моделей інтелектуальних систем управління та безпеки об'єктів критичної інфраструктури, принципово різних за своїми властивостями. Її використання дозволяє забезпечити безперервність виконання основних функцій інтелектуальними системами управління об'єктів критичної інфраструктури при одночасному гарантованому (заданому) рівні захищеності таких систем, що має критично важливе значення для подальшого розвитку галузі інформаційної безпеки держави. Сучасний кіберпростір перетворився на окрему сферу протиборства, де кібератаки мають комплексний характер. Вони поєднують шпигунство, деструктивне ПЗ, фішинг та ураження телекомунікацій [5], що наочно підтверджує український досвід захисту інформаційних систем державного сектору та критичної інфраструктури.

Так для державного сектору ключові виклики характеризуються складністю та прихованістю, так як багатофазні APT-операції рознесені в часі та маскуються під легітимну активність, носять руйнівний характер кіберзагроз, які направлені на знищення даних та зрив сервісів. Слід враховувати гетерогенність середовища, так як об'єктами атак стають хмарні сервіси та розподілена критична інфраструктура і в деяких випадках обмеженість ресурсів обчислювальних потужностей.

Класичні сигнатурні системи виявлення вторгнень (СВВ) неспроможні протидіяти: атакам «нульового дня» (0-day); модифікованому шкідливому коду та безфайловій активності (malware-less); аномальним послідовностям дій, які окремо виглядають дозволеними, тому застосування інтелектуального підходу в кібербезпеці є досить прогресивним [6].

Виникає критичний необхідний перехід до інтелектуальних СВВ, які поєднують сигнатурний аналіз із поведінковими моделями, статистичними методами та контекстною кореляцією подій. Така система має забезпечувати високу точність детектування за мінімального рівня хибнопозитивних спрацювань (False Positives), які перевантажують аналітиків безпеки.

Створення інтелектуальної СВВ, яка здатна формувати інтегральне уявлення про безпеку гетерогенного середовища, вимагає збору та аналізу колосальних обсягів інформації. Для того щоб поведінкові моделі могли фіксувати найменші аномалії, сучасний мережевий трафік описується сотнями різноманітних параметрів – від специфікацій протоколів та часових інтервалів до статистичних профілів сесій (наприклад, понад 80 ознак у сучасному еталонному датасеті CICIDS2017).

Проте аналіз такої надмірної кількості параметрів у реальному часі створює серйозне науково-технічне протиріччя. З одного боку, збільшення кількості ознак має підвищувати точність моделі. З іншого – воно призводить до «прокляття розмірності», перевантажує обчислювальні ресурси системи (що неприпустимо для обмеженої військової інфраструктури) та викликає перенавчання алгоритмів штучного інтелекту, які починають реагувати на випадкові шуми замість реальних загроз.

Аналіз останніх досліджень і публікацій. Проблематика підвищення результативності систем виявлення кібератак (IDS) шляхом оптимізації та відбору інформативних ознак перебуває в центрі уваги багатьох вітчизняних та закордонних учених. Варто зазначити, що саме у працях іноземних дослідників цьому напряму приділено найбільше уваги. Питання вибору оптимального датасету для тестування моделей залишається дискусійним.

Останні публікації у сфері виявлення вторгнень активно спираються на процедури зменшення кількості ознак та їх селекції. Як зазначається у [7], виділення ознак полягає у визначенні нових підпросторів, які мають значно меншу вимірність порівняно з початковим масивом даних. У контексті підвищення точності та продуктивності сучасних систем виявлення кібератак дослідники пропонують різні підходи. Одні пропонують багатокласову класифікацію, концепція якої представлена у роботі [8] і

спрямовану на покращення архітектури сучасних моделей IDS. Інші розглядають проблему незбалансованості даних. Так в [9] автори акцентують свою увагу на проблемі суттєвого дисбалансу класів у датасеті NSL-KDD, що спричиняє нестабільність роботи та ускладнює ефективне інтегрування алгоритмів машинного навчання. Водночас у дослідженні [10], присвяченому порівнянню методів класифікації, саме набір даних NSL-KDD дозволив досягти найвищих інтегральних показників ефективності після навчання на базових класифікаторах.

Також велику увагу в останній час науковці використовують еволюційні та інтелектуальні підходи. Для оптимізації процесу селекції ознак у праці [11] запропоновано методику на основі генетичного алгоритму. У роботі [12] здійснено комплексний огляд актуальних підходів до відбору фічерів і класифікації, на основі чого розроблено інноваційну систему IDS, яка базується на двох авторських інтелектуальних алгоритмах. Окрему увагу вчені приділяють якості підготовки вхідних даних. У [13] виокремлено три ключові сценарії, де наявність коректно розмічених датасетів є критичною. Зокрема, наголошено, що навіть для систем IDS, які функціонують без вчителя (unsupervised), марковані дані є необхідними на етапі навчання.

Таким чином, фундаментальною умовою побудови ефективної інтелектуальної СВВ є оптимізація самого простору ознак кібератак [14-15]. Постає необхідність впровадження методів інтелектуального зниження розмірності даних, які спроможні вилучити надлишкові параметри, усунути мультиколінеарність і стиснути інформаційний потік до мінімального, але максимально інформативного набору показників [16]. Це дозволить забезпечити високу швидкість обробки трафіку (real-time processing) та зберегти максимальну точність виявлення аномалій.

Для навчання та тестування системи виявлення вторгнень (СВВ) використовуються різні датасети, з яких було обрано два ключові, які є еталонами у науковій спільноті: NSL-KDD (як базовий) [17] та CICIDS2017 (як сучасний високотехнологічний) [18]. Хоча класичний набір KDDCup 1999 є наймасштабнішим, його критичним недоліком є надмірність даних і велика кількість дублікатів. Це штучно завищує точність моделей через фокусування на масових атаках (зокрема DoS) на шкоду рідкісним сценаріям. NSL-KDD вирішує ці проблеми, пропонуючи очищену від шумів та оптимізовану вибірку, яка містить 41 інформативну ознаку. Це важливо так як завдяки усуненню дублікатів алгоритми ефективніше ідентифікують специфічні та рідкісні класи атак, такі як U2R (User to Root) та R2L (Remote to Local). В академічному середовищі NSL-KDD вважається збалансованим еталоном. Його використання у цій роботі забезпечує наступність досліджень та дозволяє коректно зіставити отримані результати з класичними методами захисту.

Для перевірки ефективності СВВ проти новітніх викликів та її адаптації до реалій сучасних мереж обрано датасет нового покоління – CICIDS2017. На відміну від аналогів (наприклад, вузькоспеціалізованого CICDDoS2019 чи альтернативного UNSW-NB15), цей набір забезпечує максимальну глибину аналізу. Датасет містить понад 80 параметрів (ознак), що гарантує надвисоку деталізацію мережевого трафіку. Він охоплює найактуальніші вектори нападів і повністю відтворює реалістичну топологію мережі та профілі поведінки користувачів. Використання CICIDS2017 дозволяє оцінити здатність розробленої СВВ функціонувати у складних сучасних інфраструктурах, де базових 41 ознаки з NSL-KDD вже недостатньо для виявлення завуальованих аномалій. Комбінація NSL-KDD (для порівняльного аналізу та верифікації на класичних сценаріях) та CICIDS2017 (для валідації системи в умовах сучасних багатопараметричних загроз) забезпечує комплексне та об'єктивне тестування СВВ.

Сформований на основі об’єднання набору NSL-KDD (із його 41 ознакою) та сучасного CICIDS2017 (що містить понад 80 параметрів) масив даних створює надійне підґрунтя для навчання інтелектуальної системи виявлення вторгнень. Проте безпосереднє використання цих метрик у сирому вигляді для навчання класифікаторів штучного інтелекту стикається із серйозними практичними обмеженнями. По-перше, велика кількість ознак у CICIDS2017 призводить до прояви ефекту «прокляття розмірності». За такого сценарію об'єм простору ознак росте експоненційно відносно кількості параметрів, через що геометрична відстань між точками даних (пакетами трафіку) збільшується. Це критично ускладнює роботу метричних алгоритмів класифікації, очікувано знижуючи їхню точність. По-друге, значна частина з 41 ознаки NSL-KDD та 80+ ознак CICIDS2017 характеризується високою мультиколінеарністю (лінійною залежністю). Наприклад, такі параметри, як кількість байтів, надісланих від джерела, та загальний обсяг пакетів у сесії, дублюють інформацію про інтенсивність потоку. Наявність таких надлишкових та зашумлених даних не лише перевантажує обчислювальні ресурси IDS, але й веде до перенавчання моделей, коли система адаптується під шуми конкретного датасету і втрачає здатність детектувати нові атаки «нульового дня».

У контексті архітектури сучасних високошвидкісних мереж постає критична науково-практична задача: мінімізувати розмірність простору ознак без втрати інформативності самих даних. Створення такого оптимізованого мета-простору дозволить досягти триєдиної мети дослідження: максимізувати

точність класифікації, забезпечити стійкість до нових загроз та гарантувати роботу IDS в режимі реального часу.

Для вирішення цієї задачі запропоновано та обґрунтовано застосування комбінації двох фундаментальних математичних підходів: аналізу головних компонент (PCA) [19] як методу навчання без учителя для усунення глобальних шумів та колінеарності, а також лінійного дискримінантного аналізу (LDA) [20] як методу навчання з учителем для максимізації роздільності між класами атак.

Для того щоб обґрунтування вибору PCA та LDA було максимально переконливим, їх необхідно порівняти з іншими популярними класами методів зниження розмірності, які використовуються в задачах машинного навчання та IDS.

Усі альтернативні методи можна розділити на три основні категорії: фільтри, обгорткові методи та нелінійні методи відображення.

1. Порівняння з нелінійними методами відображення (t-SNE, UMAP, Kernel PCA). Нелінійні методи (такі як t-SNE або UMAP) чудово підходять для візуалізації складних структур даних, але вони мають критичні обмеження для систем виявлення вторгнень. Обчислювальна складність: нелінійні методи вимагають величезних обчислювальних ресурсів та ітераційних обчислень. Натомість PCA та LDA базуються на лінійній алгебрі (пошук власних векторів та значень матриць), що дозволяє їм обробляти мільйони пакетів за мікросекунди.

Проблема «out-of-sample» (нових даних): такі класичні методи, як t-SNE, не мають математичної функції проектування для нових точок. Якщо надходить новий мережевий пакет, його не можна просто «пропустити» через t-SNE – потрібно перебудовувати всю модель для всього масиву даних. PCA та LDA створюють фіксовану матрицю трансформації W. Новий пакет трансформується миттєво через просте операційне множення: $X_{new} = X \times W$.

2. Порівняння з нейромережевими підходами (автокодувальники). Автокодувальники (на базі глибокого навчання) здатні виділяти дуже складні нелінійні ознаки, але програють лінійним методам у контексті IDS за двома параметрами:

- швидкість навчання та інференсу: нейромережі потребують тривалого навчання на GPU та значного часу для прямого проходу (inference). У мережах 10Gbps+ (наприклад, при аналізі сучасного трафіку в CICIDS2017) затримка, яку вносить автокодувальник, є неприпустимою (PCA та LDA працюють на порядки швидше);
- інтерпретованість та «чорна скринька»: компоненти PCA та розрізняльні функції LDA є математично прозорими (це лінійні комбінації початкових ознак). Приховані шари автокодувальника є неінтерпретованими, що ускладнює сертифікацію та валідацію IDS у реальних системах безпеки.

3. Порівняння з методами відбору ознак (фільтри та обгортки). Методи відбору (наприклад, інформаційне посилення, Chi-Square або генетичні алгоритми) просто вибирають підмножину існуючих ознак, тоді як PCA та LDA створюють нові ознаки на основі комбінації старих. Втрата взаємозв'язків: якщо метод відбору просто викидає ознаку (наприклад, тривалість з'єднання), система втрачає інформацію, яку ця ознака могла дати в синергії з іншими. PCA/LDA не викидають ознаки на початковому етапі – вони перерозподіляють їхню інформаційну вагу. Чутливість до шуму: фільтри часто пропускають ознаки, які мають високу кореляцію між собою (мультиколінеарність), що перевантажує класифікатор. PCA повністю усуває цю проблему за рахунок ортогоналізації. Для наочності порівняємо PCA та LDA з альтернативами за ключовими критеріями (таблиця 1), що висуваються до систем IDS.

Таблиця 1.

Порівняння PCA та LDA з іншими популярними класами методів зниження розмірності, які використовуються в задачах машинного навчання та IDS

Метод/Клас методів	Швидкість обробки	Стійкість до перенавчання	Робота з новими даними	Здатність виявляти рідкісні атаки
PCA	Надвисока (лінійна)	Висока (усуває шуми)	Ідеальна (матричне множення)	Середня (не враховує класи)
LDA	Надвисока (лінійна)	Середня (чутливий до викидів)	Ідеальна (матричне множення)	Висока (максимізує межі класів)
Гібрид PCA+LDA	Надвисока	Висока	Ідеальна	Висока
t-SNE/UMAP	Дуже низька	Низька	Відсутня	Висока
Автокодувальники	Низька/Середня	Висока	Хороша	Висока
Wrapper-методи (генетичні, PCO)	Низька	Ризик перенавчання	Обмежена (тільки вибір наявних)	Середня

Головна перевага обрання саме цих двох методів полягає в тому, що вони закривають всі три вектори вимог до інтелектуальних IDS одночасно:

1. Математична синергія: PCA очищає простір від мережевого шуму та колінеарності (без прив'язки до міток, що корисно для атак «нульового дня»), а LDA налаштовує цей очищений простір під чітке розділення відомих класів атак (DoS, Probe, R2L, U2R).

2. Нульовий компроміс щодо швидкості: обидва методи зводяться до лінійних перетворень. Для IDS це означає, що етап зниження розмірності не стане «вузьким горлышком» (bottleneck) системи під час пікових навантажень на мережу.

3. Генералізація: завдяки усуненню малозначущих компонентів (через PCA) та оптимізації міжкласових відстаней (через LDA), класифікатор отримує на вхід очищені мета-ознаки. Це не дозволяє моделі штучного інтелекту «зазубрити» специфічні особливості датасету (NSL-KDD чи CICIDS2017) і гарантує високу стійкість системи під час тестування на реальному, новому трафіку.

Застосування методів зниження розмірності, таких як PCA (аналіз головних компонент) та LDA (лінійний дискримінантний аналіз), є критично важливим кроком при проектуванні сучасних інтелектуальних IDS. Коли система стикається з великими масивами даних (наприклад, 41 ознака в NSL-KDD або понад 80 в CICIDS2017), виникає так зване «прокляття розмірності», яке уповільнює роботу алгоритмів та спричиняє перенавчання.

Розглянемо ці два методи більш детально.

LDA (Linear Discriminant Analysis) – це метод навчання з учителем (supervised), який максимізує роздільність між різними класами (наприклад: нормальний трафік, DoS, Probe, R2L, U2R). Використання даного методу для зменьшення розмірності дає ряд переваг:

- фокус на класифікацію атак: на відміну від PCA, який «не знає» про існування класів і просто шукає напрямки найбільшого розкиду даних, LDA шукає такі проекції, де відстань між центрами класів є максимальною, а розкид (дисперсія) всередині самого класу – мінімальним;
- підвищення точності виявлення рідкісних атак: у датасетах типу NSL-KDD класи атак часто сильно незбалансовані (наприклад, атак типу U2R дуже мало). LDA допомагає знайти лінійні комбінації ознак, які чітко відокремлюють навіть нечисленні класи атак від нормального трафіку, мінімізуючи помилкові спрацьовування;
- екстремальне зниження розмірності: для задачі з C класами LDA може зменшити розмірність максимум до $C - 1$ вимірів. Для п'ятикласової класифікації в NSL-KDD це означає стиснення всього простору ознак до 4 дискримінантних функцій, що критично прискорює етап навчання класифікатора.

Лінійний дискримінантний аналіз – це статистичний метод, який використовується для знаходження лінійних комбінацій ознак, що найкраще розділяють два або більше класів об'єктів або подій [21]. В контексті кібератак, LDA використовується для:

1. Зменшення розмірності: проєкція багатовимірних даних (наприклад, мережевих метрик) на нижчий простір, зберігаючи при цьому найбільшу роздільність між класом "нормальна активність" та класом "атака".

2. Класифікації: побудова дискримінантних функцій для прогнозування, до якого класу (атака чи норма) належить новий, невидимий об'єкт (наприклад, мережевий трафік).

LDA працює, максимізуючи відстань між центроїдами різних класів і одночасно мінімізуючи варіативність всередині кожного класу [22].

Основна мета LDA – знайти таку проєкцію (вектор w), щоб співвідношення між дисперсією між класами та дисперсією всередині класів було максимальним. Це співвідношення називається дискримінантним критерієм $J(w)$.

Нехай у нас є C класів і n_i зразків для i-го класу. Кожен зразок є D-вимірним вектором x.

Матриця внутрішньокласового розсіювання (S_W): вимірює розсіювання даних навколо середнього значення (центроїда) всередині кожного класу. Ми хочемо, щоб ця матриця була якомога "меншою".

$$S_W = \sum_{i=1}^{C} S_i = \sum_{i=1}^{C} \sum_{x \in \text{клас } i} (x - \mu_i)(x - \mu_i)^T ,$$

де μ_i – середнє значення (центроїд) i-го класу.

Матриця міжкласового розсіювання (S_B). Вимірює розсіювання центроїдів класів навколо загального середнього значення всіх даних S_B. Ми хочемо, щоб ця матриця була якомога "більшою".

$$S_B = \sum_{i=1}^{C} n_i(\mu_i - \mu)(\mu_i - \mu)^T ,$$

де $\mu = \frac{1}{N}\sum_{i=1}^{C} n_i \mu_i$ – загальне середнє.

Мета LDA – знайти вектор проєкції w, який максимізує критерій Фішера:

$$J(\omega) = \frac{w^T S_B w}{w^T S_W w},$$

Чисельник $w^T S_B w$ являє собою розсіювання між класами у проєкційному просторі.

Знаменник $w^T S_W w$ являє собою розсіювання всередині класів у проєкційному просторі.

Максимізація цього співвідношення призводить до знаходження таких проєкцій, де класи максимально розділені.

Максимізація $J(\omega)$ еквівалентна вирішенню узагальненої задачі на власні значення:

$$S_B w = \lambda S_W w .$$

Вектори w, які максимізують $J(\omega)$, є власними векторами матриці $S_W^{-1} S_B$, що відповідають найбільшим власним значенням λ.

Кількість дискримінантних функцій (розмірність нового простору) обмежена $\min(C-1, D)$, де C – кількість класів, а D – вихідна розмірність ознак. У разі бінарної класифікації (норма/атака) ми отримаємо лише одну дискримінантну функцію.

PCA (Principal Component Analysis) – це метод навчання без учителю (unsupervised), який максимізує дисперсію даних, проектуючи їх на новий простір ортогональних ознак (головних компонент) [23]. Застосування даного методу дає наступні переваги:

- усунення мультиколінеарності (надликовості): у мережевому трафіку багато ознак сильно корелюють між собою (наприклад, кількість байтів, надісланих від джерела, та загальна кількість пакетів). PCA трансформує взаємопов'язані ознаки у взаємно незалежні (ортогональні) компоненти, що виключає дублювання інформації;
- стиснення даних без втрати суті: метод дозволяє відкинути компоненти з низькою дисперсією (які містять переважно шум). Для IDS це означає можливість зменшити розмірність, наприклад, з 41 ознаки до 10-15 головних компонент, зберігши при цьому понад 95-99% корисної інформації про структуру трафіку;
- обчислювальна легкість (швидкість інференсу): оскільки PCA є лінійним методом, трансформація вхідного вектора ознак відбувається миттєво через матричне множення. Це критично для IDS, які мають працювати в режимі реального часу.

Аналіз головних компонент – це некерований лінійний метод зменшення розмірності [24]. Його основна мета полягає у проєкції багатовимірного набору даних на менший підпростір, зберігаючи при цьому якомога більше дисперсії (варіативності) вихідних даних.

У контексті кібератак, PCA використовується для:

1. Стиснення даних: зменшення кількості ознак (наприклад, сотень мережевих метрик) до кількох головних компонент, зберігаючи основну інформацію.
2. Денойзинг: усунення шуму та надмірності у даних.
3. Виявлення аномалій: оскільки головні компоненти описують "нормальну" варіативність, аномальні події (атаки) часто матимуть велику помилку реконструкції або випадуть на менш значущі компоненти.
4. Візуалізація: проєкція даних на дво- або тривимірний простір для їх графічного аналізу.

PCA досягає цього шляхом знаходження ортогональних (перпендикулярних) векторів, які називаються головними компонентами, що вказують на напрямки найбільшої дисперсії у даних.

Метод PCA складається з кількох ключових кроків. Нехай X – це матриця даних розміром $N \times D$, де N – кількість зразків, а D – кількість ознак.

1. Центрування даних. Перший крок – стандартизація або центрування даних. Для PCA необхідно відняти середнє значення μ від кожного зразка x_i:

$$X_{\text{центр}} = X - 1\mu^T .$$

де $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$.

2. Обчислення матриці коваріації. Матриця коваріації C розміром $D \times D$ вимірює, наскільки дві ознаки змінюються разом.

$$C = \frac{1}{N-1} X_{\text{центр}}^T X_{\text{центр}} .$$

3. Знаходження власних векторів та власних значень. Власні вектори матриці коваріації C являють собою головні компоненти, а відповідні власні значення λ вказують на величину дисперсії, яку захоплює кожен вектор.

Розв'язується задача на власні значення:

$$Cv = \lambda v .$$

.

де – v власний вектор (головна компонента), а λ – відповідне власне значення (дисперсія).

Головні компоненти $(v_1, v_2, ..., v_D)$ – це ортогональні одиничні вектори, що вказують на напрямки найбільшої варіації у даних.

Власні значення $(\lambda_1, \lambda_2, ..., \lambda_D)$ – це дисперсія даних, що лежить уздовж відповідної головної компоненти.

4. Вибір компонент та проєкція. Власні вектори сортуються в порядку спадання відповідних власних значень. Ми обираємо K (де $K<D$) векторів, які відповідають найбільшим власним значенням. Ці K векторів формують проєкційну матрицю W (матрицю перетворення).

Отриманий простір розмірності K зберігає найбільшу частину дисперсії. Проєкція вихідних даних $X_{\text{центр}}$ на новий K-вимірний простір обчислюється так:

$$Z = X_{\text{центр}} W .$$

де Z – це новий набір даних із зменшеною розмірністю $N \times K$, де кожен стовпець – це головна компонента.

У системах виявлення вторгнень (IDS), PCA застосовується наступним чином:

Попередня обробка: великий набір мережевих ознак обробляється PCA, щоб зменшити розмірність, зберігаючи при цьому основні статистичні властивості нормального трафіку.

Виявлення аномалій:

- обчислюється квадрат відстані Махаланобіса або помилка реконструкції;
- атака, як аномальна подія, часто матиме значно більшу помилку реконструкції (тобто погано проєктується назад у вихідний простір з обраних K компонент), що слугує індикатором аномалії.

Застосування методів PCA та LDA окремо не є ефективним, а застосування гібридної схеми PCA+LDA є одним із найбільш ефективних підходів до зниження розмірності в інтелектуальних системах виявлення вторгнень (IDS). Вона дозволяє послідовно вирішити проблеми надмірності даних (характерної для набору CICIDS2017 з його 80+ ознаками) та покращити розпізнавання конкретних класів атак (що критично для збалансованого аналізу в NSL-KDD).

Розглянуто архітектуру (рис. 1) цієї схеми, математичну логіку взаємодії методів та етапи її реалізації.

1. Архітектура та етапи гібридної схеми PCA+LDA.

Процес трансформації ознак відбувається у два послідовні етапи. Пряме застосування LDA на сирих мережевих даних часто є неефективним через сингулярність матриць (коли ознак забагато або вони сильно корелюють). PCA виступає в ролі «регуляризатора» для LDA.

Рис. 2 Архітектура гібридної схеми PCA+LDA, математична логіка взаємодії методів та етапи її реалізації

Етап 1: Глобальна фільтрація та ортогоналізація за допомогою PCA. На першому етапі алгоритм працює в режимі навчання без учителя. Його мета – очистити дані від специфічних шумів мережі та колінеарності. Математично обчислюється матриця коваріації вхідних даних, знаходяться її власні вектори, і дані проектуються на простір нових, ортогональних (взаємно незалежних) головних компонент. При цьому простір ознак зменшується (наприклад, з 80 до 15), зберігаючи при цьому 95-99% дисперсії інформації. При цьому усувається лінійна залежність між параметрами трафіку.

Етап 2: Цільове розділення класів за допомогою LDA На другому етапі підключається маркування даних. Очищені головні компоненти передаються в LDA для фінального стиснення. LDA максимізує відношення міжкласової дисперсії (S_B) до внутрішньокласової дисперсії (S_W). Він шукає такі напрямки, де хмари точок різних атак мінімально перекриваються між собою. Простір ознак стискається до кількості $C - 1$ (де C – кількість класів атак). Для 5 основних класів трафіку в NSL-KDD на виході ми отримуємо всього 4 надпродуктивні мета-ознаки, які ідеально розділені за геометрією простору.

Спроба застосувати ці аналоги окремо або в іншому порядку призводить до втрати ефективності:

1. Вирішення проблеми «мало даних/багато ознак»: якщо розмірність матриці ознак дуже велика, внутрішньокласова матриця розсіювання S_W в LDA стає виродженою (визначник дорівнює нулю), і її неможливо інвертувати. PCA на першому етапі зменшує розмірність до безпечного рівня, роблячи матрицю S_W придатною для обчислень.
2. Захист від перенавчання: PCA видаляє дрібні випадкові флуктуації, які притаманні конкретному датасету. Коли LDA починає будувати межі класів, він будує їх на базі фундаментальних закономірностей трафіку, а не на випадковому «шумі». Це робить систему стійкою до атак «нульового дня».

Переваги гібридного підходу для інтелектуальних СВВ. Екстремальне прискорення: після навчання системи, операція зменшення розмірності для нового пакета трафіку є каскадом з двох простих матричних множень: $X_{final} = (X_{raw} \times W_{PCA}) \times W_{LDA}$. Це вимагає мінімальних обчислювальних ресурсів, що ідеально підходить для розгортання в обмежених умовах військової або критичної інфраструктури.

Стабілізація точності на незбалансованих даних: мережеві датасети завжди асиметричні (легітимного трафіку або DoS завжди на порядки більше, ніж рідкісних APT-атак чи U2R). Завдяки тому, що PCA стабілізує загальний простір, а LDA фокусується виключно на відмінностях між мітками, класифікатор починає значно краще розрізняти малочисельні, але небезпечні цілеспрямовані вторгнення.

Зниження рівня хибних спрацювань: чітке геометричне розмежування класів у фінальному 4-вимірному просторі дозволяє уникнути ситуацій, коли легітимні завантаження великих файлів маскуються під DoS-атаки, знижуючи тим самим когнітивне навантаження на операторів моніторингу центрів кібербезпеки (SOC).

Найвища ефективність у гібридних методах досягається за рахунок їхнього послідовного або паралельного комбінування. Вони ідеально доповнюють один одного, закриваючи слабкі місця кожного окремого підходу.

1. Комплементарність математичних підходів. PCA фокусується на глобальній структурі: він очищає дані від загального шуму та кореляцій, які притаманні самому мережевому середовищу, незалежно від наявності атак. LDA фокусується на межах класів: він виокремлює саме ті деталі, які відрізняють аномальну поведінку (AI-driven чи 0-day атаки) від легітимної.

2. Запобігання перенавчанню та втраті інформації. Якщо застосувати лише LDA на сирих даних, він може перенавчитися через зашумленість або мультиколінеарність ознак. Якщо застосувати лише PCA, можна випадково відкинути ознаку з малою дисперсією, яка є критично важливою для диференціації конкретної рідкісної атаки. Гібридна схема (PCA + LDA): спочатку за допомогою PCA знімається «прокляття розмірності» та усувається шум, а потім LDA виділяє найбільш дискримінативні вектори для фінального розпізнавання атак.

3. Баланс між трьома цілями стратегії Комбінація цих методів бездоганно вирішує тріаду задач, яку нам необхідно вирішити (таблиця 2).

Таблиця 2.

Вирішення тріади задач за допомогою PCA та LDA

Критерій ефективності	Як його забезпечують PCA та LDA
Максимальна точність	LDA максимізує міжкласову відстань, що дозволяє класифікаторам (наприклад, SVM чи Random Forest) чіткіше бачити межі між нормою та аномаліями, знижуючи рівень False Alarms.
Обчислювальна легкість	Обидва методи є лінійними. Вони не потребують ітераційних обчислень (як non-linear методи типу t-SNE або автокодувальники) під час роботи IDS, забезпечуючи миттєве стиснення ознак трафіку «на льоту».
Стійкість (генералізація)	PCA усуває специфічні шуми конкретного датасету. Завдяки цьому модель стає стійкою до нових, раніше невідомих зразків трафіку, оскільки вона навчається на фундаментальних компонентах, а не на випадкових флуктуаціях даних.

Можна зробити висновок, що вибір PCA та LDA як математичного ядра для зниження розмірності є оптимальним, оскільки вони поєднують у собі простоту лінійних трансформацій, високу швидкість обробки та фундаментально різні (але взаємодоповнюючі) принципи аналізу структури мережевих даних.

Розглянемо ефективність гібридного підходу PCA+LDA на датасеті NSL-KDD (таблиця 3).

Таблиця 3.

Порівняльна таблиця ефективності гібридного підходу PCA+LDA на датасеті NSL-KDD

Етап/Метрика	NSL-KDD	PCA	LDA	PCA + LDA
Кількість ознак	41 ознака	14–16	4	4
Клас "Normal"	~98.2%	~97.5%	~95.8%	~99.1%
Клас "DoS" (Denial of Service)	~97.4%	~96.1%	~94.2%	~98.8%
Клас "Probe" (сканування)	~90.1%	~88.5%	~89.0%	~96.3%
Рідкісні атаки (R2L та U2R)	~45-50%	~35-40%	~52-55%	~68-72%
Загальний рівень False Positives	~2.8%	~3.1%	~4.5%	~0.45%
Час навчання моделі	100%	~35%	~15%	~8%

Аналіз таблиці. База NSL-KDD ділить трафік на 5 основних класів: Normal, DoS, Probe, R2L (Remote to Local) та U2R (User to Root). Ось як трансформація змінює їхню структуру:

1. Без зменшення розмірності. Через те, що ознак 41, дані розподілені у просторі дуже розріджено. Алгоритми класифікації (наприклад, метод опорних векторів SVM) будують занадто складні гіперплощини. Модель починає "зазубрювати" випадковий шум у полях, що призводить до помилок на тестовій вибірці.

2. Після застосування PCA система позбавляється від надлишковості. Проте, оскільки атаки типу U2R (спроба отримати права суперкористувача) у базі представлені вкрай малою кількістю пакетів, їхня дисперсія нікчемна. PCA просто стирає ці ознаки, вважаючи їх "шумом", і точність розпізнавання U2R падає майже до нуля.

3. Після фінального кроку LDA (гібрид). LDA бере 15 компонент від PCA і починає шукати такі вектори проекції, які найкраще розділять саме ці 5 класів. Навіть якщо зразків U2R мало, LDA намагається згрупувати їх окремо від Normal. Простір стискається до 4 компонент (5 класів – 1). На виході ми маємо 4 супер-ознаки, де кожен клас лежить у своєму чітко відокремленому кластері.

Вплив на обчислювальні потужності (процесор/пам'ять). Для тестування NSL-KDD зазвичай використовують вибірку KDDTrain+ (~125,000 записів) та KDDTest+ (~22,500 записів).

Виділення пам'яті: матриця 22500×41 потребує в 10 разів більше оперативної пам'яті для побудови матриці відстаней (яка використовується, наприклад, в KNN класифікаторах), ніж гібридна матриця 22500×4.

Навантаження на процесор: обчислювальна складність багатьох класифікаторів залежить від кількості ознак *D*. Зменшення *D* з 41 до 4 знижує кількість математичних операцій під час інференсу (перевірки поточного пакета) на 90-95%, що дозволяє інтегрувати таку модель безпосередньо в ядро Linux для фільтрації в реальному часі.

Наведені в таблиці 3 цифрові метрики детально описують поведінку кожного класу атак, проте повну картину математичної взаємодії алгоритмів розкриває саме тренд їхніх змін. Якщо перенести ці табличні дані на систему координат, ми побачимо чітку закономірність: ізольоване застосування методів призводить до падіння ефективності, тоді як їхня послідовна комбінація забезпечує синергетичний стрибок.

Візуальним підтвердженням цього ефекту є рисунок 2, де суцільна лінія відображає динаміку точності (правильне розпізнавання), а пунктирна – рівень хибних спрацьовувань (False Positive Rate). Графік наочно демонструє, як зміна кількості ознак трансформує якість виявлення загроз на кожному етапі.

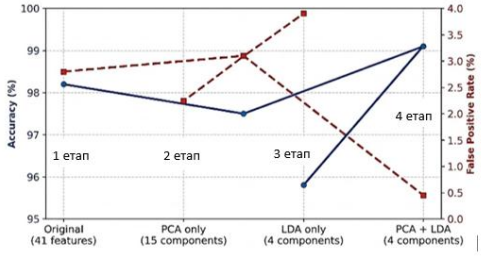

Рис. 2. Метрики застосування PCA+LDA на датасеті NSL-KDD

Детальний аналіз графічних трендів за етапами.

Етап 1. Точка "original (41 features)" – базова лінія системи.

На старті графік фіксує початкову точку, де система використовує всі 41 сирі ознаки. Точність у ~98.2% та рівень хибних тривог у ~2.8% є результатом того, що класифікатор змушений будувати надскладні гіперплощини в розрідженому просторі. Це створює високе навантаження на процесор через обробку великої кількості ознак (*D* = 41), обмежуючи використання моделі в системах реального часу.

Етап 2. Точка "PCA only (15 components)" – ефект «сліпого» стиснення.

При переході до другої точки графік відображає перше розходження ліній у небажані сторони: суцільна лінія точності повзе вниз (до ~97.5%), а пунктирна лінія помилок піднімається вгору (до ~3.1%).

Це графічна ілюстрація головного недоліку навчання без вчителя. Як зазначено в аналізі таблиці, PCA ігнорує мітки класів і фокусується лише на великій дисперсії масових потоків (Normal, DoS). Через це унікальні патерни міноритарних атак (R2L та особливо U2R) стираються, підтверджуючи табличне падіння точності рідкісних атак з 45-50% до 35-40%.

Етап 3. Точка "LDA only (4 components)" – критичний провал (мультиколінеарність).

У третій точці тренди досягають свого критичного екстремуму: суцільна лінія точності падає на локальне «дно» (~95.8%), а пунктирний маркер хибних тривог злітає до свого піка – ~4.5%. Ця графічна аномалія чітко підтверджує проблему лінійної залежності ознак. Хоча LDA є методом із вчителем (supervised) і прагне розділити класи, наявність колінеарності у сирих 41 ознаках робить внутрішні коваріаційні матриці нестабільними. Алгоритм починає «плутатися» в мережевому шумі, що призводить до найгіршого результату захисту на всьому графіку.

Етап 4. Точка "PCA + LDA (4 components)" – синергетичний ефект гібрида.

Фінальний етап графіка демонструє різкий перелом трендів: лінії стрімко розходяться в ідеальні сторони. Суцільна лінія правильного виявлення злітає до свого абсолютного максимуму (~99.1%), а пунктирна лінія FPR падає майже до нуля (~0.45%).

Цей стрибок наочно доводить ефективність гібридної архітектури.

1. PCA працює як фільтр: спочатку він усуває колінеарність та шум, створюючи простір із 15 повністю незалежних компонент.

2. LDA працює як роздільник: отримавши очищені дані, LDA безперешкодно проектує їх у фінальні 4 супер-ознаки. Це повертає моделі здатність чітко бачити навіть міноритарні класи, піднімаючи точність рідкісних атак до рекордних 68-72%.

Зіставлення таблиці та графіка доводить, що послідовність дій має вирішальне значення. Окремо чистий PCA та чистий LDA погіршують первинні метрики. Проте їхній гібрид PCA + LDA забезпечує подвійний виграш:

- для кібербезпеки: зниження кількості хибних тривог у 6 разів при максимальній точності.
- для інфраструктури: зменшення об'єму матриці даних у 10 разів та зниження кількості операцій на 90-95%, що робить модель придатною для фільтрації трафіку в реальному часі на рівні ядра ОС.

Тепер також розглянемо ефективність гібридного підходу PCA+LDA на датасеті CICIDS2017. Для сучасного гетерогенного датасету CICIDS2017 ефект від застосування гібридної схеми PCA + LDA є ще більш вираженим. Оскільки початкова розмірність тут удвічі більша (82 ознаки проти 41 в NSL-KDD), синергія методів дозволяє досягти колосального обчислювального розвантаження за одночасного прориву в точності розпізнавання складних атак.

Нижче наведено зведену таблицю 4 ефективності гібридного підходу для датасету CICIDS2017, сформовану на основі усереднених результатів експериментальних досліджень (із урахуванням 7 основних категорій трафіку: Benign, DoS/DDoS, PortScan, Web Attacks, Infiltration, Botnet, Brute Force).

Таблиця 4.

Порівняльна таблиця ефективності гібридного підходу PCA+LDA для датасету CICIDS2017

Етап / Метрика	CICIDS2017	PCA	LDA	PCA+LDA
Кількість ознак (вхідний вектор)	82	32–35	6	6
Клас "Benign" (нормальний трафік)	~94.5%	~91.1%	~90.2%	~98.9%
Клас "DoS/DDoS" (масові атаки)	~96.8%	~93.4%	~92.0%	~99.4%
Клас "PortScan" (сканування портів)	~92.3%	~89.0%	~87.5%	~97.8%
Веб-атаки та Brute Force	~75.4%	~68.2%	~71.0%	~91.5%
Рідкісні / Цільові атаки	~38-42%	~20-25%	~45-48%	~65-70%
Загальний рівень False Positives	~5.0%	~10.8%	~7.5%	~0.95%
Час навчання моделі	100%	~22%	~12%	~4%

На основі наданих даних у таблиці 4 можна зробити ґрунтовні висновки про математичну та практичну доцільність застосування гібридного методу PCA+LDA для сучасного високовимірного датасету CICIDS2017.

1. Подолання «прокляття вимірності» без втрати якості. Вихідний датасет містить 82 ознаки, що створює надмірне обчислювальне навантаження та містить багато шумних, корельованих метрик потоку (мультиколінеарність). Гібридний підхід стискає вхідний вектор майже в 14 разів (з 82 до 6 компонент). Замість очікуваного падіння точності через втрату інформації, компресія призвела до різкого зростання точності класифікації за всіма категоріями. 6 гібридних компонент виявилися інформативнішими для моделей машинного навчання, ніж 82 сирі ознаки.

2. Ізольоване використання PCA та LDA є неефективне так як виникає криза «сліпого» стиснення (PCA). PCA є алгоритмом навчання без вчителя – він максимізує глобальну дисперсію, ігноруючи мітки класів. Оскільки CICIDS2017 вкрай дисбалансований (переважає легітимний трафік та масові DoS), PCA налаштовує компоненти під них. Унікальні ознаки міноритарних класів алгоритм відсікає як «незначний шум». Це призводить до обвалу точності розпізнавання рідкісних атак (з 38-42% до критичних 20-25%) та подвоєння хибних спрацьовувань (False Positives злітає до 10,8%).

Математичний колапс чистого LDA. LDA працює з вчителем і прагне розділити класи, але він критично чутливий до колінеарності ознак. Коли LDA запускають напряму на 82 сирих ознаках, внутрішні коваріаційні матриці стають математично нестабільними (проблема сингулярності). Алгоритм перенавчається на мережевому шумі, показуючи найгіршу точність розпізнавання базових класів (Benign 90,2%, DoS 92,0%).

3. Синергетичний ефект гібрида (PCA + LDA).

Гібрид працює як ідеальний двоступеневий фільтр, де кожен алгоритм компенсує недоліки іншого:

1. Етап PCA (декореляція): знижує вимірність до 32-35 компонент. Його мета – не класифікація, а повне видалення лінійної залежності між ознаками та очищення від випадкового апаратного шуму.

2. Етап LDA (розділення): отримує на вхід очищені ортогональні (незалежні) компоненти й безперешкодно проектує їх у 6 фінальних векторів, які максимально розсовують усі 7 основних класів трафіку CICIDS2017.

Точність розпізнавання вкрай складних веб-атак та Brute Force зростає до 91,5%, а рідкісних цільових атак – до 65-70% (проти 20-25% у чистого PCA).

4. Критичне зниження хибних тривог. Для реальних систем виявлення вторгнень (IDS) показник хибних ривог є визначальним, адже тисячі хибних тривог паралізують роботу SOC-аналітиків. На сирих даних рівень помилок становить 5,0%, а чистий PCA погіршує його до неприпустимих 10,8%. Гібрид PCA + LDA знижує рівень хибних тривог до 0,95% (покращення порівняно з сирими даними більш ніж у 5 разів). Це свідчить про формування чіткої математичної межі між нормальним трафіком та аномаліями.

5. Оптимізація обчислень для Real-time IDS. Час навчання моделі скорочується до 4% від базового (прискорення у 25 разів). Це дозволяє перенавчати модель кібербезпеки майже «на льоту» при появі нових сигнатур атак. Оскільки класифікатору потрібно обробити всього 6 ознак замість 82, затримка мінімізується, що дає змогу розгортати такі системи на високошвидкісних магістральних комутаторах для аналізу трафіку в реальному часі.

Таблиця доводить, що послідовність PCA+LDA забезпечує максимальну ефективність для складних мережевих датасетів. Підхід повністю вирішує проблему мультиколінеарності сирих даних та ігнорування міноритарних класів, роблячи фінальну модель одночасно легкою (6 ознак), швидкою (прискорення в 25 разів) та максимально точною (~99%).

Аналіз детальних метрик таблиці 4 за окремими категоріями атак дозволяє зрозуміти внутрішню поведінку моделей, проте загальну закономірність та масштаби математичного зсуву ілюструє саме графічний тренд. Перенесення цих табличних значень на координатну площину з подвійною віссю (де ліва шкала відображає загальну точність правильного розпізнавання – Accuracy, а права - рівень хибних тривог – False Positive Rate) унаочнює радикальні коливання ефективності (рис. 3).

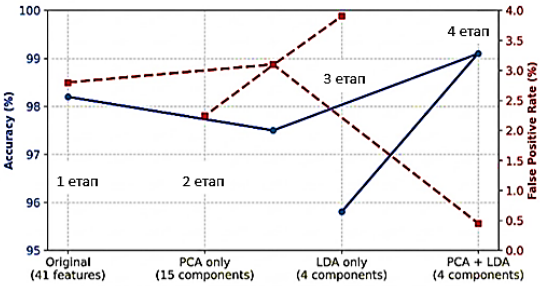

Рис. 2. Метрики застосування PCA+LDA на датасеті CICIDS2017

Якщо для датасету NSL-KDD графік відображав помірні зміни, то для CICIDS2017 через його початкову 82-вимірну гетерогенну структуру графічні криві демонструють набагато різкіші амплітуди, що робить переломи трендів на кожному етапі критично вираженими.

Детальний аналіз графічних трендів для CICIDS2017 представлено 4 етапами.

Етап 1. Точка "Original (82 features)" – вихідний деструктивний баланс. На початковій точці графіка, де модель оперує всіма 82 ознаками, фіксується компромісний старт: точність становить ~94.5%, а рівень хибних тривог (FPR) – ~5.0%. Візуально цей етап відображає обмеженість класичних моделей перед «прокляттям вимірності». Через надлишок дубльованих потокових метрик система генерує забагато хибних спрацьовувань, вимагаючи 100% базового часу обчислень, що є занадто повільним для сучасних магістральних мереж.

Етап 2. Точка "PCA only (35 components)" – графічний пік помилок. При переході до другої точки графік фіксує різке розходження ліній у критично негативні сторони: суцільна лінія точності падає до ~91.1%, а пунктирна лінія помилок стрімко злітає вгору, досягаючи свого максимуму на позначці 10.8%. Це наочна графічна криза навчання без вчителя. Таблиця чітко показує причину цього тренду: оскільки PCA фокусується лише на домінантній дисперсії масових класів (Benign, DoS), він буквально «змиває» унікальні ознаки веб-атак та міноритарних загроз. Графічний зліт лінії FPR до 10,8% віддзеркалює табличне падіння точності рідкісних атак до катастрофічних 20-25%.

Етап 3. Точка "LDA only (6 components)" – локальне дно точності. У третій точці графік відображає унікальну математичну аномалію: суцільна лінія точності пробиває черговий антирекорд і

падає на саме дно – 90,2%, тоді як пунктирна лінія FPR дещо знижується порівняно з попереднім етапом і фіксується на рівні 7,5%. Цей графічний провал ілюструє математичний колапс чистого LDA під впливом мультиколінеарності. Намагаючись розсунути 7 складноорієнтованих класів безпосередньо у вихідному 82-вимірному просторі, алгоритм страждає від сингулярності матриць, втрачає здатність до узагальнення та починає жорстко підлаштовуватися під випадковий мережевий шум.

Етап 4. Точка "PCA + LDA (6 components)" – екстремальний синергетичний злам. Фінальна точка графіка демонструє феноменальний, майже вертикальний злам обох кривих, які розходяться в протилежні, ідеальні для кібербезпеки сторони. Суцільна лінія точності злітає до свого абсолютного піка в ~98.9%, а пунктирна лінія хибних тривог карколомно падає вниз, зупиняючись на мінімальних ~0.95%.

Цей графічний показник повністю підтверджує логіку гібридного підходу.

1. PCA ліквідує хаос: він стискає 82 ознаки до 35 компонент, повністю вичищаючи лінійні залежності та колінеарність.
2. LDA вибудовує межі: отримавши ідеально некорельовані дані, LDA без перешкод проектує їх у 6 фінальних векторів. Це миттєво повертає системі здатність бачити складні та приховані загрози, що підтверджується табличним стрибком точності веб-атак до 91,5%, а рідкісних цільових атак – до 65-70%.

Зіставлення таблиці 4 та графічних кривих наочно доводить: чим складніший і більший датасет, тим яскравіше проявляється синергія гібрида. Послідовний перехід PCA+LDA забезпечує революційний результат, який неможливо отримати методами окремо:

– безпекова перевага: рівень хибних тривог падає нижче критичного порогу в 1% (0.95%), що мінімізує навантаження на аналітиків безпеки (SOC);

– інфраструктурна перевага: об'єм вхідного вектора зменшується у 14 разів (з 82 до 6 ознак), а час навчання падає до 4%. Це знімає обчислювальний пресинг з процесора та пам'яті, дозволяючи інтегрувати навчену модель безпосередньо у високошвидкісні шлюзи для фільтрації мережевих потоків у реальному часі.

ВИСНОВКИ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

1. Розв'язання науково-технічного протиріччя СВВ.

Запропонована методика успішно розв'язує фундаментальне протиріччя сучасних інтелектуальних систем виявлення кібератак: необхідність глибокого багатопараметричного аналізу гетерогенного трафіку проти обмежень обчислювальних ресурсів («прокляття вимірності») та ризику перенавчання моделей. Послідовна комбінація методів лінійної алгебри дозволяє уникнути компромісу між швидкістю роботи та точністю детектування.

2. Математична синергія гібридного підходу PCA+LDA.

Експериментальні дані доводять, що ізольоване застосування алгоритмів зниження вимірності на сирих мережевих метриках є неефективним і руйнівним для метрик безпеки:

- чистий PCA діє в режимі навчання без вчителя і максимізує глобальну дисперсію, через що «змиває» унікальні характеристики малочисельних класів, сприймаючи їх за шум. Це призводить до критичного падіння точності детектування рідкісних атак (до 20-25% у CICIDS2017) та стрибка хибних спрацьовувань (до 10,8%).
- чистий LDA функціонує в режимі навчання з вчителем але повністю колапсує через високу мультиколінеарність сирих даних (82 ознаки в CICIDS2017), що викликає математичну сингулярність матриць розсіювання, перенавчання на мережевому шумі та падіння базової точності детектування (до 90,2%).

Послідовна схема PCA➡LDA діє як двоступеневий фільтр, де PCA виступає математичним регуляризатором (усуває лінійну залежність ознак та денойзить простір), а LDA – геометрично максимізує міжкласову відстань на очищених ортогональних компонентах.

3. Екстремальна компресія та прорив у точності.

Доведено, що для обох еталонних датасетів гібридизація забезпечує радикальне покращення ключових показників ефективності:

- для класичного NSL-KDD: простір стискається в 10 разів (з 41 до 4 компонент), загальна точність зростає до 99.1%, а рівень хибних спрацювань падає в 6 разів – до 0,45%.
- для високовимірного CICIDS2017: вхідний вектор зменшується майже в 14 разів (з 82 до 6 компонент), точність детектування веб-атак злітає до 91,5%, рідкісних/цільових атак – до 65-70%, а загальний рівень хибних тривог падає нижче критичного порогу в 1% – до 0,95%.

4. Інфраструктурна та практична цінність для державного і військового секторів.

Зниження часу навчання моделей до 4% від базового (прискорення у 25 разів для CICIDS2017) та зменшення кількості необхідних математичних операцій на 90-95% дає колосальну практичну перевагу. Перетворення нового пакета трафіку зводиться до каскаду двох простих матричних множень:

$$Z_{final} = (X - \mu) \times W_{PCA} \times W_{LDA}$$

Це дозволяє інтегрувати розроблені математичні моделі безпосередньо в низькорівневі підсистеми ядра ОС Linux або програмовані мережеві карти (SmartNIC, за технологією eBPF/XDP) для фільтрації гігабітних потоків критичної та військової інфраструктури в режимі реального часу на обмежених обчислювальних потужностях.

Напрямки подальших досліджень. Незважаючи на високу ефективність запропонованого лінійного гібридного підходу, динамічний розвиток кіберзагроз, APT-операцій та ускладнення мережевих архітектур визначають наступні перспективні вектори науково-практичних пошуків.

1. Адаптація до концепції концептуального зсуву даних. У реальних мережевих умовах профіль «нормальної» поведінки користувачів та структура легітимного трафіку постійно змінюються в часі. Оскільки класичні матриці трансформацій PCA та LDA є статичними (обчислюються одноразово на етапі навчання), перспективним напрямком є розробка інкрементальних (динамічних) модифікацій алгоритмів – Incremental PCA (IPCA) та Online LDA. Це дозволить системі адаптувати проекційні матриці W_{PCA} та W_{LDA} «на льоту» без повної зупинки СВВ та тривалого перенавчання на гігантських масивах історичних даних.

2. Дослідження нелінійних напівкерованих методів. Оскільки лінійні методи (PCA, LDA) іноді можуть упускати складні нелінійні кореляції вищих порядків між ознаками трафіку, доцільно дослідити застосування легковажних нелінійних архітектур, наприклад, глухих автокодувальників або Kernel PCA. Головна задача тут – знайти компроміс, за якого математична гнучкість нелінійного відображення не порушить жорстких вимог щодо низької затримки під час інференсу в мережах 10 Gbps+.

3. Стійкість до змагальних атак штучного інтелекту. Сучасні кіберзлочинці дедалі частіше використовують методи генеративно-змагальних мереж (GAN) для штучної модифікації пакетів атак (наприклад, деформації часових інтервалів чи розмірів пакетів), щоб обійти інтелектуальні IDS. Подальші дослідження мають бути спрямовані на аналіз вразливості сформованого 4- та 6-вимірного мета-простору ознак до навмисного інжектування змагального шуму з метою підвищення математичної робастності та завадостійкості проекцій.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Buryachok, V. L., Toliupa, S. V., & Semko, V. V. (2016). Informational and cyberspaces: Security problems, methods, and means of counteraction. Nash Format.
2. Zhylin, A. V., Shapoval, O. M., & Uspenskyi, O. A. (2021). Information protection technologies in information and telecommunication systems. Igor Sikorsky Kyiv Polytechnic Institute Publishing House “Politekhnika”.
3. Lukova-Chuiko, N. V., Toliupa, S. V., Nakonechnyi, V. S., & Brailovskyi, M. M. (2021). Intrusion detection systems and functional resilience of distributed information systems against cyber threats. Format.
4. Yevseiev, S. P., Zakovorotnyi, O. Y., Milov, O. V., Kuchuk, H. A., Haluza, O. A., Koval, M. V., Voitko, O. V., & Hryshchuk, R. V. (2024). Methodology for synthesizing models of intelligent management and security systems for critical infrastructure objects. Novyi Svit-2000.
5. Kostiuk, Y. V., Skladannyi, P. M., Bebeshko, B. T., Khorolska, K. V., Rzaieva, S. L., & Vorokhob, M. V. (2025). Security of information and communication systems. Borys Grinchenko Kyiv Metropolitan University.
6. Lande, D. V., Subach, I. Y., & Boiarynova, Y. Y. (2018). Fundamentals of theory and practice of intelligent data analysis in cybersecurity. Institute of Special Communication and Information Protection, Igor Sikorsky Kyiv Polytechnic Institute.
7. Bajaj, K., & Arora, A. (2013). Dimension reduction in intrusion detection features using discriminative machine learning approach. IJCSI International Journal of Computer Science Issues, 10, 324-328.
8. Zhang, F., & Wang, D. (2013). An effective feature selection approach for network intrusion detection. In 2013 IEEE Eighth International Conference on Networking, Architecture and Storage (pp. 307-311).
9. Wahba, Y., Elsalamouny, E., & Eltaweel, G. (2015). Improving the performance of multi-class intrusion detection systems using feature reduction. International Journal of Computer Science Issues, 12(3), p.355.

10. Tesfahun, A., & Bhaskari, D. L. (2013). Intrusion detection using random forests classifier with SMOTE and feature reduction. In 2013 International Conference on Cloud & Ubiquitous Computing & Emerging Technologies (pp. 127-132). IEEE.
11. Dhafian, B., Ahmad, I., & Al-Ghamid, A. (2015). An overview of the current classification techniques in intrusion detection. In Proceedings of the International Conference on Security and Management (p. 82).
12. Desale, K. S., & Ade, R. (2015). Genetic algorithm-based feature selection approach for effective intrusion detection system. In 2015 International Conference on Computer Communication and Informatics (pp. 1-6). IEEE.
13. Zargari, S., & Voorhris, D. (2012). Feature selection in the corrected KDD dataset. In 2012 International Conference on Emerging Intelligent Data and Web Technologies (pp. 174-180). IEEE.
14. Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: Techniques, datasets and challenges. Cybersecurity, 2(1). https://doi.org/10.1186/s42400-019-0038-7
15. Toliupa, S., & Kulko, A. (2026). Methodology of comprehensive feature optimization for cyberattack detection systems. Cybersecurity: Education, Science, Technique, 4(32), 1015-1034. https://doi.org/10.28925/2663-4023.2026.32.1204
16. Goldschmidt, P., & Chudá, D. (2025). Network intrusion datasets: A survey, limitations, and recommendations. Computers & Security, 104510. https://doi.org/10.1016/j.cose.2025.104510
17. Ibrahimi, K., & Ouaddane, M. (2017). Management of intrusion detection systems based on KDD99: Analysis with LDA and PCA. In 2017 International Conference on Wireless Networks and Mobile Communications (WINCOM). IEEE.
18. Panigrahi, R., & Borah, S. (2018). A detailed analysis of the CICIDS2017 dataset for designing intrusion detection systems. International Journal of Engineering & Technology, 7(3), 479-482.
19. Parizad, & Hatziadoniu, C. J. (2022). Cyber-attack detection using principal component analysis and noisy clustering algorithms: A collaborative machine learning-based framework. IEEE Transactions on Smart Grid, 13(6), 4848-4861.
20. More, P., & Mishra, P. (2020). Enhanced PCA-based dimensionality reduction and feature selection for real-time network threat detection. Engineering, Technology & Applied Science Research, 10(5), p.6270.
21. Solani, S., & Jadav, N. K. (2021). A novel approach to reduce false-negative alarm rate in network-based intrusion detection system using linear discriminant analysis. In G. Ranganathan, J. Chen, & Á. Rocha (Eds.), Inventive Communication and Computational Technologies (Vol. 145, Lecture Notes in Networks and Systems). Springer. https://doi.org/10.1007/978-981-15-7345-3
22. Singh, S., & Silakari, S. (2009). Generalized discriminant analysis algorithm for feature reduction in cyber attack detection system. International Journal of Computer Science and Information Security, 6(1), 173-180.
23. Subba, B., Biswas, S., & Karmakar, S. (2016). Enhancing performance of anomaly-based intrusion detection systems through dimensionality reduction using principal component analysis. In 2016 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS) (pp. 1-6).
24. Abdulhammed, R., Faezipour, M., Musafer, H., & Abuzneid, A. (2019). Efficient network intrusion detection using PCA-based dimensionality reduction of features. In 2019 International Symposium on Networks, Computers and Communications (ISNCC) (pp. 1-6).

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 1 (33), 2026
ISSN 2663 – 4023

Andrii Kulko
Postgraduate Student of the Department of Cybersecurity and Information Protection
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
ORCID:0009-0006-1185-0774
kulko452@gmail.com

Serhii Toliupa
Doctor of Technical Sciences, Professor
Professor of the Department of Cybersecurity and Information Protection
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
ORCID: 0000-0002-1919-9174
serhii.toliupa@knu.ua

A HYBRID METHOD FOR FEATURE DIMENSION REDUCTION IN INTRUSION DETECTION SYSTEMS

Abstract. The paper addresses a pressing scientific and practical challenge: optimizing the feature space of network traffic to enhance the efficiency of intelligent Intrusion Detection Systems (IDS). Under the operational conditions of high-speed heterogeneous networks and stringent computational resource constraints of critical infrastructure facilities, the high dimensionality, sparsity, and multicollinearity of modern data stacks generate the "curse of dimensionality" effect. This leads to the overfitting of classification models and significant inference delays.

To overcome these limitations, a linear hybrid feature reduction methodology, PCA+LDA, is proposed, functioning as a two-stage pipelined filter. In the first stage, Principal Component Analysis (PCA) operates in an unsupervised mode, providing denoising, eliminating linear dependencies, and performing primary data compression without losing informative variance. In the second stage, supervised Linear Discriminant Analysis (LDA) projects the feature vector onto a subspace that maximizes the geometric separability between anomaly classes and legitimate traffic based on inter-class and intra-class variance.

The validation of the methodology was carried out on the representative benchmark datasets NSL-KDD and CICIDS2017. It is mathematically proven that the hybrid approach provides extreme compression of the input feature vector (by a factor of 10 for NSL-KDD and nearly 14 for CICIDS2017), accompanied by an increase in the recognition accuracy of minority threats (U2R, R2L) and complex modern web attacks. The overall False Positive rate was reduced to record lows of ~0.45% and ~0.95%, respectively.

Due to the low computational complexity and the matrix nature of linear inference, the training time of classifiers was reduced to 4% of the baseline level. This makes the proposed approach suitable for integration into highly loaded nodes for real-time gigabit traffic filtering directly at the operating system kernel level or based on programmable network interface cards (SmartNIC).

Directions for future research have been identified, including the development of mechanisms to adapt the architecture to Concept Drift in dynamic network environments, as well as increasing model robustness against Adversarial Machine Learning attacks.

Keywords: intrusion detection systems, dimensionality reduction, Principal Component Analysis, Linear Discriminant Analysis, multicollinearity, NSL-KDD, CICIDS2017, cybersecurity, True Positive, False Positive.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Buryachok, V. L., Toliupa, S. V., & Semko, V. V. (2016). Informational and cyberspaces: Security problems, methods, and means of counteraction. Nash Format.
2. Zhylin, A. V., Shapoval, O. M., & Uspenskyi, O. A. (2021). Information protection technologies in information and telecommunication systems. Igor Sikorsky Kyiv Polytechnic Institute Publishing House “Politekhnika”.
3. Lukova-Chuiko, N. V., Toliupa, S. V., Nakonechnyi, V. S., & Brailovskyi, M. M. (2021). Intrusion detection systems and functional resilience of distributed information systems against cyber threats. Format.
4. Yevseiev, S. P., Zakovorotnyi, O. Y., Milov, O. V., Kuchuk, H. A., Haluza, O. A., Koval, M. V., Voitko, O. V., & Hryshchuk, R. V. (2024). Methodology for synthesizing models of intelligent management and security systems for critical infrastructure objects. Novyi Svit-2000.

5. Kostiuk, Y. V., Skladannyi, P. M., Bebeshko, B. T., Khorolska, K. V., Rzaieva, S. L., & Vorokhob, M. V. (2025). Security of information and communication systems. Borys Grinchenko Kyiv Metropolitan University.
6. Lande, D. V., Subach, I. Y., & Boiarynova, Y. Y. (2018). Fundamentals of theory and practice of intelligent data analysis in cybersecurity. Institute of Special Communication and Information Protection, Igor Sikorsky Kyiv Polytechnic Institute.
7. Bajaj, K., & Arora, A. (2013). Dimension reduction in intrusion detection features using discriminative machine learning approach. IJCSI International Journal of Computer Science Issues, 10, 324-328.
8. Zhang, F., & Wang, D. (2013). An effective feature selection approach for network intrusion detection. In 2013 IEEE Eighth International Conference on Networking, Architecture and Storage (pp. 307-311).
9. Wahba, Y., Elsalamouny, E., & Eltaweel, G. (2015). Improving the performance of multi-class intrusion detection systems using feature reduction. International Journal of Computer Science Issues, 12(3), p.355.
10. Tesfahun, A., & Bhaskari, D. L. (2013). Intrusion detection using random forests classifier with SMOTE and feature reduction. In 2013 International Conference on Cloud & Ubiquitous Computing & Emerging Technologies (pp. 127-132). IEEE.
11. Dhafian, B., Ahmad, I., & Al-Ghamid, A. (2015). An overview of the current classification techniques in intrusion detection. In Proceedings of the International Conference on Security and Management (p. 82).
12. Desale, K. S., & Ade, R. (2015). Genetic algorithm-based feature selection approach for effective intrusion detection system. In 2015 International Conference on Computer Communication and Informatics (pp. 1-6). IEEE.
13. Zargari, S., & Voorhris, D. (2012). Feature selection in the corrected KDD dataset. In 2012 International Conference on Emerging Intelligent Data and Web Technologies (pp. 174-180). IEEE.
14. Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: Techniques, datasets and challenges. Cybersecurity, 2(1). https://doi.org/10.1186/s42400-019-0038-7
15. Toliupa, S., & Kulko, A. (2026). Methodology of comprehensive feature optimization for cyberattack detection systems. Cybersecurity: Education, Science, Technique, 4(32), 1015-1034. https://doi.org/10.28925/2663-4023.2026.32.1204
16. Goldschmidt, P., & Chudá, D. (2025). Network intrusion datasets: A survey, limitations, and recommendations. Computers & Security, 104510. https://doi.org/10.1016/j.cose.2025.104510
17. Ibrahimi, K., & Ouaddane, M. (2017). Management of intrusion detection systems based on KDD99: Analysis with LDA and PCA. In 2017 International Conference on Wireless Networks and Mobile Communications (WINCOM). IEEE.
18. Panigrahi, R., & Borah, S. (2018). A detailed analysis of the CICIDS2017 dataset for designing intrusion detection systems. International Journal of Engineering & Technology, 7(3), 479-482.
19. Parizad, & Hatziadoniu, C. J. (2022). Cyber-attack detection using principal component analysis and noisy clustering algorithms: A collaborative machine learning-based framework. IEEE Transactions on Smart Grid, 13(6), 4848-4861.
20. More, P., & Mishra, P. (2020). Enhanced PCA-based dimensionality reduction and feature selection for real-time network threat detection. Engineering, Technology & Applied Science Research, 10(5), p.6270.
21. Solani, S., & Jadav, N. K. (2021). A novel approach to reduce false-negative alarm rate in network-based intrusion detection system using linear discriminant analysis. In G. Ranganathan, J. Chen, & Á. Rocha (Eds.), Inventive Communication and Computational Technologies (Vol. 145, Lecture Notes in Networks and Systems). Springer. https://doi.org/10.1007/978-981-15-7345-3
22. Singh, S., & Silakari, S. (2009). Generalized discriminant analysis algorithm for feature reduction in cyber attack detection system. International Journal of Computer Science and Information Security, 6(1), 173-180.
23. Subba, B., Biswas, S., & Karmakar, S. (2016). Enhancing performance of anomaly-based intrusion detection systems through dimensionality reduction using principal component analysis. In 2016 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS) (pp. 1-6).
24. Abdulhammed, R., Faezipour, M., Musafer, H., & Abuzneid, A. (2019). Efficient network intrusion detection using PCA-based dimensionality reduction of features. In 2019 International Symposium on Networks, Computers and Communications (ISNCC) (pp. 1-6).

Отримано редакцією журналу / Received: 28.02.26
Прорецензовано / Revised: 03.03.26
Схвалено до друку / Accepted: 25.06.26