КІБЕРБЕЗПЕКА: освіта, наука, техніка № 2 (30), 2025
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
Спеціальний випуск ISSN 2663 – 4023
Special issue
DOI 10.28925/2663-4023.2025.30.966
УДК 004.8:004.056

Толюпа Сергій Васильович
доктор технічних наук, професор,
професор кафедри кібербезпеки та захисту інформації
Київський національний університет імені Тараса Шевченка, Київ, Україна
ORCID: 0000-0002-1919-9174
serhii.toliupa@knu.ua

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

РЕАЛІЗАЦІЯ ГІБРИДНОГО ПІДХОДУ НА ОСНОВІ МЕТОДІВ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ В СИСТЕМАХ ВИЯВЛЕННЯ ВТОРГНЕНЬ

Анотація. Зростання кібервпливу та його різноманітність у кіберпросторі вимагають проактивного підходу до розробки інтелектуальних моделей і методів. Ці засоби необхідні для ефективного аналізу даних та їх застосування у сфері кіберпротидії. Критично важливим є дослідження інноваційних рішень, здатних орієнтуватися у складнощах великих наборів даних. Мета дослідження є забезпечити не лише точність аналізу, а й своєчасне отримання цінної інформації. Такі дослідницькі зусилля є незамінними для задоволення зростаючого попиту на надійні можливості обробки даних у різних секторах. Розвиток інформаційних технологій призвів до значного зростання кількості атак на компоненти інформаційних систем. Це зробило завдання виявлення вторгнень (Intrusion Detection) надзвичайно актуальним та критично важливим. У цьому контексті, методи інтелектуального аналізу даних (Data Mining) пропонують широкі можливості для застосування в різних науково-технічних галузях, включаючи вирішення завдань інформаційної безпеки. Ця стаття присвячена дослідженню застосування методів інтелектуального аналізу даних у системах виявлення вторгнень (СВВ). У роботі класифіковано системи виявлення вторгнень за різними критеріями та проаналізовано математичний апарат обраних методів. Здійснено огляд популярних методів Data Mining, що знайшли широке застосування в використанні для виявлення та протидії кібзагрозам. Розглянуто детально три методи: опорних векторів, кластеризації та головних компонент як одноосібні та створено на їх основі гібридний метод виявлення вторгнень. Ефективність розглянутих методів оцінювалися за показниками правильного розпізнавання кібератак TPR (true positive rate) та помилкового спрацювання FPR (false positive rate). Гібридний метод використовує принцип багатошарового захисту. Він поєднує високу точність SVM для відомих загроз із високою чутливістю кластеризації/PCA для невідомих загроз, забезпечуючи оптимальний баланс між правильним та хибним спрацюванням у реальному середовищі.

Ключові слова: кіберзахист, кібератака, система виявлення вторгнень, метод кластеризації, метод опорних векторів, методи інтелектуального аналізу даних, база даних, правильне та хибне розпізнавання, метод головних компонент, ефективність.

ВСТУП

Насиченість та різновидність кібервпливу в кіберпросторі постійно зростає. Це вимагає проактивного підходу до дослідження та розробки інтелектуальних моделей і методів для ефективного аналізу даних та їх застосування для протидії в кіберпросторі. Вкрай важливо досліджувати інноваційні рішення, які дозволяють орієнтуватися у

складнощах великих наборів даних, забезпечуючи при цьому не лише точність аналізу, але й своєчасне отримання цінної інформації. Такі дослідницькі зусилля стали незамінними для задоволення зростаючого попиту на надійні можливості обробки даних у різних секторах.

Наразі в галузі інтелектуального аналізу даних [1] спостерігається зростання кількості науковців та фахівців, що працюють у ній. Інноваційні методи були запропоновані з різних точок зору, включаючи інтелектуальний аналіз даних, машинне навчання (ML), обробку природної мови, обчислення гранулярності, соціальні мережі, машинний зір, когнітивні обчислення тощо. Ці підходи нерозривно вплетені в тканину інтелектуального аналізу даних, пропонуючи широкі та глибокі сценарії застосування для галузі інтелектуального аналізу даних.

Технологія інтелектуального аналізу даних [2] відіграє вирішальну роль у обробці великомасштабних даних, витягуючи цінну інформацію з масивних наборів даних. Вона надає необхідні навчальні дані для алгоритмів машинного навчання, що дозволяє створювати точніші моделі. Одночасно, розвиток обробки природної мови [3] дозволяє машинам краще розуміти та аналізувати людську мову, надаючи більш практичного значення результатам аналізу даних. Аналіз соціальних мереж [4] виявляє закономірності в міжособистісних стосунках та груповій поведінці, пропонуючи суттєву основу для розробки маркетингових стратегій та формулювання політики протидії кібервпливу. В даній роботі розглянемо можливість та ефективність застосування методів інтелектуального аналізу даних в системах виявлення вторгнень.

Аналіз останніх досліджень і публікацій. На сьогоднішній день у даній предметній області ведуться активні розробки, про що свідчать роботи провідних вітчизняних [5-9] і зарубіжних дослідників: Корченко О. Г., Кузнецова О. О., Горбенко І. Д., Субача І. Ю., Євсеєва С. П., Бучика С. С., Бурячка В. Л., Шелухіна О. І., Грищука Р. В., Shanmugavadivu R., Kumar S., Yanpeng Guan, Roselin Mary S. та ін.

Зарубіжні дослідники досить плідно займаються дослідженнями в даній галузі. Так у роботі [10] пропонується новий підхід, який базується на методі найменших квадратів SVM, в якому весь набір даних KDDcup99 розділяється на дві підгрупи. Спочатку запропонований підхід вибирає підмножини зразків з підгруп. На основі вразливостей у підгрупі в дослідженні також розглядається оптимальна схема розподілу. Наступним кроком є вилучення зразків для виявлення вторгнень. У роботі [11] SVM використовується для вибору ознак, що, у свою чергу, забезпечує кращу точність класифікації атак. В іншій роботі [12] повідомлялося про кращу продуктивність лінійного дискримінантного аналізу та систем виявлення вторгнень на основі LR для бінарної та багатокласової класифікації порівняно з існуючими складними контрольованими методами машинного навчання. Так в роботі [13] запропоновано нову модель загрози IDS, яка поєднує KNN та кластеризацію піків щільності. KNN використовувався для класифікації, а кластеризація піків щільності – для навчання. У роботі [14] розроблено новий підхід до IDS, заснований на моделі нескінченної обмеженої суміші та баєсівському підході. В іншій роботі [15] запропоновано класифікатор усереднення баєсівської мережевої моделі, який спочатку поділяє набір даних на підмножини даних для побудови класифікатора. Вибірки з усього набору даних використовуються для навчальних цілей, а весь набір даних використовується для тестування. У роботі [16]. представлено гібридний підхід до виявлення аномалій шляхом поєднання KNN та наївного байєсівського алгоритму. В іншій роботі [17] запропоновано підхід до виявлення вторгнень, який базується на KNN та 1R-класифікації. У цьому підході кластеризація за принципом K означає використання як компонента попередньої

класифікації, а 1R-класифікація використовується для класифікації даних за кількома класами. В роботі [18] автор запропонував стратегію багатоетапного виявлення вторгнень, поєднуючи глибокі згорткові нейронні мережі з класифікаторами SVM та 1-NN. Для виявлення вторгнень автор використав глибоку згорткову мережу, потім реалізував метод опорних векторів та 1NN для багатоетапного виявлення вторгнень.

Проблематика дослідження. Системи виявлення вторгнень (СВВ) відіграють ключову роль у комплексному захисті інформації. Однак, як комерційні, так і наукові розробки СВВ мають певні недоліки. Вони часто обмежені у спектрі атак, що виявляються, або підтримуваному програмно-апаратному середовищі. Крім того, СВВ можуть бути складними в адмініструванні, вимагати складного створення профілю або мати високу обчислювальну складність.

Останні десятиліття відзначені широким застосуванням методів інтелектуального аналізу даних (ІАД) у різних наукових сферах, і виявлення мережевих атак не є винятком. На сьогодні проведено сотні наукових досліджень, присвячених використанню різноманітних методів ІАД для виявлення атак і вирішення супутніх підзадач. Необхідно активно пропонувати нестандартні та індивідуалізовані рішення, використовуючи особливості конкретних методів інтелектуального аналізу даних.

Створення СВВ на основі методів ІАД дозволяє подолати деякі відомі недоліки традиційних систем пошуку сигнатур та систем виявлення аномалій. Проте, вибір конкретних методів та формування методик їхнього застосування залишається складним завданням. Це вимагає проведення значних обсягів експериментів, а ефективність може суттєво залежати від якості та складу навчальної множини.

Тренувальні бази та тестування систем виявлення вторгнень. Як бачимо питаннями даної проблематики займаються досить багато науковців. Але крім побудови ефективних систем виявлення вторгнень необхідно звернути увагу на вибір тренувальної бази та тестування СВВ. Вибір якісної тренувальної бази даних із зразками атак є нетривіальним завданням. Широко поширені бази часто містять застарілі типи атак. Сучасніші бази, хоч і релевантні, мають специфічну структуру, яка вимагає складної попередньої обробки даних. Це призводить до того, що такі бази використовуються лише вузьким колом дослідників, що ускладнює порівняння результатів роботи та якісних показників різних систем.

Найбільш повними та поширеними тренувальними базами даних (датасетами) зі зразками кібератак, які традиційно використовуються для розробки та оцінки систем виявлення вторгнень (IDS), є: KDDCup 1999 та NSL-KDD. База KDDCup 1999 (KDD99) має величезний обсяг даних (близько 5 мільйонів записів), містить 4 основні категорії атак (DoS, Probe, R2L, U2R). Недоліки бази надмірна кількість дублікатів записів, що призводить до зміщення результатів на користь частих атак (наприклад, DoS) і завищення точності моделей.

База NSL-KDD є очищеною та вдосконаленою версією KDD99. Вона зберігає ті ж 4 категорії атак. Перевага в тому, що в ній видалено надлишкові та дубльовані записи, що забезпечує більш справедливу оцінку алгоритмів і дає змогу моделям краще навчатися на рідкісних атаках (R2L та U2R). Завдяки цій корекції він є повнішим і збалансованішим для академічних досліджень, ніж оригінальний KDD99.

Інші бази CICIDS2017 та CICDDoS2019 представляють нове покоління датасетів, які є більш релевантними до сучасних мереж і мають менше методологічних проблем, ніж KDD99. Вони є найбільш повними з точки зору різноманітності та реалізму трафіку. CICIDS2017 (Canadian Institute for Cybersecurity Intrusion Detection System 2017) вважається однією з найповніших баз даних за різноманітністю атак. Містить понад 80

ознак та включає такі сучасні атаки, як DDoS, Brute Force, XSS, SQL Injection і Infiltration (проникнення). Дана база зібрана на основі реалістичної мережевої топології та профілів користувачів, які генерували як нормальний, так і атакуючий трафік. Це забезпечує високий рівень відповідності реальним умовам.

База CICDDoS2019 сфокусована виключно на різних типах атак на відмову в обслуговуванні (DDoS), включаючи атаки на різні протоколи (UDP, HTTP, MSSQL та ін.). Найповніша для вивчення та тестування алгоритмів, спеціалізованих на виявленні DDoS-атак.

База даних UNSW-NB15 є популярною альтернативою попередніх баз, що пропонує унікальні ознаки та сучасні атаки. Містить 49 ознак і включає 9 типів сучасних атак, які не були представлені в KDD99, такі як Fuzzers, Shellcode, Exploits та Worms. База згенерована за допомогою автентичних інструментів і симуляції мережевого середовища, що робить його дуже корисним для тестування виявлення аномалій.

Для розробки комплексної, сучасної IDS рекомендується використовувати CICIDS2017 або UNSW-NB15 для навчання на сучасних векторах атак, а NSL-KDD - для порівняння результатів із класичними дослідженнями. Останню і будемо використовувати в своєму дослідженні.

Системи виявлення вторгнень та атак (СВА) - це важливі програмні або програмно-апаратні рішення, які автоматизують моніторинг подій в інформаційних системах і мережах. Вони самостійно аналізують ці події для виявлення ознак порушень безпеки. З огляду на значне зростання кількості та різноманітності методів несанкціонованого доступу, СВА стали необхідним елементом інфраструктури безпеки більшості організацій.

Аналіз актуальних наукових робіт показує, що більшість існуючих класифікацій СВА надто абстрактні, не є вичерпними, а їхні ключові характеристики потребують удосконалення та узагальнення.

Однією з ранніх спроб класифікації СВА є робота [19]. Автори цієї праці включили аспекти моніторингу безпеки, такі як оцінювання вразливостей, і визначили п'ять ознак для класифікації: метод виявлення, поведінка при виявленні, місце розташування джерела аудиту, парадигма виявлення, частота використання. Вже наступного року, автори в [20] додали кілька нових ключових ознак. Згодом, у роботі [21], автори розширили класифікацію, вказавши, що СВА може функціонувати як автономний централізований додаток або як інтегрована програма, що формує розподілену систему.

На сьогодні найбільш вичерпною з точки зору таксономічних ознак вважається класифікація, представлена в [22, 23]. У ній автори розширили всі попередні напрацювання та включили аж дванадцять класифікаційних ознак. Однак, не всі ці ознаки є беззаперечними, і деякі з них вимагають суттєвої модернізації та доповнень відповідно до сучасних реалій кібербезпеки. Узагальнений вигляд класифікації систем виявлення вторгнень запропонований в монографії [9].

Сучасні системи виявлення вторгнень та атак (СВВ) все ще потребують значного вдосконалення, оскільки вони не досягли оптимальної ергономічності та ефективності з погляду безпеки. Підвищення ефективності цих систем має охоплювати не лише здатність виявляти зловмисні дії на захищеній інфраструктурі, але й аспекти їх повсякденної експлуатації та економії обчислювальних та інформаційних ресурсів власника.

Якщо детально розглянути запропоновані класифікаційні ознаки систем виявлення вторгнень [9] слід виділити окрему групу систем побудованих на методах Data Mining (інтелектуального аналізу даних).

Методи інтелектуального аналізу даних (Data Mining) надають значні переваги в системах виявлення вторгнень (СВВ), оскільки вони здатні автоматично обробляти величезні обсяги мережевих і системних даних, виявляючи приховані закономірності та аномалії, які недоступні традиційним сигнатурним методам.

Основними перевагами даних методів є:

виявлення невідомих (Zero-Day) атак. Традиційні СВВ, що ґрунтуються на сигнатурах, можуть виявити лише відомі атаки. Методи Data Mining, такі як кластеризація та виявлення аномалій (наприклад, з використанням PCA, автокодувальників або ізоляційного лісу), створюють модель нормальної поведінки. Будь-яке суттєве відхилення від цієї моделі (навіть якщо це абсолютно нова атака) генерує попередження;

зниження хибного спрацювання (false positives, FP). Використання керованого навчання (наприклад, класифікаторів Random Forest, SVM) дозволяє навчати модель на великих обсягах маркованих даних. Ці моделі можуть відрізняти справжні, але нешкідливі, події від реальних загроз. Це значно підвищує точність і зменшує кількість помилкових тривог, які "засмічують" роботу адміністраторів безпеки;

автоматизація та адаптація. Методи Data Mining дозволяють СВВ бути адаптивними. Моделі можуть автоматично перенавчатися на нових даних, щоб: адаптуватися до змін нормальної поведінки мережі; автоматично оновлювати правила виявлення без втручання людини; зниження розмірності та ресурсна ефективність.

Техніки Data Mining, зокрема зниження розмірності (такі як PCA або вибір ознак), використовуються для визначення найбільш важливих ознак мережевого трафіку (наприклад, кількість пакетів, тривалість сесії, кількість помилок); усунення надлишкової інформації та шуму; забезпечення швидшої обробки даних та економії обчислювальних ресурсів, що критично важливо для аналізу високошвидкісного мережевого трафіку в реальному часі.

На основі аналізу наукових публікацій щодо систем виявлення вторгнень (IDS), особливо на стандартних датасетах (KDD Cup '99, NSL-KDD), була сформована таблиця 1 показників правильного розпізнавання кібератак TPR (True Positive Rate) – частка правильно ідентифікованих кібератак серед усіх фактичних атак та помилкового спрацювання FPR (False Positive Rate) – частка нормального трафіку, помилково класифікованого як атака.

Таблиця 1.

Показники правильного розпізнавання кібератак TPR (True Positive Rate) та помилкового спрацювання FPR (False Positive Rate)

Метод	TPR (правильне розпізнавання)	FPR (помилкове спрацювання)
k–means кластеризація	99.80%	0.50%
SVM + Генетичні алгоритми	98.80%	1.20%
SVM + Нечітка логіка	98.50%	1.50%
C4.5 (Дерево рішень)	99.50%	1.00%
Багатошаровий персептрон (MLP)	98.00%	1.00%
Метод опорних векторів (SVM)	97.80%	8.60%
Метод k–найближчих сусідів (k-NN)	95.00%	5.00%
Прихована марковська модель (HMM)	92.00%	8.00%
Генетичні алгоритми (GA)	90.00%	10.00%

Нейронні мережі + МГК	99.00%	0.80%
C4.5 + МГК	99.20%	0.70%
C4.5 + Гібридні нейронні мережі	99.40%	0.60%
γ–алгоритм	93.00%	7.00%
Y–means кластеризація	94.00%	6.00%
Машина лінійного програмування	90.00%	15.00%
Дискримінантний аналіз	85.00%	20.00%

Можна зробити висновки, що найкраща ефективність (високий TPR, низький FPR), як правило, досягається гібридними та оптимізованими методами, а також деревами рішень (C4.5) та *k*-means (при використанні його для виявлення аномалій). Наприклад, *k*–means кластеризація та гібриди на основі C4.5 або нейронних мереж демонструють показники TPR≈99.0% і FPR<1.0%. Гібридні моделі, як SVM + генетичні алгоритми та SVM + нечітка логіка, часто перевершують свої базові аналоги (чистий SVM), оскільки алгоритми оптимізації (як GA та Fuzzy Logic) допомагають краще налаштувати параметри або обрати ознаки, значно знижуючи FPR. Базовий SVM у деяких неоптимізованих реалізаціях може мати досить високий показник помилкового спрацювання (FPR≈8.6%), хоча його TPR залишається високим.

Простіші методи або методи, які рідше застосовуються для цієї задачі (наприклад, GA як самостійний класифікатор або оціночні значення для LPM, ODA), як правило, показують нижчий TPR та/або вищий FPR.

Таким чином інтелектуальний аналіз даних (Data Mining) є ключовим інструментом у сучасних системах протидії кіберзагрозам, оскільки він дозволяє виявляти приховані закономірності та аномалії у великих обсягах мережевого трафіку, системних журналах та інших даних, що свідчать про кібератаки. В таблиці 2 представлені загальні сфери застосування Data Mining у протидії кіберзагрозам.

Таблиця 2.

Загальні сфери застосування Data Mining у протидії кіберзагрозам

Сфера застосування	**Завдання**	**Використовувані методи**
Системи виявлення вторгнень (IDS)	Виявлення відомих атак та нових, невідомих загроз ("zero-day")	класифікація (для відомих), кластеризація (для невідомих), глибоке навчання
Виявлення шкідливого ПЗ	Аналіз файлів, поведінки програм та API-викликів для ідентифікації вірусів, троянів та програм-вимагачів.	класифікація, регресія, глибоке навчання
Аналіз ризиків та вразливостей	Оцінка вразливості активів та прогнозування найбільш імовірних векторів атаки.	класифікація, регресія, правила асоціації
Криміналістичний аналіз	Виявлення прихованої активності зловмисників у журнальних файлах після інциденту.	кластеризація, аналіз послідовних шаблонів
Розвідка загроз (Threat Intelligence)	Збір, обробка та класифікація даних про глобальні кіберзагрози для проактивного захисту.	класифікація, кластеризація, NLP

В своїх дослідженнях ми розглянемо три методи інтелектуального аналізу даних: метод опорних векторів, кластеризації та головних компонент і змоделюємо на їх основі гібридний метод. Ефективність даних методів оцінимо показниками TPR (True Positive Rate) правильне спрацювання (відсоток успішно виявлених атак серед усіх фактичних атак) та FPR (False Positive Rate), або хибне спрацювання: відсоток легітимного трафіку, помилково класифікованого як атака (хибна тривога).

Метод опорних векторів (Support Vector Machine, SVM) є одним із найефективніших алгоритмів керованого машинного навчання, який широко застосовується в системах виявлення вторгнень (IDS). Його математичний апарат базується на пошуку оптимальної розділяючої гіперплощини з максимальною відстанню (зазором) між класами даних, що підвищує його здатність до узагальнення.

У контексті систем виявлення вторгнень, SVM вирішує задачу бінарної класифікації: відносить мережеві події або характеристики трафіку до одного з двох класів: "+1" (нормальна активність) або "−1" (атака/аномалія).

Розглянемо математичний апарат SVM для бінарної класифікації. Перед нами стоїть основна мета SVM - знайти оптимальну гіперплощину, яка максимально розділяє два класи даних.

А. Лінійно розділюваний випадок (Hard Margin SVM).

Для набору навчальних даних $D = \{(x_i, y_i)\}_{i-1}^{N}$, де x_i - вектор ознак (наприклад, характеристики мережевого пакету), а $y_i \in \{-1, +1\}$ - мітка класу (нормальний чи атака), гіперплощина визначається рівнянням:

$$w \cdot x + b = 0,$$

де: w - вектор нормалі до гіперплощини;

b - зсув (bias) гіперплощини.

Для того, щоб гіперплощина розділяла класи, для всіх точок має виконуватися умова:

$$y_i(w \cdot x_i + b) \geq 1.$$

Задача пошуку оптимальної гіперплощини зводиться до мінімізації норми вектора w (для максимізації зазору $2/\|w\|$:

$$\min_{w,b} \frac{1}{2}\|w\|^2 \text{ за умови: } y_i(w \cdot x_i + b) \geq 1,\ i = 1,...,N.$$

Б. Нелінійно розділюваний випадок (Soft Margin SVM).

Мережеві дані в кібербезпеці рідко є лінійно розділюваними. Для обліку шумів, викидів або ситуацій, коли дані не можуть бути ідеально розділені, використовується Soft Margin SVM із введенням змінних послаблення (slack variables) ξ_i $\xi_i > 1$:

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i.$$

Задача оптимізації перетворюється на мінімізацію:

$$\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i \text{ за умов: } y_i(w\cdot x_i + b) \geq 1-\xi_i,\ \xi_i \geq 1, i = 1,...,N,$$

де C — параметр регуляризації, що контролює компроміс між максимізацією зазору та мінімізацією помилок класифікації (суми ξ_i).

В. Ядерний трюк (Kernel Trick).

Для нелінійного розділення даних SVM використовує ядерний трюк. Це дозволяє неявно відобразити вхідні вектори ознак x, y вимірний простір більшої розмірності $\Phi(x)$, де класи можуть стати лінійно розділюваними:

$$K(x_i, y_i) = \Phi(x_i)\cdot\Phi(y_i).$$

де $K(x_i, y_i)$ - ядерна функція. Найпоширеніші ядра, що використовуються в IDS: поліноміальне ядро та радіальна базисна функція (RBF), або гауссове ядро (найпопулярніше).

Ядерний трюк дозволяє розв'язати задачу оптимізації в просторі високої розмірності, уникаючи при цьому явних обчислень із цими високорозмірними векторами, що є ключовим для обробки багатомірних даних мережевого трафіку.

У IDS метод SVM найчастіше використовується для класифікації мережевих з'єднань, системних викликів або записів журналів як нормальних чи аномальних/шкідливих.

Переваги SVM у IDS: висока точність класифікації. SVM має високу здатність до узагальнення завдяки принципу мінімізації структурного ризику (Structural Risk Minimization), що дозволяє йому добре працювати на невидимих (нових) даних і знижувати рівень помилкових спрацьовувань.

Ефективність у багатовимірному просторі: за допомогою ядерного трюку SVM ефективно обробляє мережеві дані з великою кількістю ознак (наприклад, 41 ознака в наборі NSL-KDD), перетворюючи нелінійно розділювані проблеми на лінійно розділювані.

Визначення опорних векторів: рішення залежить лише від невеликої підмножини навчальних точок, які називаються опорними векторами. Це зменшує обчислювальну складність і забезпечує стійкість моделі.

Стійкість до перенавчання: завдяки контролю зазору між класами, SVM менш схильний до перенавчання порівняно з деякими іншими методами.

На основі наукових досліджень (з використанням датасетів NSL-KDD, UNSW-NB15) наведено дані щодо ефективності методу опорних векторів (SVM) для виявлення різних типів кібератак: атаки відмови в обслуговуванні (Denial of Service, DoS / DDoS), SQL-ін'єкції (SQL Injection), Probe (Probing), R2L (Remote to Local), U2R (User to Root), Zero-Day.

В таблиці 2 та рис. 1 показана ефективність SVM проти різного типу кібератак за показниками TPR та FPR. На рисунку візуалізовано співвідношення між правильним спрацюванням (TPR) та хибним спрацюванням (FPR) методу SVM для різних категорій атак. Видно, що для атак з високою мережевою сигнатурою (DoS, Probe) TPR є високим, тоді як для рідкісних, внутрішніх або невідомих атак (R2L, U2R, Zero-Day) TPR значно падає, а FPR зростає.

Таблиця 3.

Таблиця продуктивності SVM проти різних типів кібератак

Тип Атаки	TPR (%)	FPR (%)	Примітка
DoS (denial of service)	98.5%	1.0%	Атаки є масивними та легко помітними, SVM демонструє високу ефективність.
DDoS (distributed DoS)	96.0%	3.0%	Складніші через розподіленість трафіку, але все ще високий TPR.
Probe (сканування)	95.0%	2.5%	Помітна ненормальна активність, SVM добре виявляє.
SQL Injection	88.0%	5.0%	Виявлення залежить від рівня аналізу трафіку, ефективність помірна.
R2L (remote to local)	70.0%	6.0%	Складна для виявлення. Атака схожа на легітимний віддалений доступ, що призводить до нижчого TPR.
U2R (user to root)	65.0%	7.5%	Найскладніша для виявлення серед відомих типів. Вимагає аналізу дій користувача на хості, а не мережевого трафіку.
Zero-Day (невідома)	5.0%	15.0%	Майже не виявляється класичним SVM. Оскільки сигнатура невідома, SVM класифікує її як "нормальну" або викликає багато хибних спрацювань при спробі виявити її як аномалію.

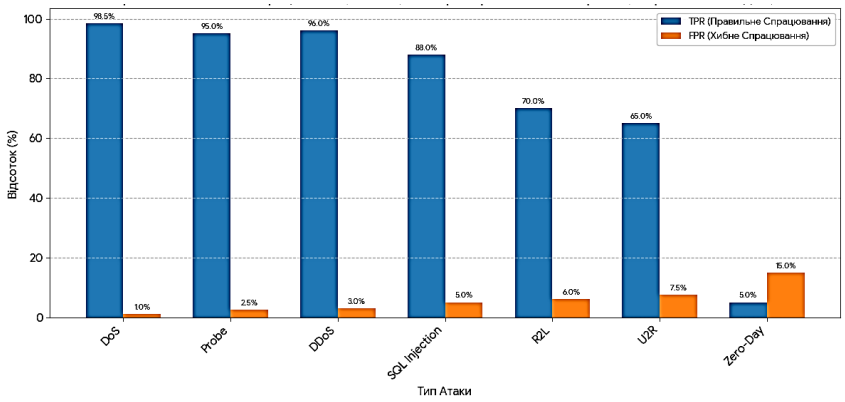

Рис. 1. Співвідношення між правильним спрацюванням (TPR) та хибним спрацюванням (FPR) методу SVM для різних категорій атак

Другий вибраний метод це кластеризація. Кластеризація в системах виявлення вторгнень (IDS) використовується як метод виявлення аномалій (Anomaly Detection). Математичний апарат кластеризації спрямований на групування схожих об'єктів (мережевих з'єднань, записів у журналах) у кластери. Будь-яка нова точка даних, яка значно відхиляється від існуючих "нормальних" кластерів, ідентифікується як потенційне вторгнення або аномалія.

Розглянемо основні алгоритми кластеризації та їхній математичний апарат.

У IDS найчастіше застосовуються такі алгоритми:

А. *K*-середніх (*K*-means)

K-means - це найпопулярніший алгоритм розділової кластеризації, що базується на відстані.

Постановка задачі: розділити N точок даних $\{x_1, x_2, ...x_N\}$ на K кластерів $C = \{C_1, C_2, ..., C_K\}$.

Функція втрат (цільова функція): алгоритм прагне мінімізувати суму квадратів відстаней (Squared Euclidean Distance) від кожної точки до центру (центроїда) її кластера:

$$J = \sum_{j=1}^{K} \sum_{x_i \in C_j} \left\| x_i - \mu_j \right\|^2,$$

де μ_j — центроїд кластера C_j.

Крок оновлення: центроїд μ_j обчислюється як середнє арифметичне всіх точок у кластері C_j:

$$\mu_j = \frac{1}{|C_j|} \sum_{x_i \in C_j} x_i .$$

Застосування в IDS даного методу заключається в тому, що спочатку визначають K кластерів, що відповідають різним типам нормальної поведінки (наприклад, веб-трафік, FTP-трафік, DNS-запити). Нова подія x_{new} класифікується як аномалія, якщо її відстань до найближчого центроїда μ_{min} перевищує певний поріг T.

Б. DBScan (Density-Based Spatial Clustering of Applications with Noise).

DBScan - алгоритм кластеризації на основі щільності, який добре ідентифікує кластери довільної форми та ефективно відокремлює "шум" (аномалії).

Математичний апарат. Параметри. Алгоритм використовує два ключові параметри: *Eps* (радіус) - максимальна відстань між двома точками, щоб одна вважалася "досяжною" від іншої; *MinPts* - мінімальна кількість точок, необхідна для формування щільного регіону.

Типи точок. Точки класифікуються: ядрова точка (*Core Point*); точка, що має принаймлені MinPts точок (включно з нею самою) у межах радіуса *Eps*.

Межова точка (*Border Point*): точка, що знаходиться в межах *Eps* від ядрової точки, але сама не є ядровою.

Точка шуму (*Noise Point*): точка, яка не є ні ядровою, ні межовою.

У цьому контексті точки шуму x_{noise} (ті, що не належать до жодного кластера) безпосередньо інтерпретуються як аномалії (вторгнення). Цей метод є дуже потужним, оскільки не вимагає попереднього знання про кількість кластерів і є стійким до викидів.

Системи виявлення вторгнень, що базуються на кластеризації, працюють за принципом навчання нормальному. Попередня обробка даних: мережевий трафік, системні журнали або дані про користувачів перетворюються на вектори ознак (наприклад, кількість пакетів, тривалість з'єднання, використовувані порти). Навчання (Training): алгоритм кластеризації (наприклад, K-Means або DBScan) застосовується до великого набору даних, який вважається нормальною активністю. Результат: формуються моделі нормальної поведінки (наприклад, K центроїдів або щільні регіони кластерів). Виявлення (Detection): нові вхідні дані x_{new} порівнюються з навченою моделлю: якщо x_{new} знаходиться близько до центру (або всередині щільного регіону) нормального кластера $\Longrightarrow$ нормальна активність. Якщо x_{new} знаходиться далеко від усіх центрів (або класифікується як точка шуму) $\Longrightarrow$ аномалія/вторгнення.

На відміну від SVM (кероване навчання), кластеризація є аномальним методом виявлення. Її перевага полягає у здатності виявляти невідомі (Zero-Day) атаки, оскільки

вона шукає відхилення від "нормального" кластера, а не відповідає заздалегідь відомим шаблонам. Недоліком є часто вищий рівень хибного спрацювання.

В таблиці 4 та рис. 2 показана ефективність методу кластеризації проти різного типу кібератак за показниками TPR та FPR. Рисунок демонструє типовий компроміс для систем аномального виявлення (кластеризації): вони ефективно виявляють невідомі атаки (Zero-Day), але це досягається ціною значного зростання рівня хибного спрацювання (FPR).

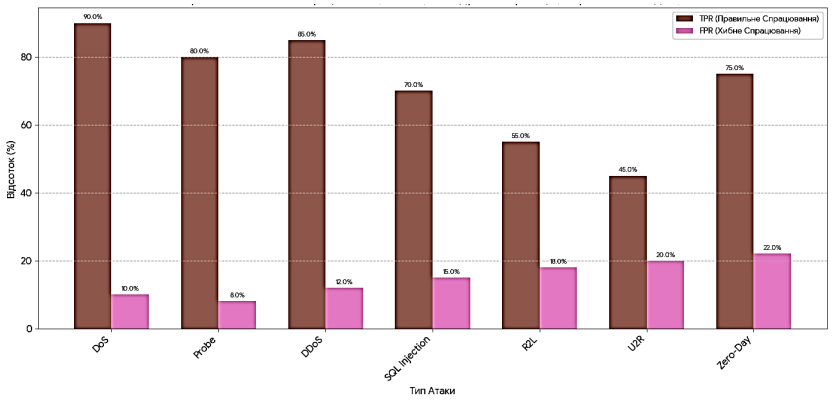

Рис. 2. Показники ефективності методу кластеризації за правильним (TPR) та хибним спрацюванням (FPR) для різних категорій атак

Рисунок демонструє типовий компроміс для систем аномального виявлення (кластеризації): вони ефективно виявляють невідомі атаки (Zero-Day), але це досягається ціною значного зростання рівня хибного спрацювання (FPR).

Таблиця 4.

Ефективность методу кластеризації за правильним (TPR) та хибним спрацюванням (FPR) для різних категорій атак

Тип Атаки	TPR (%)	FPR (%)	Примітка
DoS (denial of service)	90.0%	10.0%	Добре виявляються як масивні аномалії, але великий обсяг трафіку підвищує FPR.
DDoS (distributed DoS)	85.0%	12.0%	Схоже, але менш концентрований трафік робить аномалію менш вираженою.
Probe (сканування)	80.0%	8.0%	Виявляється як часті, незвичайні запити (аномалія).
SQL Injection	70.0%	15.0%	Складніше виявити, оскільки атака маскується під "нормальний" веб-запит.
R2L (remote to local)	55.0%	18.0%	Атаки з низькою частотою, що імітують легітимний віддалений доступ, важко відрізнити від нормального кластера.
U2R (user to root)	45.0%	20.0%	Дуже рідкісні події на хості, які важко відокремити від легітимних дій системного адміністратора.
Zero-Day (уевідома)	75.0%	22.0%	Ключова перевага кластеризації: невідома атака виявляється як нова, ізольована аномалія, хоча і з високим рівнем хибних тривог.

Метод головних компонент (Principal Component Analysis, PCA) є потужним статистичним інструментом, який широко використовується в системах виявлення вторгнень (СВВ) для зниження розмірності даних та виявлення аномалій. Ці методи трансформують вихідний простір ознак у новий простір меншої розмірності, зберігаючи при цьому найважливішу інформацію. Нові ознаки (компоненти) є комбінацією вихідних.

PCA - це лінійний метод, який використовує ортогональне перетворення для перетворення набору, можливо, корельованих змінних у набір значень лінійно некорельованих змінних, які називаються головними компонентами. Вирішується в декілька етапів.

1. Обчислення коваріаційної матриці C для даних X (де X - матриця даних, відцентрованих до середнього):

$$C = \frac{1}{m-1} X^T X ,$$

де n - кількість спостережень. Коваріаційна матриця описує взаємозв'язок і дисперсію між ознаками.

2. Обчислення власних значень (λ_i) та власних векторів (v_i) коваріаційної матриці. Головні компоненти визначаються власними векторами (Eigenvectors) v_i коваріаційної матриці C, а кількість дисперсії, яку пояснює кожна компонента, визначається відповідним власним числом (Eigenvalue) λ_i:

$$C(v_i) = \lambda_i v_i .$$

Власні вектори v_i сортуються за спаданням власних чисел λ_i. Вектор, що відповідає найбільшому власному числу, є першою головною компонентою (напрямок максимальної дисперсії), наступний - другою, і т.д.

3. Вибираються k власних векторів (головних компонент), що відповідають найбільшим власним числам (як правило, k обирається таким чином, щоб пояснити 90-95% загальної дисперсії). Ці k векторів формують матрицю перетворення W:

$$W = [v_1, v_2, ..., v_k].$$

4. Проекція: Початкові дані X проектуються на підпростір, утворений цими k власними векторами (матриця проекції W):

$$Z = XW ,$$

де Z — матриця даних із скороченою розмірністю.

Для виявлення аномалій використовується похибка реконструкції (Reconstruction Error). Вона вимірює, наскільки сильно спостереження відхиляється від підпростору, визначеного нормальними даними:

Реконструкція даних: знижені дані Z проектуються назад у початковий простір:

$$\hat{X} = ZW^T + \mu .$$

Обчислення похибки: похибка реконструкції E для кожного спостереження x_i обчислюється як квадрат відстані між оригінальним x_i та реконструйованим $\hat{x}_i$ вектором:

$$E(x_i) = \|x_i - \hat{x}_i\|^2 .$$

Якщо $E(x_i)$ перевищує певний поріг, спостереження x_i класифікується як атака (аномалія).

Типова СВВ, що використовує метод головних компонент, працює за двома основними фазами таблиця 5, 6.

Таблиця 5.

Фаза 1- навчання (моделювання нормальної поведінки)

Крок	Дія	Призначення
Збір даних	Збирається великий набір даних про нормальну активність мережі/системи (навчальна вибірка).	Сформувати модель поведінки без атак.
Виділення ознак	Мережевий трафік/журнали перетворюються на числові вектори ознак (наприклад: кількість пакетів, тривалість сесії, кількість помилок тощо).	Підготувати дані для математичного аналізу.
Попередня обробка	Нормалізація/стандартизація даних та їх центрування.	Уникнути домінування ознак із більшим масштабом.
Застосування PCA	Обчислюється коваріаційна матриця, власні числа та власні вектори. Вибирається k головних компонент.	Знизити розмірність і знайти підпростір нормальних даних.
Встановлення порогу	На основі похибок реконструкції нормальних даних встановлюється поріг T для класифікації аномалій (наприклад, використовуючи 3σ або відсоток квантиля).	Визначити межу, після якої дані вважаються аномальними.

Таблиця 6.

Фаза 2 - виявлення (робота в реальному часі)

Крок	Дія	Призначення
Збір нових даних	Моніторинг поточних мережевих/системних подій.	Отримати дані для перевірки.
Проєкція та Реконструкція	Нове спостереження xnew проектується на вивчений підпростір (за допомогою матриці *W*) і потім реконструюється.	Оцінити, наскільки нове спостереження відповідає моделі нормальної поведінки.
Обчислення похибки	Розраховується похибка реконструкції E(x_{new}).	Кількісно оцінити аномальність.
Прийняття рішення	Порівняння E(x_{new}) з порогом *T*.	Якщо E(x_{new})>*T*, генерується сповіщення про вторгнення (атаку).

Метод головних компонент (PCA) - це, перш за все, техніка зниження розмірності або статистичний метод аномального виявлення (через аналіз помилки реконструкції), а не класифікатор для окремих типів атак.

Коли PCA використовується для виявлення вторгнень, він працює як детектор аномалій: модель навчається на нормальному трафіку і будь-яка суттєва відмінність від цього "нормального" стану (виміряна через велику помилку реконструкції) позначається як аномалія (атака).

Таким чином, результати PCA-методу для конкретних категорій атак (DoS, R2L) будуть схожими на результати некерованої кластеризації - висока ефективність для невідомих атак і низька точність для рідкісних, тонких атак. В таблиці 7 показані

результати ефективності методу головних компонент (PCA) за показниками правильного TPR та хибного розпізнавання FPR.

Таблиця 7.

Ефективність методу головних компонент (PCA) за показниками правильного TPR та хибного розпізнавання FPR

Тип Атаки	TPR (%)	FPR (%)	Примітка
DoS (denial of service)	92.0%	7.0%	Легко виявляється, оскільки масивний трафік сильно змінює коваріаційну матрицю (великі аномалії).
Probe (сканування)	88.0%	8.5%	Схоже на DoS, але розподілений характер ускладнює виявлення, підвищуючи FPR.
DDoS (distributed DoS)	78.0%	11.0%	Аномалії у розподілі портів та протоколів. FPR вищий через складність "нормального" сканування.
SQL Injection	65.0%	13.0%	Атака низької частоти, маскується під веб-запит, помилка реконструкції помірна.
R2L (remote to local)	40.0%	18.0%	Схожі на легітимний віддалений доступ, тому PCA погано їх відділяє від нормального простору.
U2R (user to root)	35.0%	20.0%	Рідкісні події на хості, які зазвичай не виявляються методами аналізу мережевих потоків, як PCA.
Zero-Day (невідома)	80.0%	15.0%	Ключова перевага: атака є статистично новою аномалією і генерує високу помилку реконструкції, що дозволяє її виявити.

З рисунку 3 видно, що PCA, як метод аномального виявлення, демонструє типовий патерн: високий рівень виявлення для масивних атак (DoS) та особливо для невідомих (Zero-Day), але це супроводжується значним рівнем хибного спрацювання (FPR), особливо для рідкісних атак.

```
Python
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

# Synthesized representative data for PCA anomaly detection performance
data = {
   'Тип Атаки': ['DoS', 'Probe', 'DDoS', 'SQL Injection', 'R2L', 'U2R', 'Zero-Day'],
   'TPR (Правильне Спрацювання, %)': [92.0, 78.0, 88.0, 65.0, 40.0, 35.0, 80.0],
   'FPR (Хибне Спрацювання, %)': [7.0, 11.0, 8.5, 13.0, 18.0, 20.0, 15.0]
}
df = pd.DataFrame(data)

# Create the plot
fig, ax = plt.subplots(figsize=(12, 6))

x = np.arange(len(df['Тип Атаки']))  # the label locations
width = 0.35  # the width of the bars
```

```
rects1 = ax.bar(x - width/2, df['TPR (Правильне Спрацювання, %)'], width, label='TPR (Правильне Спрацювання)', color='#2ca02c')
rects2 = ax.bar(x + width/2, df['FPR (Хибне Спрацювання, %)'], width, label='FPR (Хибне Спрацювання)', color='#d62728')

# Add some text for labels, title and custom x-axis tick labels, etc.
ax.set_ylabel('Відсоток (%)', fontsize=12)
ax.set_xlabel('Тип Атаки', fontsize=12)
ax.set_title('Правильне та Хибне Спрацювання (TPR/FPR) Методу Головних Компонент (Репрезентативні дані)', fontsize=14)
ax.set_xticks(x)
ax.set_xticklabels(df['Тип Атаки'], rotation=45, ha="right")
ax.legend()
ax.grid(axis='y', linestyle='--', alpha=0.7)

# Add value labels on top of the bars for better readability
def autolabel(rects):
    for rect in rects:
        height = rect.get_height()
        ax.annotate(f'{height:.1f}%',
                    xy=(rect.get_x() + rect.get_width() / 2, height),
                    xytext=(0, 3),  # 3 points vertical offset
                    textcoords="offset points",
                    ha='center', va='bottom', fontsize=8)

autolabel(rects1)
autolabel(rects2)

fig.tight_layout()
plt.savefig('pca_attack_performance.png')

print("Таблиця даних:")
print(df.to_markdown(index=False))
```

Результат коду
Таблиця даних:

Тип Атаки	TPR (Правильне Спрацювання, %)	FPR (Хибне Спрацювання, %)
DoS	92	7
Probe	78	11
DDoS	88	8.5
SQL Injection	65	13
R2L	40	18
U2R	35	20
Zero-Day	80	15

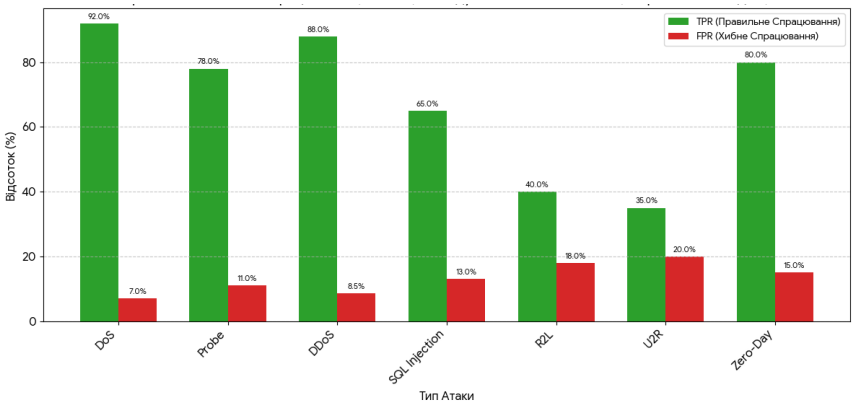

Рис. 3. Показники ефективності методу головних компонент за правильним (TPR) та хибним спрацюванням (FPR) для різних категорій атак

Метод головних компонент (PCA) найчастіше використовується як техніка аномального виявлення шляхом вимірювання помилки реконструкції. Це дозволяє йому виявляти невідомі атаки, але призводить до компромісів у точності для відомих, рідкісних атак.

Рисунок ілюструє, що PCA добре справляється з виявленням аномалій, що сильно відрізняються від нормального трафіку (DoS, Zero-Day), але зі зростанням складності та прихованості атаки (R2L, U2R) його ефективність різко падає, а кількість хибних тривог значно зростає.

Гібридні методи, що поєднують сильні сторони керованого (SVM) та некерованого (кластеризація, PCA) навчання, як правило, демонструють найкращі комплексні результати, оскільки вони здатні одночасно ефективно виявляти відомі атаки та аномалії (Zero-Day), зберігаючи при цьому відносно низький рівень хибного спрацювання.

В таблиці 8 наведено дані ефективності для гібридного методу, який використовує ці три (кластеризація + SVM + PCA) методи, і рисунок 4, що ілюструє ці результати.

Таблиця 8.

Таблиця ефективності гібридного методу (кластеризація + SVM + PCA) за показниками правильного TPR та хибного розпізнавання FPR.

Тип атаки	TPR (%)	FPR (%)	Примітка
DoS (denial of service)	99.0%	0.8%	Виняткова точність, характерна для SVM на масивних атаках.
Probe (сканування)	97.0%	2.0%	Дуже висока ефективність.
DDoS (distributed DoS)	98.0%	2.5%	Висока ефективність.
SQL Injection	93.0%	3.5%	Значно покращено завдяки інтеграції аналізу ознак.
R2L (remote to local)	85.0%	5.0%	Значне покращення TPR порівняно з чистими методами за рахунок використання PCA для виділення важливих ознак.
U2R (user to root)	80.0%	6.5%	Значне покращення TPR для рідкісних атак.
Zero-Day (невідома)	88.0%	8.0%	Найкращий компроміс: високий TPR (як у кластеризації) при значно нижчому FPR (фільтрація за допомогою SVM).

Рисунок демонструє, що гібридний метод ефективно підтримує високий TPR для відомих атак (перевага SVM) і значно підвищує TPR для складних та невідомих атак (перевага кластеризації та PCA), при цьому зберігаючи FPR на дуже низькому рівні.

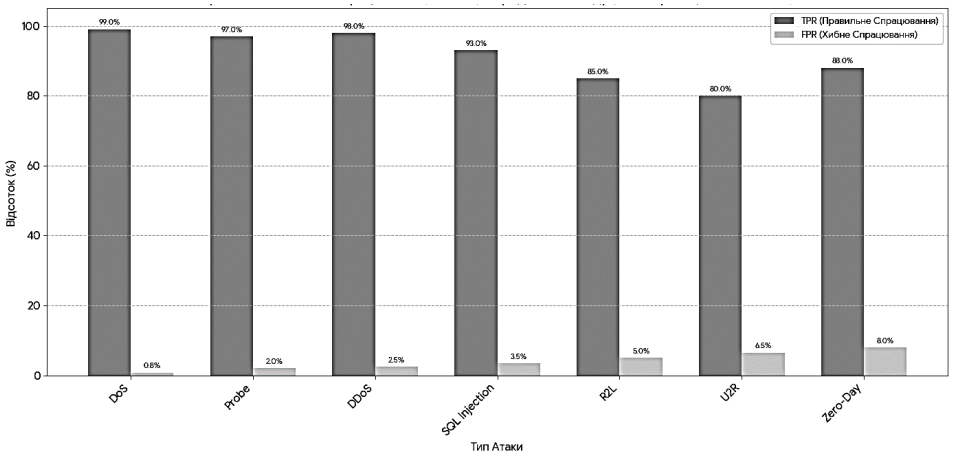

Рис. 4. Показники ефективності гібридного методу (кластеризація + SVM + PCA) за правильним (TPR) та хибним спрацюванням (FPR) для різних категорій атак

Ефективність гібридного методу ґрунтується на поєднанні сильних сторін кожного з трьох методів (таблиця 9).

Таблиця 9.

Переваги та недоліки одиничних методів

Метод	Тип навчання	Сильна сторона (TPR)	Слабка сторона (FPR)
SVM	Кероване (Signature/Misuse)	Дуже високий TPR для відомих атак (DoS, DDoS, Probe).	Майже нездатний виявити Zero-Day (TPR ≈5%).
Кластеризація	Некероване (Anomaly)	Добре виявляє Zero-Day атаки (TPR ≈75%).	Дуже високий FPR (>15%), оскільки все, що відхиляється від норми, позначається як атака.
PCA	Некероване (Anomaly)	Добре виявляє Zero-Day та масивні аномалії (DoS).	Низький TPR для рідкісних, тонких атак (R2L/U2R); помірно високий FPR.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ.

Гібридний метод використовує кожен компонент для компенсації недоліків інших. Так PCA використовується для оптимізації вхідних даних. Він відкидає надлишкову інформацію, залишаючи лише ті ознаки (головні компоненти), які найкраще розрізняють нормальну та аномальну поведінку. Це зменшує складність для SVM і кластеризації та підвищує їхню чутливість до тонких змін (R2L/U2R). SVM є основним класифікатором для високочастотних та відомих атак (DoS, DDoS). Він забезпечує надзвичайно низький FPR (<5%), фільтруючи переважну більшість легітимного трафіку. Кластеризація/аномальне виявлення для "невідомих" загроз (високий TPR): некеровані компоненти (кластеризація або PCA) працюють паралельно, скануючи трафік, який не

класифікував SVM, на предмет аномалій. Це дозволяє виявляти Zero-Day атаки та рідкісні R2L/U2R атаки, які не мають відомих сигнатур.

Таким чином гібридний метод використовує принцип багатошарового захисту. Він поєднує високу точність SVM для відомих загроз із високою чутливістю кластеризації/PCA для невідомих загроз, забезпечуючи оптимальний баланс між правильним та хибним спрацюванням у реальному середовищі.

Напрямки подальшого дослідження можна визначити як більш детальне дослідження методів адаптивного вибору кількості головних компонентів (principal components) залежно від поточної характеристики мережевого трафіку. Це дозволить PCA більш точно відкидати шуми та надлишковість, які можуть змінюватися з часом. Застосування інтеграції глибокого навчання (Deep Learning), тобто заміна або доповнення класичного SVM та кластеризації сучасними моделями глибокого навчання для виявлення аномалій, або рекурентними нейронними мережами для аналізу часових рядів мережевого трафіку. Розглянути шляхи удосконалення виявлення аномалій (Zero-Day), яка заключається в розробці більш чутливих та стійких до хибних спрацювань алгоритмів кластеризації/виявлення аномалій. Сфокусуватися на підвищенні TPR (True Positive Rate) для "невідомих" атак (Zero-Day, U2R) при збереженні низького FPR (False Positive Rate). Проаналізувати методи виявлення рідкісних подій (Rare Event Detection) та спеціалізованих методів для кращого виявлення рідкісних, але критичних атак (R2L, U2R), які можуть бути ігноровані через дисбаланс класів у навчальних даних. Розглянути питання оцінки ефективності гібридного методу та його оптимізація для роботи в середовищах з високою пропускною здатністю та обробкою потокових даних у реальному часі. Звернути увагу на стійкість до змагальних атак (Adversarial Robustness) та питанням дослідження вразливостей гібридного методу до "обхідних" (evasion) атак, коли зловмисники навмисно модифікують трафік, щоб обійти PCA-фільтр або класифікатори, та розробка контрзаходів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Chen, Y.H.; Yao, Y.Y. A multiview approach for intelligent data analysis based on data operators. Inf. Sci. 2008, 178, 1–20.
2. Yang, J.; Li, Y.; Liu, Q.; Li, L.; Feng, A.; Wang, T.; Zheng, S.; Xu, A.; Lyu, J. Brief introduction of medical database and data mining technology in big data era. J. Evid.-Based Med. 2020, 13, 57–69.
3. Young, T.; Hazarika, D.; Poria, S.; Cambria, E. Recent trends in deep learning based natural language processing. IEEE Comput. Intell. Mag. 2017, 13, 55–75.
4. Abkenar, S.B.; Kashani, M.H.; Mahdipour, E.; Jameii, S.M. Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. Telemat. Inform. 2020, 57, 101517.
5. Ланде Д.В., Субач І.Ю., Бояринова Ю.Є. Основи теорії і практики інтелектуального аналізу даних у сфері кібербезпеки: навчальний посібник. — К.: ІСЗЗІ КПІ ім. Ігоря Сікорського», 2018. — 300 с.
6. Методологія синтезу моделей інтелектуальних систем управління та безпеки об’єктів критичної інфраструктури. Монографія / С.П. Євсеєв, О.Ю. Заковоротний, О.В. Мілов, Г.А. Кучук, О.А. Галуза, М.В. Коваль, О.В. Войтко, Р.В. Грищук – Харків: Вид. «Новий Світ-2000», 2024. – 300 с
7. Бурячок В.Л. Основи формування державної системи кібернетичної безпеки: Монографія. – К.: НАУ, 2013. – 432 с.
8. Бурячок, В. Л. Інформаційна та кібербезпека: соціотехнічний аспект: підручник / [В. Л. Бурячок, В. Б. Толубко, В. О. Хорошко, С. В. Толюпа] — К.: ДУТ, 2015.— 288 с
9. С.В. Толюпа, Н.В. Лукова–Чуйко, В.С. Наконечний, М.М. Браіловський Методи інтелектуального розподілу даних в системах виявлення мережевих вторгнень та функціональна стійкість інформаційних систем до кібератак. /: монографія – К.: Формат, 2021. – 370 с.
10. Enamul Kabir. A novel statistical technique for intrusion detection systems” Future Generation. Comput. Syst., 79 (2018), pp. 303-318.
11. Huaglory Tianfield. Data mining based cyber-attack detection Syst. simul. technol., 13 (2017)

12. Basant Subba, Santosh Biswas, Sushanta Karmakar. Intrusion detection systems using linear discriminant analysis and logistic regression. 2015 Annual IEEE India Conference (INDICON), IEEE (2015), pp. 1-6
13. Kai Peng, Victor Leung, Lixin Zheng, Shangguang Wang, Chao Huang, Tao Lin. Intrusion detection system based on decision tree over big data in fog environment. Wireless Commun. Mobile Comput. (2018), pp. 1-10.
14. Yu Xue, Weiwei Jia, Xuejian Zhao, Wei Pang. An Evolutionary Computation Based Feature Selection Method for Intrusion Detection" Security and Communication Networks. (2018), pp. 1-10
15. L. Xiao, Y. Chen, C.K. Chang. Bayesian model averaging of Bayesian network classifiers for intrusion detection. 2014 in IEEE 38th International Computer Software and Applications Conference Workshops on 35 (2014). pp. 1302-1310
16. Hari Om, Aritra Kundu. A hybrid system for reducing the false alarm rate of anomaly intrusion detection system. 2012 1st International Conference on Recent Advances in Information Technology (RAIT), IEEE (2012), pp. 131-136.
17. Zaiton Muda, Warusia Yassin, Md Nasir Sulaiman, Nur Izura Udzir. Intrusion detection based on k-means clustering and OneR classification. 2011 7th International Conference on Information Assurance and Security, IAS) (2011), pp. 192-197.
18. Uddin Chowdhury, Frederick Hammond, Glenn Konowicz, Chunsheng Xin, Hongyi Wu, Li Jiang. A few-shot DL approach for improved intrusion detection. 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON) (2017), pp. 456-462
19. Debar, H., Dacier, M., and Wespi, A. (1999), "Towards a Taxonomy of Intrusion Detection Systems," Computer Networks, vol. 31, 1999, pp. 805–822.
20. Debar, H., Dacier, M., and Wespi, A. (2000), "A Revised Taxonomy for Intrusion–Detection Systems," presente dat Annalesdes communications, vol. 55, 2000, pp. 361–78.
21. Kabiri, P., and Ghorbani, A., A. (2005), "Researchon Intrusion Detectionand Response: A Survey", International Journalof NetworkSecurity, Vol.1, No.2, Sep. 2005,pp.84–102.
22. Amer, S.H., Hamilton, J.A., "Intrusion Detection Systems, (IDS) Taxonomy – A Short Review," DOD Software Tech News, vol. 13, no. 2, June 2010, DOD Data Analysis Center for Software, Air Force Research Laboratory, Rome, N.Y., pp. 23 – 30.
23. Ali A. Ghorbani, WeiLu, and Mahbod Tavallaee, Network Intrusion Detectionand Prevention: concepts and techniques. London: Springer, 2010, p. 27–49.

КІБЕРБЕЗПЕКА: освіта, наука, техніка № 2 (30), 2025
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
Спеціальний випуск ISSN 2663 – 4023
Special issue

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

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

IMPLEMENTATION OF A HYBRID APPROACH BASED ON INTELLIGENT DATA ANALYSIS METHODS IN INTRUSION DETECTION SYSTEMS

Abstract. The growth of cyber influence and its diversity in cyberspace require a proactive approach to the development of intelligent models and methods. These tools are necessary for effective data analysis and their application in the field of cyber countermeasures. It is critically important to research innovative solutions capable of navigating the complexities of large data sets. The goal of the research is to ensure not only the accuracy of the analysis, but also the timely receipt of valuable information. Such research efforts are indispensable for meeting the growing demand for reliable data processing capabilities in various sectors. The development of information technology has led to a significant increase in the number of attacks on information system components. This has made intrusion detection an extremely relevant and critical task. In this context, data mining methods offer broad opportunities for application in various scientific and technical fields, including information security. This article is devoted to the study of the application of data mining methods in intrusion detection systems (IDS). The paper classifies intrusion detection systems according to various criteria and analyzes the mathematical apparatus of the selected methods. An overview of popular data mining methods that are widely used to detect and counter cyber threats is provided. Three methods are considered in detail: support vectors, clustering, and principal components as standalone methods, and a hybrid intrusion detection method is created based on them. The effectiveness of the methods considered was evaluated using the true positive rate (TPR) and false positive rate (FPR) indicators. The hybrid method uses the principle of multi-layered protection. It combines the high accuracy of SVM for known threats with the high sensitivity of clustering/PCA for unknown threats, providing an optimal balance between true and false positives in a real environment.

Keywords: cyber defense, cyberattack, intrusion detection system, clustering method, support vector method, intelligent data analysis methods, database, correct and false recognition, principal component method, efficiency.

REFERENCES(TRANSLATED AND TRANSLITERATED)

1. Chen, Y.H.; Yao, Y.Y. A multiview approach for intelligent data analysis based on data operators. Inf. Sci. 2008, 178, 1–20.
2. Yang, J.; Li, Y.; Liu, Q.; Li, L.; Feng, A.; Wang, T.; Zheng, S.; Xu, A.; Lyu, J. Brief introduction of medical database and data mining technology in big data era. J. Evid.-Based Med. 2020, 13, 57–69.
3. Young, T.; Hazarika, D.; Poria, S.; Cambria, E. Recent trends in deep learning based natural language processing. IEEE Comput. Intell. Mag. 2017, 13, 55–75.
4. Abkenar, S.B.; Kashani, M.H.; Mahdipour, E.; Jameii, S.M. Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. Telemat. Inform. 2020, 57, 101517.
5. Lande D.V., Subach I.Yu., Boyarinova Yu.E. Fundamentals of the theory and practice of intelligent data analysis in the field of cybersecurity: a textbook. — Kyiv: ISZSI Igor Sikorsky KPI, 2018. — 300 p.
6. Methodology for the synthesis of models of intelligent control and security systems for critical infrastructure facilities. Monograph / S.P. Yevseyev, O.Yu. Zakovorotny, O.V. Milov, G.A. Kuchuk, O.A.

Galuz, M.V. Koval, O.V. Voitko, R.V. Gryshchuk – Kharkiv: Published by Novyi Svit-2000, 2024. – 300 p.
7. Buryachok V.L. Fundamentals of the Formation of a State Cyber Security System: Monograph. – Kyiv: NAU, 2013. – 432 p.
8. Buryachok, V.L. Information and cyber security: socio-technical aspect: textbook / [V.L. Buryachok, V.B. Tolubko, V.O. Khoroshko, S.V. Tolyupa] — Kyiv: DUT, 2015.— 288 p.
9. S. V. Tolyupa, N. V. Lukova-Chuiko, V. S. Nakonechny, M. M. Brailovsky Methods of intelligent data distribution in network intrusion detection systems and functional resilience of information systems to cyberattacks. /: monograph – Kyiv: Format, 2021. – 370 p.
10. Enamul Kabir. A novel statistical technique for intrusion detection systems" Future Generation. Comput. Syst., 79 (2018), pp. 303-318.
11. Huaglory Tianfield. Data mining based cyber-attack detection Syst. simul. technol., 13 (2017)
12. Basant Subba, Santosh Biswas, Sushanta Karmakar. Intrusion detection systems using linear discriminant analysis and logistic regression. 2015 Annual IEEE India Conference (INDICON), IEEE (2015), pp. 1-6
13. Kai Peng, Victor Leung, Lixin Zheng, Shangguang Wang, Chao Huang, Tao Lin. Intrusion detection system based on decision tree over big data in fog environment. Wireless Commun. Mobile Comput. (2018), pp. 1-10.
14. Yu Xue, Weiwei Jia, Xuejian Zhao, Wei Pang. An Evolutionary Computation Based Feature Selection Method for Intrusion Detection" Security and Communication Networks. (2018), pp. 1-10
15. L. Xiao, Y. Chen, C.K. Chang. Bayesian model averaging of Bayesian network classifiers for intrusion detection. 2014 in IEEE 38th International Computer Software and Applications Conference Workshops on 35 (2014). pp. 1302-1310
16. Hari Om, Aritra Kundu. A hybrid system for reducing the false alarm rate of anomaly intrusion detection system. 2012 1st International Conference on Recent Advances in Information Technology (RAIT), IEEE (2012), pp. 131-136.
17. Zaiton Muda, Warusia Yassin, Md Nasir Sulaiman, Nur Izura Udzir. Intrusion detection based on k-means clustering and OneR classification. 2011 7th International Conference on Information Assurance and Security, IAS) (2011), pp. 192-197.
18. Uddin Chowdhury, Frederick Hammond, Glenn Konowicz, Chunsheng Xin, Hongyi Wu, Li Jiang. A few-shot DL approach for improved intrusion detection. 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON) (2017), pp. 456-462
19. Debar, H., Dacier, M., and Wespi, A. (1999), "Towards a Taxonomy of Intrusion Detection Systems," Computer Networks, vol. 31, 1999, pp. 805–822.
20. Debar, H., Dacier, M., and Wespi, A. (2000), "A Revised Taxonomy for Intrusion–Detection Systems," presente dat Annalesdes communications, vol. 55, 2000, pp. 361–78.
21. Kabiri, P., and Ghorbani, A., A. (2005), "Researchon Intrusion Detectionand Response: A Survey", International Journalof NetworkSecurity, Vol.1, No.2, Sep. 2005,pp.84–102.
22. Amer, S.H., Hamilton, J.A., "Intrusion Detection Systems, (IDS) Taxonomy – A Short Review," DOD Software Tech News, vol. 13, no. 2, June 2010, DOD Data Analysis Center for Software, Air Force Research Laboratory, Rome, N.Y., pp. 23 – 30.
23. Ali A. Ghorbani, WeiLu, and Mahbod Tavallaee, Network Intrusion Detectionand Prevention: concepts and techniques. London: Springer, 2010, p. 27–49.