КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (28), 2025
ISSN 2663 - 4023
DOI 10.28925/2663-4023.2025.28.818
УДК 004.9

Радівілова Тамара Анатоліївна
д. т. н., проф, професор кафедри інфокомунікаційної інженерії ім.В.В. Поповського
Харківський національний університет радіоелектроніки, Харків, Україна
ORCID ID: 0000-0001-5975-0269
tamara.radivilova@nure.ua

Добринін Ігор Станіславович
к.т.н, доцент, доцент кафедри інфокомунікаційної інженерії ім.В.В. Поповського
Харківський національний університет радіоелектроніки, Харків, Україна
ORCID ID: 0000-0001-8910-2609
ihor.dobrynin@nure.ua

Пантєлєєв Вадим Олегович
аспірант кафедри інфокомунікаційної інженерії ім.В.В. Поповського
Харківський національний університет радіоелектроніки, Харків, Україна
ORCID ID: 0009-0008-7824-8782
vadym.pantielieiev@nure.ua

Фісенко Дмитро Михайлович
аспірант кафедри інфокомунікаційної інженерії ім.В.В. Поповського
Харківський національний університет радіоелектроніки, Харків, Україна
ORCID ID: 0009-0006-9966-5329
dmytro.fisenko@nure.ua

Мазепа Артем Дмитрович
аспірант кафедри інфокомунікаційної інженерії ім.В.В. Поповського
Харківський національний університет радіоелектроніки, Харків, Україна
ORCID ID: 0009-0002-9977-5932
artem.mazepa@nure.ua

Білодід Володимир Григорович
науковий співробітник
Харківського національного університету
Повітряних Сил ім. І. Кожедуба,, Харків, Україна
ORCID ID: 0009-0002-8976-2310
vibos111@gmail.com

АНАЛІЗ МЕТОДІВ ПРОГНОЗУВАННЯ ВНУТРІШНІХ ЗАГРОЗ НА ОСНОВІ АНАЛІЗУ ДАНИХ СОЦІАЛЬНОЇ МЕРЕЖІ TWITTER

Анотація. Внутрішні загрози є значною проблемою для ефективного функціонування організацій. Дослідники зазвичай визначають внутрішні загрози, аналізуючи поведінку користувачів та трафік. Більшість існуючих робіт зосереджуються на методах виявлення внутрішніх загроз, а не на їх прогнозуванні. У цій роботі основна увага приділялася методам прогнозування внутрішніх загроз за допомогою аналізу даних соціальних мереж. У цій статті проаналізовано ефективність таких методів прогнозування: аналіз настроїв на основі словника, машинне навчання (Random Forest), рекурентні нейронні мережі (LSTM), трансформери (BERT), гібридний підхід (NLP CBOW + графічні нейронні мережі) та гібридний підхід (NLP CBOW + LPA). Для забезпечення надійності результатів ефективність методів прогнозування оцінювалася за допомогою декількох ключових показників: точність, відтворюваність, F1-показник, AUC-ROC, час навчання та час інференції. Для аналізу були проведені експерименти на наборі даних Sentiment140, який містить 1,6 мільйона твітів, позначених як позитивні або негативні. Для обробки даних та побудови моделей прогнозування використовувалися бібліотеки мови програмування Python. 80% даних було використано для навчання моделей, а 20% — для прогнозування. Результати аналізу показали, що найшвидшим методом прогнозування

інцидентів є аналіз тону на основі словників. Він може бути використаний для аналізу змін настроїв в організації за допомогою онлайн-аналізу соціальних мереж. Кожна організація може встановити власну шкалу реагування на внутрішні інциденти. Аналіз також показав, що гібридні підходи, які комплексно аналізують як текстовий вміст, так і структуру соціальних зв'язків, демонструють найвищі показники точності прогнозування внутрішніх інцидентів. Однак ці підходи вимагають значних обчислювальних ресурсів, великих обсягів даних для навчання та більше часу на навчання, а також мають низьку інтерпретованість.

Ключові слова: Natural Language Processing; Long Short-Term Memory; Random Forest; внутрішні загрози; оцінка якості прогнозів; прогнозування.

ВСТУП

Внутрішня загроза, як правило, означає зловмисні або ненавмисні дії з боку інсайдера, які завжди негативно впливають на конфіденційність, цілісність або доступність інформаційної системи організації [1]. Зазвичай інсайдери зазвичай добре обізнані з механізмами безпеки організації та мають доступ до системних служб, а ненавмисне завдання шкоди безпеці організації також є небезпечним, то внутрішня загроза є однією з найскладніших і найважчих для виявлення. В звітах організацій та наукових працях зазначається, що 67% організацій зазнали одного або декількох внутрішніх атак протягом останніх 12 місяців, а внутрішні атаки становлять 25% випадків кіберзлочинів [1] – [3]. Таким чином, внутрішні загрози, такі як саботаж систем, порушення безпеки даних, були визнані критичними проблемами безпеки, з якими стикаються різні установи та урядові органи.

Одним із ключових аспектів для стабільної роботи організації є контроль за виникненням інцидентів, що можуть порушити нормальний процес роботи. При цьому соціальні мережі часто стають джерелом ризиків для організації [2] – [4]. При спілкуванні в соціальних мережах співробітники обмінюються інформацією та можуть допустити її витік, або неправильно її інтерпретувати [5], [6]. Представники промисловості та наукових кіл запропонували багато підходів до виявлення внутрішніх загроз [7] – [9]. Оскільки зловмисна поведінка є дуже різноманітною, неможливо чітко охарактеризувати внутрішні загрози. Натомість більшість рішень спрямовані на створення моделей нормальної поведінки користувачів на основі аналізу їхньої поведінки в минулому, а аномалії визначаються як істотні відхилення від нормальної поведінки [1], [8], [9]. Більшість дослідницьких робіт направлені на аналіз даних соціальних мереж та виявлення вже існуючих інцидентів, а не на їхнє прогнозування [1], [7] – [12]. Таким чином важливою задачею є прогнозування інцидентів в організації на основі аналізу соціальних мереж.

Метою роботи є аналіз ефективності роботи методів прогнозування. Для виконання поставленої цілі необхідно вирішити низку задач: обрати методи прогнозування внутрішніх інцидентів на основі аналізу даних соціальних мереж, обрати стандартизовані метрики якості для розрахунку ефективності прогнозування методів, провести аналіз обраних методів прогнозування на однакових наборах даних соціальних мереж з використанням стандартизованих метрик якості.

ОГЛЯД МЕТОДІВ ПРОГНОЗУВАННЯ

Структурні методи засновані на виявленні внутрішньої структури даних та взаємозв'язків між різними елементами системи. До цієї групи відносяться нейронні мережі, ланцюги Маркова, методи опорних векторів, дерева рішень і випадкові ліси.

Штучні нейронні мережі імітують структуру та функціонування біологічних нейронних мереж і здатні виявляти складні нелінійні залежності в даних. Багатошаровий перцептрон (MLP) складається з вхідного, одного або кількох прихованих і вихідного шарів. Рекурентні нейронні мережі (RNN), особливо їх різновиди LSTM (Long Short-Term Memory) та GRU (Gated Recurrent Unit), ефективні для роботи з послідовними даними, такими як часові ряди [13] – [15]. Згорткові нейронні мережі (CNN) здатні автоматично виявляти просторові залежності в даних. У сучасних дослідженнях набувають поширення трансформери та моделі з механізмом уваги, які показують високу ефективність при аналізі послідовностей різної природи [16].

Дерева рішень розбивають простір ознак на підмножини, базуючись на правилах виду «якщо – то». Вони можуть використовуватися як для класифікації, так і для регресії. Випадкові ліси об’єднують прогнози декількох дерев рішень, побудованих на різних підмножинах даних, що підвищує стабільність і точність прогнозів.

У сучасних дослідженнях все більшої популярності набувають методи глибинного навчання, зокрема глибинні нейронні мережі, які здатні виявляти складні залежності в даних без необхідності їх попередньої трансформації [17]. Архітектури на основі трансформерів, такі як BERT, GPT і їх модифікації, демонструють високу ефективність при аналізі текстових даних, що особливо важливо для прогнозування на основі інформації з соціальних мереж.

Гібридні методи прогнозування поєднують переваги різних підходів для досягнення кращої точності [18] – [20]. Графи широко використовуються в Natural Language Processing (NLP) [21]. Можна побудувати графік тексту, що містить кілька ієрархій елементів, таких як документ, уривок, речення та слово. Такий підхід дозволяє відобразити більш багаті взаємозв’язки між елементами тексту. Даний гібридний підхід надає широкі можливості щодо прогнозування в області соціальних мереж.

Хвиля досліджень на перетині глибокого навчання на графах (Graph Neural Networks, GNN) [22] і Natural Language Processing (NLP) вплинула на різноманітні завдання включаючи завдання NLP [21] класифікацію, семантичне рольове маркування, узагальнення та прогнозування. GNN зазвичай працюють шляхом перетворення, поширення та агрегації вузлів/ребер через кілька нейронних шарів, щоб краще вивчити графічні представлення. Як загальна система навчання на основі графів, GNN можуть застосовуватися до різних завдань, пов’язаних з графами, таких як класифікація вузлів, прогнозування зв’язків та класифікація графів [23].

Алгоритми поширення міток (LPA) — це клас напівнаглядових алгоритмів на основі графів, які поширюють мітки з мічених точок даних на раніше немічені точки даних [21]. В основному LPA працюють шляхом повторного поширення та агрегації міток по графу. При кожному повторенні кожен вузол змінює свою мітку на основі міток, які мають сусідні вузли. В результаті інформація про мітки поширюється по графу. Застосування LPA широко використовуються в літературі з мережевих наук для виявлення структур спільнот у складних мережах. У літературі з NLP LPA успішно застосовуються для усунення багатозначності слів та аналізу емоційного забарвлення [21], [22]. Ці застосування зазвичай зосереджуються на напівнаглядовому навчанні, де міткованих даних недостатньо, і використовують алгоритм LPA для поширення міток з обмеженої кількості міткованих прикладів на велику кількість подібних немічкованих прикладів, виходячи з припущення, що подібні приклади повинні мати подібні мітки.

ISSN 2663 - 4023

ПАРАМЕТРИ ОЦІНКИ ЯКОСТІ ПРОГНОЗУВАННЯ

Оцінка якості прогнозування є критично важливим етапом, що дозволяє визначити надійність і достовірність отриманих прогнозів. Існує ряд стандартних метрик, які використовуються для кількісної оцінки точності методів прогнозування. Вони поділяються на абсолютні показники, порівняльні показники, а також методи декомпозиції помилки прогнозу.

Верифікація прогнозів є процесом перевірки їх обґрунтованості та точності. Основні методи верифікації включають ретроспективний прогноз (Backtesting), при якому вихідні дані поділяються на навчальну і тестову вибірки; крос-валідацію (Cross-validation), що є систематичним підходом до оцінки якості моделі; метод ковзного вікна (Rolling Window Method), який є спеціальним видом крос-валідації для часових рядів; крива AUC-ROC, що є інструментом для оцінки ефективності моделей бінарної класифікації. В бінарній класифікації є чотири можливі результати для прогнозування тесту: істинно позитивний (TP), хибно позитивний (NP), істинно негативний(TN) і хибно негативний(FN) [24].

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \quad (1)$$

Accuracy — це кількість правильно класифікованих екземплярів даних від загальної кількості екземплярів даних.

$$\Pr ecision = \frac{TP}{TP + FP} \quad (2)$$

Precision — це метрика, яка вимірює, наскільки часто модель машинного навчання правильно прогнозує позитивний клас. Вищий показник Precision вказує на те, що модель робить менше хибнопозитивних прогнозів, але вона не враховує хибнонегативні результати.

$$\mathrm{Re} call = \frac{TP}{TP + FN} \quad (3)$$

Recall (TPR) — це метрика, яка вимірює, як часто модель машинного навчання правильно ідентифікує позитивні приклади (TP) з усіх реальних позитивних зразків у наборі даних. Хибнонегативні помилки (FN) коштуватимуть дорожче, ніж хибнопозитивні (FP). Recall не враховує вартість хибних спрацьовувань (FP).

$$F1Score = 2\frac{\Pr ecision * \mathrm{Re} call}{\Pr ecision + \mathrm{Re} call} \quad (4)$$

Отже, в ідеалі в гарному класифікаторі і Precision, і Recall прагнуть до одиниці, що також означає, що FP і FN дорівнюють нулю. F1 score — це середнє гармонійне значення Precision і Recall і є кращим показником, ніж Accuracy.

Крива AUC-ROC відображає співвідношення істинних позитивних результатів (TPR) та хибнопозитивних результатів (FPR) при різних порогових значеннях, показуючи, наскільки добре модель може розрізняти два класи, наприклад, позитивні та негативні результати (наявність або відсутність загроз). В AUC-ROC використовуються TPR (True Positive Rate) — співвідношення правильно передбачених позитивних результатів; FPR (False Positive Rate) — відношення неправильно передбачених негативних результатів; Specificity — частка фактичних негативних результатів, правильно ідентифікованих моделлю (зворотний показник до FPR); Sensitivity/Recall — відсоток фактичних позитивних результатів, правильно ідентифікованих моделлю (аналогічне TPR) [25]. На рис. 2 зображено загальний вигляд метрики оцінки класифікації кривої ROC-AUC.

ISSN 2663 - 4023

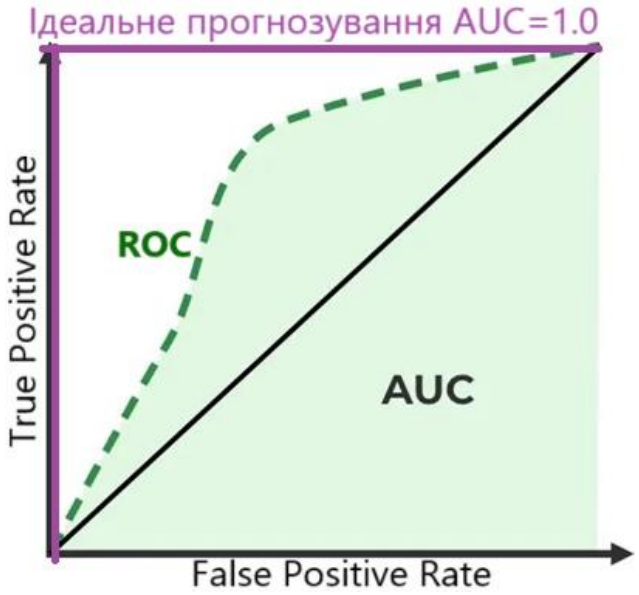

Рис. 1. Метрика оцінки класифікації ROC-AUC

Оцінка AUC-ROC коливається від 0 до 1, де 0,5 означає випадкове вгадування, а 1 — ідеальну роботу. Як правило, оцінка AUC-ROC вище 0,8 вважається хорошою, а вище 0,9 — відмінною. Однак корисність моделі залежить від конкретної проблеми та варіанту використання. Необхідно інтерпретувати показник AUC-ROC в контексті, разом з іншими показниками якості класифікації, такими як accuracy, precision, recall або F1-score. Існують різні методи розрахунку показника AUC-ROC, але найпоширенішим з них є правило трапецій. Воно передбачає апроксимацію площі під кривою ROC шляхом поділу її на трапеції з вертикальними лініями на значеннях FPR і горизонтальними лініями на значеннях TPR. Потім обчислюється площа шляхом підсумовування площ трапецій.

Важливо зазначити, що для різних типів задач прогнозування можуть бути більш релевантними різні метрики. При прогнозуванні рідкісних подій, таких як інциденти безпеки, важливими можуть бути метрики, що враховують баланс між хибнопозитивними і хибнонегативними прогнозами, такі як F1-score або Area Under the ROC Curve (AUC-ROC).

Вибір показників точності прогнозу залежить від завдань, які ставить перед собою дослідник при аналізі точності прогнозу. Про якість прогнозів, отриманих за різними методиками і моделями, можна судити лише за сукупністю зіставлень прогнозів та їх реалізації.

Для ефективного прогнозування в області соціальних мереж, що є особливо актуальним в кіберпросторі, важливо враховувати динаміку поширення інформації, структуру соціальних зв’язків, тональність повідомлень, а також поведінкові патерни користувачів [26]. Методи, які показують високу ефективність у цій області, включають моделі дифузії інформації, графові нейронні мережі, трансформери для аналізу тексту, системи виявлення аномалій, а також гібридні підходи, що поєднують різні методики.

Важливо зазначити, що в реальних задачах прогнозування часто найкращих результатів досягають гібридні методи та ансамблі моделей, які дозволяють компенсувати недоліки окремих підходів. Крім того, процес вибору оптимального методу повинен включати експериментальне порівняння декількох кандидатів на конкретному наборі даних з використанням відповідних метрик якості.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Для об'єктивної оцінки ефективності різних методів прогнозування необхідно провести їх порівняльний аналіз на однакових наборах даних з використанням стандартизованих метрик якості. У цьому розділі представлено результати порівняльного аналізу ефективності методів прогнозування. Експерименти було проведено на наборі даних використовувався набір даних Sentiment140, який містить 1,6 мільйона твітів, позначених як позитивні або негативні [27]. Для обробки даних, побудови моделей прогнозування було використано бібліотеки мови програмування Python. Для навчання моделей використовувалось 80% даних, а для прогнозування використовувалось 20% даних. Для прогнозування було обрано такі методи: аналіз тональності на основі словників, машинне навчання (Random Forest), рекурентні нейронні мережі (LSTM), трансформери (BERT), гібридний підхід (NLP CBOW + графові нейронні мережі), гібридний підхід (NLP CBOW + LPA).

Для забезпечення надійності результатів оцінка ефективності методів прогнозування проводилася за кількома ключовими метриками: Precision, Recall, F1-score та ROC-AUC, час навчання, час інференції.

Для ілюстрації ефективності різних методів при прогнозуванні внутрішніх інцидентів на основі даних соціальних мереж розглянемо порівняльний аналіз шести ключових підходів: традиційний аналіз тональності текстів з використанням словників, класифікація на основі машинного навчання (Random Forest), мережі довгої короткочасної пам'яті (Long/short termmemory, LSTM), трансформери (BERT) та гібридні підходи, що поєднує аналіз тексту (CBOW) з графовими нейронними мережами та графовий алгоритм поширення міток LPA.

В роботі ми використовуємо модель Word2Vec, що базується на використанні нейронної мережі, щоб вивчати високорозмірні вставки слів з необробленого тексту (Continuous Bag-of-Words, CBOW) шляхом передбачення цільового слова з навколишніх слів. Після відкидання останнього шару після навчання модель приймає слово на вході і виводять слово-вкладення, яке можна використовувати як вхідні дані для багатьох завдань NLP. Результати цього аналізу представлені в табл. 1.

Таблиця 1

Порівняльний аналіз методів прогнозування внутрішніх інцидентів на основі даних соціальних мереж

Метод	Precision	Recall	F1-score	AUC-ROC	Час навчання	Час інференції
Аналіз тональності на основі словників	0.65	0.48	0.55	0.71	Не потребує	Низький
Машинне навчання (Random Forest)	0.73	0.65	0.69	0.79	Середній	Низький
Рекурентні нейронні мережі (LSTM)	0.78	0.71	0.74	0.83	Високий	Середній
Трансформери (BERT)	0.82	0.76	0.79	0.87	Дуже високий	Середній-високий
Гібридний підхід (NLP CBOW) + LPA)	0.86	0.8	0.83	0.9	Дуже високий	Високий
Гібридний підхід (NLP CBOW) + графові нейронні мережі)	0.85	0.81	0.83	0.91	Дуже високий	Високий

З наведеної таблиці видно, що розглянуті методи прогнозують внутрішні інциденти з високою точністю (≥ 0,7), окрім метода аналізу тональності на основі словників, який показав значення точності прогнозу F1-score=0,55. Але цей метод не потребує часу для навчання і працює швидше за інші методи. Метод машинного навчання Random Forest дає більші значення точності прогнозування F1-score=0,69, але потребує середнє значення часу для навчання. Краще працюють методи рекурентних нейронних мереж LSTM з F1-score=0,74 та трансформерів BERT з F1-score=0,79, які потребують більше часу для навчання, більше ресурсів та більше часу інференції. Також з наведеної таблиці видно, що гібридні підходи, які комплексно аналізують як текстовий контент, так і структуру соціальних зв'язків, демонструють найвищі показники точності прогнозування внутрішніх інцидентів, хоча і вимагають значних обчислювальних ресурсів та мають низьку інтерпретованість. Це підтверджує загальну тенденцію, що для складних, багатофакторних задач прогнозування гібридні моделі є найбільш ефективними.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

В роботі було проведено аналіз ефективності роботи таких методів прогнозування: аналіз тональності на основі словників, машинне навчання (Random Forest), рекурентні нейронні мережі (LSTM), трансформери (BERT), гібридний підхід (NLP CBOW + графові нейронні мережі), гібридний підхід (NLP CBOW + LPA). Для аналізу було проведено експерименти на відкритому наборі даних Sentiment140, який містить 1,6 мільйона твітів, позначених як позитивні, нейтральні або негативні. Результати аналізу показали, що найшвидше прогнозує інциденти метод аналізу тональності на основі словників. Його можна використовувати для аналізу зміни настроїв в організації шляхом онлайн аналізу соціальних мереж. Кожна організація може виставити свою шкалу для початку реагування на внутрішні інциденти. Також в ході аналізу виявлено, що гібридні підходи, які комплексно аналізують як текстовий контент, так і структуру соціальних зв'язків, демонструють найвищі показники точності прогнозування внутрішніх інцидентів. Однак ці підходи вимагають значних обчислювальних ресурсів, великих обсягів навчальних даних та потребують більше часу для навчання та мають низьку інтерпретованість.

Поєднання GNN з Transformers для NLP Найкраще в Transformers — це просте використання елегантної архітектури моделі безпосередньо на оригінальній послідовності тексту, що відокремлює моделювання графічних даних всередині моделі Transformers від вхідних даних (і від кінцевого користувача). Однак недоліком цього вибору моделі є те, що Transformers не можуть безпосередньо працювати з більш складними даними, такими як дані з графічною структурою. На відміну від цього, GNN — це більш загальна архітектура моделі, яка безпосередньо працює з графічними даними, які, однак, повинні бути створені кінцевим користувачем за допомогою знань у конкретній галузі або інших технік моделювання графіків.

Цікавим подальшим дослідженням є аналіз методу прогнозування Graph Transformers, який є популярною моделлю, яка адаптує механізм самоуваги, що враховує структуру, для поєднання переваг Transformers і GNN. А також порівняльний аналіз гібридних підходів (NLP CBOW + графові нейронні мережі), (NLP CBOW + LPA) та методу прогнозування Graph Transformers за часовими характеристиками та якісними для різних підходів побудови графів та для різних наборів даних за об'ємом та маркуванням.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Casolaro, A., Capone, V., Iannuzzo, G., & Camastra, F. (2023). Deep learning for time series forecasting: Advances and open problems. *Information, 14(11),* 598.. https://doi.org/10.3390/info14110598
2. Hernández, R., Gutiérrez, I., & Castro, J. (2025). Social Network Analysis: A Novel Paradigm for Improving Community Detection. *International Journal of Computational Intelligence Systems, 18(1),* 87. https://doi.org/10.1007/s44196-025-00812-9
3. Rudchenko, D.V. (2017). Vyyavlennya spil′not ta yikh lideriv v sotsial′nykh merezhakh dlya zabezpechennya bezpeky [Identification of communities and their leaders in social networks to ensure security], *Information processing systems, 4(150),* 128–131. http://dx.doi.org/10.30748/soi.2017.150.26
4. Lyudmyla, K., Vitalii, B., & Tamara, R. (2017, October). Fractal time series analysis of social network activities. In *2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*. 456–459). IEEE. https://doi.org/10.1109/INFOCOMMST.2017.8246438
5. Huang, D., Song, J., & He, Y. (2024). Community detection algorithm for social network based on node intimacy and graph embedding model. *Engineering Applications of Artificial Intelligence, 132, 107947.* https://doi.org/10.1007/978-3-031-67195-1_79
6. Kirichenko, L., Radivilova, T., & Anders, C. (2017). Detecting cyber threats through social network analysis: short survey. *SocioEconomic Challenges, 1(1),* 20—34.
7. Kyriazos, T., & Poga, M. (2024). Application of machine learning models in social sciences: managing nonlinear relationships. *Encyclopedia, 4(4),* 1790–1805. https://doi.org/10.3390/encyclopedia4040118
8. Rawat, R., Mahor, V., Chirgaiya, S., & Rathore, A.S. (2021). Applications of Social Network Analysis to Managing the Investigation of Suspicious Activities in Social Media Platforms. In: Daimi, K., Peoples, C. (eds) *Advances in Cybersecurity Management. Springer Cham,* 315–335. https://doi.org/10.1007/978-3-030-71381-2_15
9. Li, W., & Law, K. L. E. (2024). Deep Learning Models for Time Series Forecasting: A Review. *IEEE Access, 12,* 92306–92327. https://doi.org/10.1109/ACCESS.2024.3422528
10. Kirichenko, L., Radivilova, T., Zinkevich, I. (2018). Comparative Analysis of Conversion Series Forecasting in E-commerce Tasks. In: Shakhovska N., Stepashko V. (eds*) Advances in Intelligent Systems and Computing II. CSIT 2017. Advances in Intelligent Systems and Computing, 689,* 230–242. https://doi.org/10.1007/978-3-319-70581-1_16
11. Mulesa, O., Povkhan, I., Radivilova, T., & Baranovskyi, O. (2021). Devising a method for constructing the optimal model of time series forecasting based on the principles of competition. *Eastern-European Journal of Enterprise Technologies, 5 (4 (113)),* 6–11. https://doi.org/10.15587/1729-4061.2021.240847
12. Mulesa, O., Batyuk, A., Geche, F., Melnyk, O., Palinchak, M., & Radivilova, T. (2021). Information technology for time series forecasting based on the evolutionary method of the forecasting scheme synthesis. In *IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT),* 258–261. IEEE. https://doi.org/10.1109/CSIT52700.2021.9648639
13. Wan, A., Chang, Q., Khalil, A. B., & He J. (2023). Short-term power load forecasting for combined heat and power using CNN-LSTM enhanced by attention mechanism. *Energy, 282, 128274.* https://doi.org/10.1016/j.energy.2023.128274
14. Alotaibi, M. A. (2022). Machine Learning Approach for Short-Term Load Forecasting Using Deep Neural Network. *Energies, 15(17), 6261.* https://doi.org/10.3390/en15176261
15. Kirichenko, L., Radivilova, T., & Bulakh, V. (2019). Machine learning in classification time series with fractal properties. *Data, 4(1(5)),* 1-13. https://doi.org/10.3390/data4010005
16. Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., & Sun, L. (2023). *Transformers in time series: A survey*. (2022). arXiv preprint arXiv:2202.07125.
17. Su, X., Xue, S., Liu, F., Wu, J., Yang, J., Zhou, C., Hu, W., Paris, C., Nepal, S., Jin, D., Sheng, Q., & Yu, P. (2024). A comprehensive survey on community detection with deep learning. *IEEE Transactions on Neural Networks and Learning Systems, 35(4),* 4682–4702. https://doi.org/10.1109/TNNLS.2021.3137396.
18. Zhang, Y., Zhou, X., Zhang, Y., Li, S., & Liu, S. (2025). Improving time series forecasting in frequency domain using a multi resolution dual branch mixer with noise insensitive ArcTanLoss. *Scientific Reports, 15(1), 12557.* https://doi.org/10.1038/s41598-025-95529-2
19. Cai, J., Hao, J., Yang, H., Yang, Y., Zhao, X., Xun, Y., & Zhang, D. (2024). A new community detection method for simplified networks by combining structure and attribute information. *Expert Systems with Applications, 246(123103).* https://doi.org/10.1016/j.eswa.2023.123103

20. Yuliansyah, H., Othman, Z. A., Bakar, A. A. (2020). Taxonomy of Link Prediction for Social Network Analysis: A Review. *IEEE Access*, *8*, 183470–183487. https://doi.org/10.1109/ACCESS.2020.3029122
21. Wu, L., Chen, Y., Shen, K., Guo, X., Gao, H., Li, S., & Long, B. (2023). Graph neural networks for natural language processing: A survey. *Foundations and Trends in Machine Learning*, *16(2)*, 119–328. https://doi.org/10.48550/arXiv.2106.06090
22. 22 Zhang, W., He, P., Qin, C., Yang, F., & Liu, Y. (2024). A graph attention network-based model for anomaly detection in multivariate time series. *The Journal of Supercomputing*, *80(6)*, 8529–8549. https://doi.org/10.1007/s11227-023-05772-5
23. Kirichenko, L., Radivilova, T., & Ryzhanov, V.(2022). Applying Visibility Graphs to Classify Time Series. In: Babichev S., Lytvynenko V. (eds) *Lecture Notes in Computational Intelligence and Decision Making. ISDMCI 2021. Lecture Notes on Data Engineering and Communications Technologies*, *77*, 397–409. https://doi.org/10.1007/978-3-030-82014-5_26
24. St-Aubin, P., & Agard, B. (2022). Precision and Reliability of Forecasts Performance Metrics. *Forecasting*, *4(4)*, 882–903. https://doi.org/10.3390/forecast4040048
25. *AUC ROC Curve in Machine Learning.* 2025. URL: https://www.geeksforgeeks.org/auc-roc-curve/
26. Shih, S. Y., Sun, F. K., Lee, H. Y. (2019). "Temporal pattern attention for multivariate time series forecasting". *Machine Learning*, *108*, 1421–1441. https://doi.org/10.1007/s10994-019-05815-0
27. *Sentiment140 dataset with 1.6 million tweets. 2018.* URL: https://www.kaggle.com/datasets/kazanova/sentiment140/data

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (28), 2025
ISSN 2663 - 4023

Tamara Radivilova
Dr. eng. sc., prof., professor of V.V. Popovskyy department of infocommunication engineering,
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID: 0000-0001-5975-0269
tamara.radivilova@gmail.com

Ihor Dobrynin
PhD, Assoc. Prof., Assoc. Prof. of the V.V. Popovsky Department of Infocommunication Engineering
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID: 0000-0001-8910-2609
ihor.dobrynin@nure.ua

Vadym Pantelieiev
Postgraduate student of the V.V. Popovskyy department of infocommunication engineering,
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID: 0009-0008-7824-8782
vadym.pantielieiev@nure.ua

Dmytro Fisenko
Postgraduate student of the V.V. Popovskyy department of infocommunication engineering,
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID: 0009-0006-9966-5329
dmytro.fisenko@nure.ua

Artem Mazepa
Postgraduate student of the V.V. Popovskyy department of infocommunication engineering,
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID: 0009-0002-9977-5932
artem.mazepa@nure.ua

Volodymyr Bilodid
Researcher
of Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine
ORCID ID: 0009-0002-8976-2310
vibos111@gmail.com

ANALYSIS OF METHODS FOR PREDICTING INTERNAL THREATS BASED ON THE ANALYSIS OF DATA FROM THE SOCIAL NETWORK TWITTER

Abstract. Internal threats are a significant problem for the effective functioning of organizations. Researchers usually identify internal threats by analyzing user behavior and traffic. Most existing works focus on methods for detecting internal threats rather than predicting them. In this work, the main focus was on methods for predicting internal threats using social network data analysis. This article analyzes the effectiveness of the following prediction methods: dictionary-based sentiment analysis, machine learning (Random Forest), recurrent neural networks (LSTM), transformers (BERT), a hybrid approach (NLP CBOW + graph neural networks), and a hybrid approach (NLP CBOW + LPA). To ensure the reliability of the results, the effectiveness of the forecasting methods was evaluated using several key metrics: Precision, Recall, F1-score, and ROC-AUC, training time, and inference time. For the analysis, experiments were conducted on the Sentiment140 dataset, which contains 1.6 million tweets labeled as positive or negative. Python programming language libraries were used to process data and build prediction models. 80% of the data was used to train the models, and 20% was used for prediction. The results of the analysis showed that the fastest method for predicting incidents is tone analysis based on dictionaries. It can be used to analyze changes in sentiment within an organization through online analysis of social networks. Each organization can set its own scale for responding to internal incidents. The analysis also found that hybrid approaches that comprehensively analyze both text content and the structure of social connections demonstrate the highest accuracy rates for predicting internal incidents. However, these

approaches require significant computing resources, large amounts of training data, and more time for training, and have low interpretability.

Keywords: Natural Language Processing; Long Short-Term Memory; Random Forest; internal threats; forecast quality assessment; forecasting.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Casolaro, A., Capone, V., Iannuzzo, G., & Camastra, F. (2023). Deep learning for time series forecasting: Advances and open problems. *Information, 14(11),* 598.. https://doi.org/10.3390/info14110598
2. Hernández, R., Gutiérrez, I., & Castro, J. (2025). Social Network Analysis: A Novel Paradigm for Improving Community Detection. *International Journal of Computational Intelligence Systems*, *18(1),* 87. https://doi.org/10.1007/s44196-025-00812-9
3. Rudchenko, D.V. (2017). Vyyavlennya spil′not ta yikh lideriv v sotsial′nykh merezhakh dlya zabezpechennya bezpeky [Identification of communities and their leaders in social networks to ensure security], *Information processing systems, 4(150),* 128–131. http://dx.doi.org/10.30748/soi.2017.150.26
4. Lyudmyla, K., Vitalii, B., & Tamara, R. (2017, October). Fractal time series analysis of social network activities. In *2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*. 456–459). IEEE. https://doi.org/10.1109/INFOCOMMST.2017.8246438
5. Huang, D., Song, J., & He, Y. (2024). Community detection algorithm for social network based on node intimacy and graph embedding model. *Engineering Applications of Artificial Intelligence*, *132*, *107947*. https://doi.org/10.1007/978-3-031-67195-1_79
6. Kirichenko, L., Radivilova, T., & Anders, C. (2017). Detecting cyber threats through social network analysis: short survey. *SocioEconomic Challenges*, *1(1),* 20—34.
7. Kyriazos, T., & Poga, M. (2024). Application of machine learning models in social sciences: managing nonlinear relationships. *Encyclopedia*, *4(4),* 1790–1805. https://doi.org/10.3390/encyclopedia4040118
8. Rawat, R., Mahor, V., Chirgaiya, S., & Rathore, A.S. (2021). Applications of Social Network Analysis to Managing the Investigation of Suspicious Activities in Social Media Platforms. In: Daimi, K., Peoples, C. (eds) *Advances in Cybersecurity Management. Springer Cham*, 315–335. https://doi.org/10.1007/978-3-030-71381-2_15
9. Li, W., & Law, K. L. E. (2024). Deep Learning Models for Time Series Forecasting: A Review. *IEEE Access*, *12,* 92306–92327. https://doi.org/10.1109/ACCESS.2024.3422528
10. Kirichenko, L., Radivilova, T., Zinkevich, I. (2018). Comparative Analysis of Conversion Series Forecasting in E-commerce Tasks. In: Shakhovska N., Stepashko V. (eds*) Advances in Intelligent Systems and Computing II. CSIT 2017. Advances in Intelligent Systems and Computing*, *689,* 230–242. https://doi.org/10.1007/978-3-319-70581-1_16
11. Mulesa, O., Povkhan, I., Radivilova, T., & Baranovskyi, O. (2021). Devising a method for constructing the optimal model of time series forecasting based on the principles of competition. *Eastern-European Journal of Enterprise Technologies*, *5 (4 (113)),* 6–11. https://doi.org/10.15587/1729-4061.2021.240847
12. Mulesa, O., Batyuk, A., Geche, F., Melnyk, O., Palinchak, M., & Radivilova, T. (2021). Information technology for time series forecasting based on the evolutionary method of the forecasting scheme synthesis. In *IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT)*, 258–261. IEEE. https://doi.org/10.1109/CSIT52700.2021.9648639
13. Wan, A., Chang, Q., Khalil, A. B., & He J. (2023). Short-term power load forecasting for combined heat and power using CNN-LSTM enhanced by attention mechanism. *Energy, 282, 128274*. https://doi.org/10.1016/j.energy.2023.128274
14. Alotaibi, M. A. (2022). Machine Learning Approach for Short-Term Load Forecasting Using Deep Neural Network. *Energies, 15(17), 6261*. https://doi.org/10.3390/en15176261
15. Kirichenko, L., Radivilova, T., & Bulakh, V. (2019). Machine learning in classification time series with fractal properties. *Data, 4(1(5)),* 1-13. https://doi.org/10.3390/data4010005
16. Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., & Sun, L. (2023). *Transformers in time series: A survey*. (2022). arXiv preprint arXiv:2202.07125.
17. Su, X., Xue, S., Liu, F., Wu, J., Yang, J., Zhou, C., Hu, W., Paris, C., Nepal, S., Jin, D., Sheng, Q., & Yu, P. (2024). A comprehensive survey on community detection with deep learning. *IEEE Transactions on Neural Networks and Learning Systems*, *35(4),* 4682–4702. https://doi.org/10.1109/TNNLS.2021.3137396.

18. Zhang, Y., Zhou, X., Zhang, Y., Li, S., & Liu, S. (2025). Improving time series forecasting in frequency domain using a multi resolution dual branch mixer with noise insensitive ArcTanLoss. *Scientific Reports, 15(1), 12557*. https://doi.org/10.1038/s41598-025-95529-2
19. Cai, J., Hao, J., Yang, H., Yang, Y., Zhao, X., Xun, Y., & Zhang, D. (2024). A new community detection method for simplified networks by combining structure and attribute information. *Expert Systems with Applications, 246(123103).* https://doi.org/10.1016/j.eswa.2023.123103
20. Yuliansyah, H., Othman, Z. A., Bakar, A. A. (2020). Taxonomy of Link Prediction for Social Network Analysis: A Review. *IEEE Access, 8,* 183470–183487. https://doi.org/10.1109/ACCESS.2020.3029122
21. Wu, L., Chen, Y., Shen, K., Guo, X., Gao, H., Li, S., & Long, B. (2023). Graph neural networks for natural language processing: A survey. *Foundations and Trends in Machine Learning, 16(2),* 119–328. https://doi.org/10.48550/arXiv.2106.06090
22. 22 Zhang, W., He, P., Qin, C., Yang, F., & Liu, Y. (2024). A graph attention network-based model for anomaly detection in multivariate time series. *The Journal of Supercomputing, 80(6),* 8529–8549. https://doi.org/10.1007/s11227-023-05772-5
23. Kirichenko, L., Radivilova, T., & Ryzhanov, V.(2022). Applying Visibility Graphs to Classify Time Series. In: Babichev S., Lytvynenko V. (eds) *Lecture Notes in Computational Intelligence and Decision Making. ISDMCI 2021. Lecture Notes on Data Engineering and Communications Technologies, 77,* 397–409. https://doi.org/10.1007/978-3-030-82014-5_26
24. St-Aubin, P., & Agard, B. (2022). Precision and Reliability of Forecasts Performance Metrics. *Forecasting, 4(4),* 882–903. https://doi.org/10.3390/forecast4040048
25. *AUC ROC Curve in Machine Learning.* 2025. URL: https://www.geeksforgeeks.org/auc-roc-curve/
26. Shih, S. Y., Sun, F. K., Lee, H. Y. (2019). "Temporal pattern attention for multivariate time series forecasting". *Machine Learning, 108,* 1421–1441. https://doi.org/10.1007/s10994-019-05815-0
27. *Sentiment140 dataset with 1.6 million tweets. 2018.* URL: https://www.kaggle.com/datasets/kazanova/sentiment140/data