ДОДАТОК А

Графічний матеріал кваліфікаційної роботи

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ

КВАЛІФІКАЦІЙНА РОБОТА

Модель виявлення аномалій в мікросервісних інфраструктурах на основі методу головних компонент

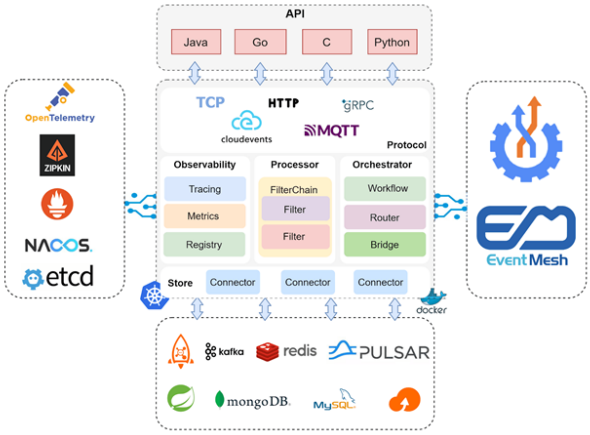
ВИКОНАВ:
Студент гр СПм-23-4 Гудзинський І В

КЕРІВНИК:
доц. Мартовицький В.О.

ХАРКІВ
2025р.

Актуальність дослідження

Зростання складності та динамічності мікросервісних інфраструктур у хмарних та розподілених середовищах створює значні виклики для своєчасного виявлення аномалій, що можуть негативно впливати на продуктивність, надійність та безпеку систем.



Мета та завдання

Мета роботи:
Розробити та дослідити модель виявлення аномалій у мікросервісних інфраструктурах з використанням робастного методу головних компонент (RPCA), яка дозволяє ефективно і стійко виявляти відхилення в поведінці системи на основі аналізу високовимірних технічних метрик.

Завдання роботи:
Провести аналіз особливостей мікросервісної архітектури та типів аномалій, що виникають у таких системах.
Розглянути сучасні підходи до виявлення аномалій та обґрунтувати доцільність застосування RPCA.
Зібрати та підготувати набір даних, що відображає роботу мікросервісної інфраструктури.
Реалізувати модель виявлення аномалій на основі алгоритму RPCA.
Провести тестування моделі на реальних або симульованих даних, оцінити точність та ефективність виявлення аномалій.
Порівняти результати RPCA з іншими методами.

3

Виявлення аномалій

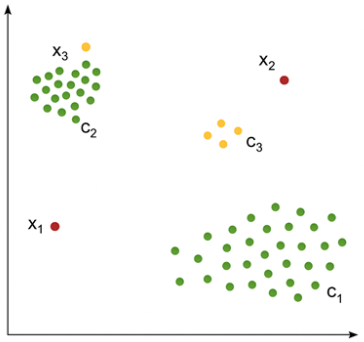


Рисунок 1 – Приклад аномалій

4

Виявлення аномалій

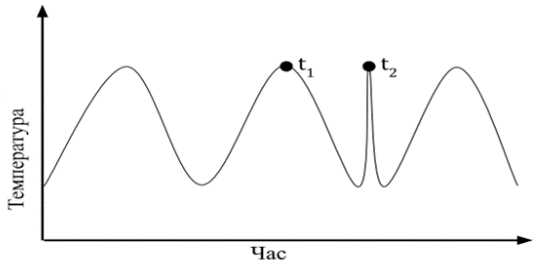
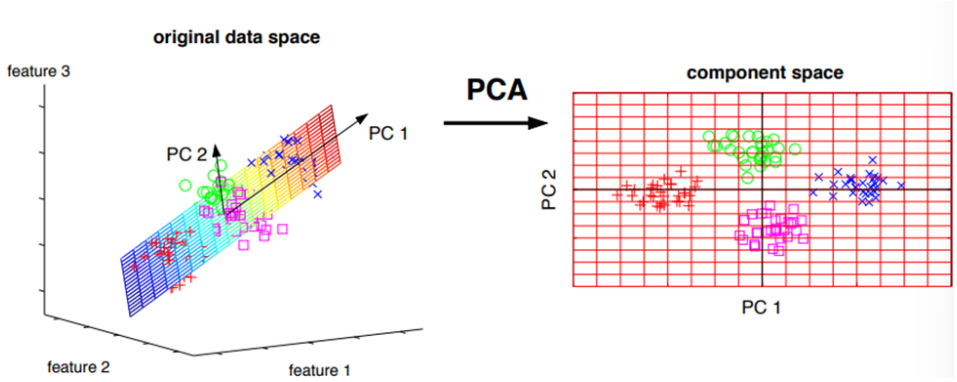


Рисунок 2 – Точки t1 і t2 мають однакове значення, але в різні моменти часу

5

Методи скорочення даних для виявлення аномалій



6

Алгоритм RPCA (Robust Principal Component Analysis)

Мета: розкласти матрицю D на:

- L — низькорівнева матриця (нормальні дані)
- S — розріджена матриця (аномалії)

Оптимізаційна задача:

min ||L||_* + λ ||S||_1 при умові D = L + S

λ = 1 / √max(m, n)

Кроки реалізації (метод ALM):

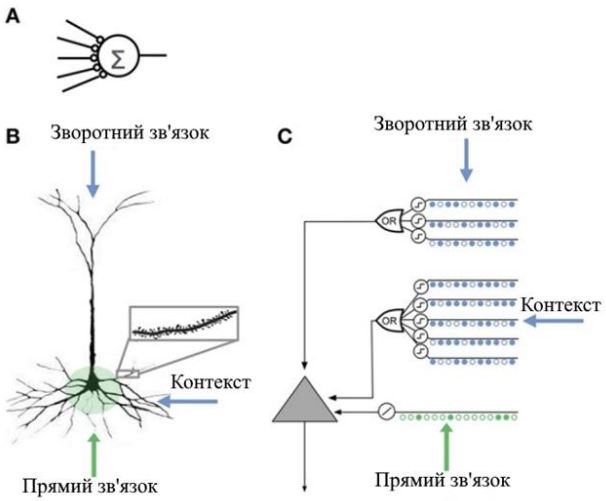
1. Ініціалізація: L = 0, S = 0, Y = 0, μ > 0
2. Повторювати до збіжності:
   - L ← SVT_{1/μ}(D - S + Y/μ)
   - S ← Shrink_{λ/μ}(D - L + Y/μ)
   - Y ← Y + μ(D - L - S)
3. Зупинка: ||D - L - S||_F / ||D||_F < ε

Результат:

- L — відновлена структура
- S — аномалії

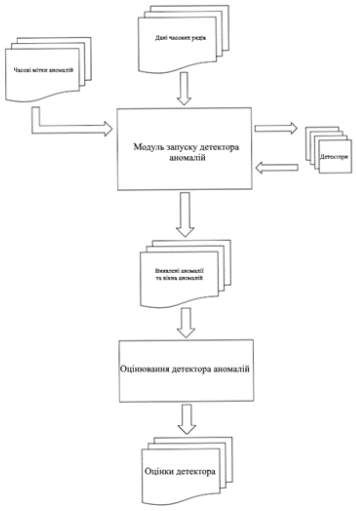
7

Hierarchical Temporal Memory



8

Система бенчмаркінгу детекторів аномалій



9

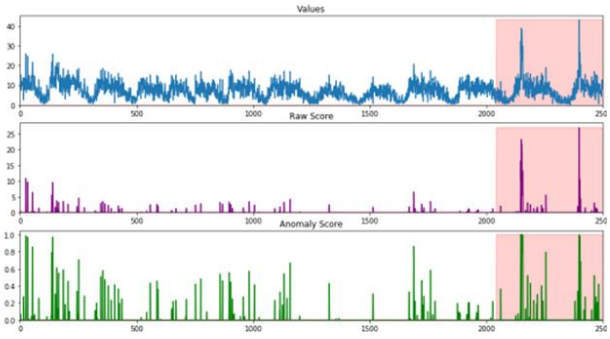


Рисунок 1 – Приклад детектора RPCA на циклічному часовому ряді

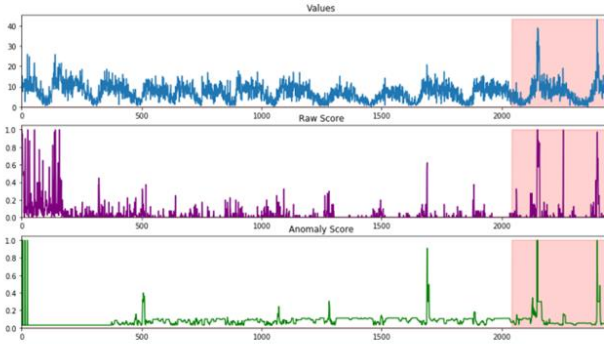


Рисунок 2 – Приклад детектора HTM на циклічному виході

10

Таблиця 1 – Вимірювання бінарної класифікації на наборі даних NAB за допомогою детектора аномалій HTM для різних порогових значень

ε	F1	Точність (precision)	Точність (accuracy)	P_частка	N_частка
0.0	0.306	0.192	0.192	1.000	0.000
0.1	0.247	0.314	0.708	0.226	0.774
0.2	0.132	0.484	0.802	0.028	0.972
0.3	0.104	0.403	0.805	0.010	0.990
0.4	0.093	0.412	0.807	0.006	0.994
0.5	0.097	0.372	0.807	0.003	0.997
0.6	0.129	0.356	0.807	0.003	0.997
0.6827	0.145	0.337	0.807	0.003	0.997
0.7	0.162	0.331	0.807	0.003	0.997
0.8	0.179	0.317	0.807	0.003	0.997
0.9	0.196	0.310	0.807	0.002	0.998
0.9545	0.195	0.309	0.807	0.002	0.998
0.9973	0.212	0.301	0.807	0.002	0.998

Таблиця 2 – Вимірювання бінарної класифікації на наборі даних NAB за допомогою детектора аномалій RPCA для різних порогових значень

ε	F1	Точність (precision)	Точність (accuracy)	P_частка	N_частка
0.0	0.306	0.192	0.192	1.000	0.000
0.1	0.170	0.294	0.653	0.229	0.771
0.2	0.163	0.320	0.676	0.195	0.805
0.3	0.152	0.335	0.718	0.140	0.860
0.4	0.161	0.343	0.752	0.099	0.901
0.5	0.170	0.360	0.760	0.085	0.915
0.6	0.164	0.376	0.771	0.069	0.931
0.6827	0.159	0.394	0.791	0.046	0.954
0.7	0.156	0.395	0.792	0.043	0.957
0.8	0.145	0.415	0.803	0.020	0.980
0.9	0.174	0.527	0.805	0.016	0.984
0.9545	0.129	0.601	0.807	0.014	0.986
0.9973	0.147	0.645	0.807	0.007	0.993

11

Таблиця 3 – Результат виявлення аномального вікна на наборі даних NAB за допомогою HTM Anomaly Detector для різних порогових значень

ε	Знайдено (total_found)	Пропущено (total_missed)	TPR загальний (total_tpr)
0.0	99	0	1.000
0.1	94	5	0.949
0.2	89	10	0.899
0.3	86	13	0.869
0.4	84	15	0.848
0.5	83	16	0.838
0.6	77	22	0.778
0.6827	74	25	0.747

Таблиця 4 – Результат виявлення аномального вікна на наборі даних NAB за допомогою RPCA для різних порогових значень

ε	Знайдено (total_found)	Пропущено (total_missed)	TPR загальний (total_tpr)
0.0	99	0	1.000
0.1	98	1	0.990
0.2	97	2	0.980
0.3	97	2	0.980
0.4	95	4	0.960
0.5	93	6	0.939
0.6	93	6	0.939
0.6827	93	6	0.939

12

Висновки

Як видно з результатів, і RPCA-детектор, і HTM-детектор добре справлялися з виявленням аномалій, і обидва алгоритми змогли працювати принаймні так само добре, як і поточні налаштування виявлення інцидентів для тестованих часових рядів. Обидва детектори змогли знайти всі відомі інциденти в наборі даних, і навіть при високих порогах ймовірності обидва алгоритми змогли знайти більшість позначених вікон аномалій. Алгоритм RPCA здається загалом більш чутливим, ніж HTM-детектор, однак, оскільки виявлення аномалій і чутливість детекторів також залежать від порогових значень, обидва алгоритми можна налаштувати на більшу чутливість за бажанням. Чутливість також залежить від набору даних, оскільки детектор HTM мав більше позначених значень, ніж детектор RPCA на наборі даних компанії, тоді як для набору даних NAB спостерігалася протилежна картина. Для набору даних компанії детектор RPCA зміг виявити 94% всіх вікон аномалій, при цьому рівень помилкових спрацьовувань склав 0,2%. Детектор HTM ніколи не досягав такого рівня продуктивності для набору даних компанії, маючи мінімальний рівень помилкових спрацьовувань 0,6% і знаходячи 88% вікон аномалій.