https://doi.org/10.31649/1997-9266-2025-182-5-75-82

УДК 517.444

П. В. Гринченко[1]
А. В. Бакурова[1]

ЗАСТОСУВАННЯ ВЕЙВЛЕТ-АНАЛІЗУ ДЛЯ ВАЛІДАЦІЇ ДАНИХ У СИСТЕМАХ ОЦІНКИ РИЗИКУ

[1]Національний університет «Запорізька політехніка»

Об'єктом дослідження є дві інформаційні системи. Перша система (S1) — це функціонуюча система, стан якої може бути нормальним або ненормальним, що описується набором параметрів системи в певний момент часу. Тобто існують часові ряди, що описують нормальну поведінку цієї системи та відхилення від неї. Друга система (S2) призначена для виявлення аномальної активності в першій системі. Виходи цієї системи виявлення аномалій представлені часовим рядом, кожне зі значень якого характеризує стан системи в певний момент часу двома параметрами: впевненістю (confidence) другої системи виявлення аномалій в оцінці стану першої системи та серйозністю (severity) аномалії, що виникла в першій системі.

У дослідженні детально розглянуто методологію застосування вейвлет-аналізу, що включає декомпозицію сигналу на коефіцієнти апроксимації та деталізації, реконструкцію сигналу, розрахунок кореляції між показниками впевненості та серйозністю виявленої аномалії системи, а також аналіз енергії компонентів для оцінки нестабільності роботи системи.

Запропонований підхід дозволяє встановити граничні значення для визначення нестабільності в роботі системи виявлення аномалій та розробити метрики якості її роботи. Особливу увагу приділено оцінюванню стабільності впевненості системи через енергію компонентів та визначення ступеня узгодженості між впевненістю та серйозністю через кореляційний аналіз.

Показано, як цей метод інтегрується з методологією оцінки ризику ACRAM (S3), забезпечуючи попередню валідацію даних, їхню фільтрацію, комплексну оцінку та валідацію результатів.

На основі тестових даних продемонстровано результати застосування розробленого алгоритму, зокрема, розрахунок енергії компонентів, кореляцію між впевненістю та серйозністю, а також встановлення порогів для визначення рівня стабільності системи (S2). Отримані результати показують, що система виявлення аномалій демонструє середній рівень стабільності, а кореляційний аналіз виявляє значну негативну кореляцію між впевненістю системи та серйозністю виявлених аномалій, що свідчить про зниження впевненості системи під час переходу до серйозніших станів.

Запропонований метод валідації даних є ефективним інструментом для підвищення надійності систем оцінювання ризику і може знайти застосування в різних галузях, що потребують обробки часових рядів та виявлення аномалій.

Ключові слова: вейвлет-аналіз, валідація даних, оцінка ризику, виявлення аномалій, енергія компонентів, кореляційний аналіз, вейвлет.

Вступ

Основним завданням дослідження є розробка алгоритму аналізу результатів роботи системи виявлення аномалій з метою оцінки стабільності та надійності її висновків. Особливість завдання полягає у використанні вейвлет-перетворення DB4 як інструменту для відокремлення нормальної поведінки системи від потенційних аномалій через аналіз різних частотних компонент сигналу.

Метою роботи є розробка методу оцінки надійності системи виявлення аномалій на основі аналізу змін «впевненості» і «серйозності» виявлених аномалій з використанням вейвлет-аналізу для валідації вхідних даних системи оцінки ризиків (S3) [1].

Подібні підходи до виявлення аномалій з використанням вейвлет-перетворень останнім часом знаходять широке застосування у сфері кібербезпеки. Наприклад, у роботі D.-W. Kim та ін. [2] застосовано дискретне вейвлет-перетворення для класифікації внутрішніх загроз на основі CERT-даних, що дозволило значно зменшити кількість хибнопозитивних спрацьовувань. H. G. Mohamed та ін. [3] запропонували вейвлет-нейронну мережу з оптимізацією Bird Swarm Algorithm для вияв-

лення атак в IoT-середовищах, досягнувши точності 99,64 % на датасеті UNSW-NB15. У роботі M. Wan та ін. [4] вейвлет-мережа застосована до SCADA-систем, де моделюється аномальна активність у протоколах Modbus через характеристику мережевої поведінки. А автори [5], R. Purohit та ін., поєднують безперервне вейвлет-перетворення з автоенкодером для реального часу виявлення аномалій у мережевому трафіку. Проте вказані підходи орієнтовані переважно на класифікацію подій або зменшення похибок виявлення, не охоплюючи аналіз стабільності самої системи виявлення. Для узагальнення ключових характеристик розглянутих підходів подано порівняльний аналіз у табличній формі.

Порівняльний аналіз аналогів запропонованого методу

Автори, рік опублікування	Тип	Сфера застосування	Що аналізується	Відмінність запропонованого методу
Kim et al., 2023	Discrete Wavelet Transform	внутрішні загрози	класифікація подій, зниження false positives	цей метод не враховує стабільність системи
Mohamed et al., 2023	Wavelet Neural Network	IoT, IDS	висока точність виявлення атак	фокус лише на точності, без аналізу впевненості
Wan et al., 2018	Wavelet Neural Network	SCADA / промислові мережі	виявлення аномалій у modbus-комунікаціях	не аналізує енергію або кореляцію параметрів
Purohit et al., 2025	Continuous Wavelet Transform + + Автоенкодер	мережевий трафік	виявлення аномалій у реальному часі	немає оцінки узгодженості висновків системи

Запропонований у цій роботі метод, на відміну від згаданих, фокусується саме на валідації достовірності висновків системи шляхом оцінки її стабільності у часі — через аналіз енергії вейвлет-компонентів сигналу впевненості та кореляції з рівнем серйозності виявлених аномалій.

Вхідні дані представлені результатами роботи інструмента виявлення аномалій робота:

– тимчасовий ряд значень впевненості у діапазоні [0,1];

– серйозність: OFF (аномалія відсутня), YELLOW, RED.

Метод розроблено з використанням вейвлет-перетворення Добеші 4-го порядку (DB4), яке дозволяє:

– відокремити нормальну поведінку системи та аномальні стани за допомогою аналізу різних частотних компонентів сигналу;

– розкласти сигнал на коефіцієнти деталізації та апроксимації, що дозволяє виявити нестабільність системи.

Основні етапи аналізу:

1. Застосування вейвлет-перетворення DB4;
2. Розрахунок енергії компонентів;
3. Статистичний аналіз;
4. Аналіз поведінки системи (S2);
5. Визначення рівня нестабільності (S2).

Критерії стабільності для комплексної оцінки продуктивності системи (S2):

– коефіцієнти кореляції між впевненістю системи та серйозністю аномалії в S2;

– результати аналізу енергії складових сигналу, а також оцінка рівня нестабільності S2 через порогові значення енергії.

Поведінка системи (S2) аналізується під час переходів між станом нормального функціонування та її аномальною поведінкою, що дозволяє здійснювати аналіз характеристик системи, таких як взаємозв’язок між рівнем впевненості у виявленні аномалій та ступенем серйозності.

Визначення рівнів нестійкості (S2) здійснюється на основі аналізу енергії компонентів, що дає можливість кількісно оцінити стійкість системи (S2) до зовнішніх збурень. До того ж виконується кореляційний аналіз впевненості та серйозності аномалій, що допомагає визначити ступінь узгодженості висновків системи.

Розроблений метод перевірки даних пройшов тестування на інтеграцію з системою оцінки ризиків ACRAM (S3) [1]. ACRAM — це методологія оцінки ризиків, яка розробляється в рамках проекту Telemetry і спрямована на адаптивне управління загрозами та прогнозний аналіз ризиків у складних інформаційних системах.

Основні сфери застосування розробленого методу в ACRAM:

– попередня перевірка даних, що дозволяє визначити достовірність вхідних даних та виявити періоди нестабільності (S2);

– фільтрація даних дозволить виключити з аналізу оцінки ризику періоди з високою нестабільністю;

– інтегрована оцінка може використовуватися в поєднанні з ACRAM для дворівневого аналізу ризиків;

– перевірка результатів оцінювання ризику підвищить точність висновків (S3).

Таким чином, розв'язання задачі аналізу стабільності виведення системи виявлення аномалій створює надійний фундамент для подальшої оцінки ризику, підвищуючи достовірність та обґрунтованість отриманих результатів.

Результати дослідження

Використання вейвлетів DB4 у цьому дослідженні зумовлено їхніми унікальними властивостями:

– компактний носій — дозволяє ефективно аналізувати локальні особливості сигналу;

– хороша частотна локалізація — вейвлети здатні розкласти сигнал на частотні складові без втрати інформації;

– ортогональність — забезпечує можливість точної реконструкції сигналу;

– гладкість — DB4 достатньо гладка для аналізу часових рядів [6].

Вейвлети четвертого порядку забезпечують баланс між часовою локалізацією, частотною роздільною здатністю та обчислювальною складністю. Вони особливо підходять для аналізу даних, пов'язаних з різкими змінами, що характерно для систем виявлення аномалій.

1. Застосування вейвлет-перетворення DB4

Для практичного використання вейвлет-перетворення необхідно визначити числові значення коефіцієнтів фільтрів, що лежать в основі вейвлетів Добеши-4. Запишемо систему рівнянь, яка дозволяє визначити ці коефіцієнти з урахуванням властивостей вейвлет-функцій.

Для знаходження коефіцієнтів фільтра $h[n]$, які визначають згортку та декомпозицію сигналу, застосовуються математичні умови, що забезпечують нормування, ортогональність та згладжування.

Розглянемо основні умови для коефіцієнтів $h[n]$.

Умова нормування (нормування для дискретних випадків) [7]

$$\sum_n |c_n|^2 = \|g\|^2 = 1, \tag{1}$$

де g — фільтр високих частот (для виділення аномальної поведінки); c_n — елементи фільтра g.

Умова ортогональності [8]

$$h[0]h[2] + h[1]h[3] = 0, \tag{2}$$

де h — фільтр низьких частот (для виділення нормальної поведінки).

Умова моментів (для $k = 0,1$) (3) [9], що виникає з нормування

$$\sum_n N^{h_n} = \sqrt{2}; \tag{3}$$

$h[0]\cdot 0 + h[1]\cdot 1 + h[2]\cdot 2 + h[3]\cdot 3 = 0$ — виникає з ортогональності.

Підставимо коефіцієнти, які визначені в фундаментальній праці Ingrid Daubechies [9].

У цьому дослідженні підставляємо значення ($h[0]...h[3]$) до відповідних рівнянь (зокрема рівняння моментів), щоб сформувати фільтр нижніх частот ($h[n]$) для декомпозиції сигналу, а також на основі цих значень побудуємо фільтр високих частот ($g[n]$). Ці фільтри є основою для подальшого вейвлет-перетворення та аналізу сигналу: $h[0] = 0,4829629131445341$;
$h[1] = 0,8365163037378079$; $h[2] = 0,2241438680420134$; $h[3] = -0,1294095225512604$.

Визначимо коефіцієнти фільтра g за формулою [9]

$$g_n = \left(\varphi_1 \varphi_{-1,n}\right) = (-1)^n h_{-n+1}. \tag{4}$$

Підставляємо h

$$g[0]=(-1)^0\cdot h[3]=-0{,}1294095225512604;\quad g[1]=(-1)^1\cdot h[2]=-0{,}2241438680420134;$$
$$g[2]=(-1)^2\cdot h[1]=0{,}8365163037378079;\quad g[3]=(-1)^3\cdot h[0]=-0{,}4829629131445341.$$

Ці коефіцієнти забезпечують:

– умову ортогональності між масштабувальною функцією та вейвлетом;

– компактний носій (ненульові значення лише для $n=0,1,2,3$);

– точну реконструкцію сигналу у разі декомпозиції та відновлення;

– достатню гладкість для аналізу сигналів з різкими змінами.

Саме через ці властивості вейвлет Добеші-4 є придатним для аналізу аномалій у цьому випадку, оскільки він може ефективно відокремлювати нормальну поведінку від потенційних аномалій.

	№	Базова впевненість	Зміни впевненості
2	1	0,358	-0,404
3	2	0,42	-0,333
4	3	0,38	-0,413
5	4	0,434	-0,337
6	5	0,409	-0,426

Фрагмент вхідних даних

Виконуємо декомпозицію сигналу на коефіцієнти апроксимації та деталізації (вихідна впевненість з набору тестових даних, стовпець «базова впевненість» рис.): $c_1 = 0{,}358$;

$c_2 = 0{,}420$; $c_3 = 0{,}380$; $\ldots c_{100}=0{,}382$ (c_k — це коефіцієнти апроксимації).

Коефіцієнти h: $h_1=0{,}4829629131445341$; $h_2=0{,}8365163037378079$;

$h_3=0{,}2241438680420134$; $h_4=-0{,}1294095225512604$.

Розрахунок коефіцієнтів апроксимації виконується за формулою (використовуючи ковзне вікно з циклічним продовженням) [10]

$$y[n]=\sum_k a_k y_k[n], \tag{5}$$

де $y[n]$ — апроксимований сигнал; $y_k[n]$ — інформація про впевненість системи; a_k — коефіцієнти апроксимації.

У контексті вейвлет-аналізу використовується апроксимація з кроком 2, тобто зі 100 вхідних значень отримуємо 50 значень апроксимації. Таке зменшення розмірності зумовлене децимацією, яка є складовою алгоритму вейвлет-перетворення та забезпечує збереження ключової інформації сигналу на нижчому рівні представлення:

1. *Децимація (дозволяє зменшити надмірність подання сигналу):*

– на кожному рівні розкладання довжина сигналу зменшується у 2 рази, це не означає втрату інформації, а є частиною алгоритму;

– зі 100 значень отримуємо 50 коефіцієнтів апроксимації.

2. *Процес обчислення.*

Для 100 вхідних значень: $a_1=h_1\cdot c_1+h_2\cdot c_2+h_3\cdot c_3+h_4\cdot c_4$;

$a_2=h_1\cdot c_3+h_2\cdot c_4+h_3\cdot c_5+h_4\cdot c_6$; $\ldots a_{50}=h_1\cdot c_{99}+h_2\cdot c_{100}+h_3\cdot c_1+h_4\cdot c_2$ (a_k — вхідні значення сигналу).

– кожна апроксимація враховує 4 послідовні значення (довжина фільтра);

– зміщення на 2 позиції забезпечує необхідне проріджування;

– при цьому зберігається вся суттєва інформація про тренд сигналу.

Тому досить визначити 50 значень апроксимації, це забезпечує коректне представлення сигналу на нижчому рівні деталізації.

Далі на цьому етапі (визначаємо від 1 до 50): $a_1 = 0{,}553188306$; $a_2 = 0{,}582861547$;

$a_{50} = 0{,}579012345$; $\overline{a}=28{,}93456789/50=0{,}578691358$.

2. Розрахунок енергії компонентів

Проводимо реконструкцію сигналу. Для реконструкції сигналу з апроксимації використовуємо формулу [11]

$$x[n] = \sum_{j=1}^{+\infty} \sum_{k \in Z} c_{j,k} \tilde{h}_j \left[n - 2^j k \right] + \sum_{k \in Z} b_{j,k} \tilde{g}^j \left[n - 2^j k \right], \quad (6)$$

де $x[n]$ — реконструйований сигнал; $c_{j,k}$, $b_{j,k}$ — коефіцієнти розкладання;

$h = [0,4829629131445341,\ 0,8365163037378079,\ 0,2241438680420134,\ -0,1294095225512604]$;

$g = [-0,1294095225512604, -0,2241438680420134,\ 0,8365163037378079,\ -0,4829629131445341]$.

Це дає нам реконструйовані значення сигналу, які використовуються для подальшого аналізу.

Тепер розрахуємо енергію компонентів. У контексті вейвлет-перетворення можна сказати, що сигнал є сумою різних компонентів, кожна з яких відображає певні характеристики сигналу на різних масштабних рівнях (або частотах). Ці компоненти містять інформацію про високочастотні та низькочастотні зміни сигналу.

Сума енергії цих компонентів дорівнює загальній енергії вихідного сигналу, а це означає, що вейвлет-перетворення є енергоощадним процесом.

У цьому випадку є два способи отримати компоненти (високочастотні компоненти):

1. Безпосередньо з декомпозиції;
2. Через реконструкцію (наш поточний спосіб).

Використано другий спосіб, тому що:

1) у разі декомпозиції отримуємо лише 50 коефіцієнтів компонентів (через децимацію);
2) у разі реконструкції та подальшому відніманні отримуємо 100 значень компонентів, що:
   – відповідає вихідній розмірності сигналу;
   – дозволяє точно зіставити кожне відхилення з конкретним моментом часу;
   – дає можливість аналізувати компоненти у тому масштабі, як і вихідний сигнал.

Тобто використання реконструйованого сигналу дозволяє зберегти тимчасову локалізацію компонентів та працювати у вихідному масштабі даних.

Вейвлет-аналіз дозволяє розрахувати енергію компонентів — метрику, яка відображає рівень нестабільності системи [12] за формулою

$$E_x = \sum_{n=-\infty}^{\infty} |x(n)|^2, \quad (7)$$

де x — дійсне число.

Розрахунок компонентів деталізації для перших трьох значень (d_n):

$d_1 = 0,358 - 0,80530155814 = -0,44730155814$; ($x_1 = 0,358$ — базова впевненість;

$\hat{x} = 0,80530155814$ — реконструйований сигнал); $d_2 = -0,333$; $d_3 = -0,413$.

3. Статистичний аналіз

Встановимо пороги за допомогою статистичних методів. Спочатку потрібно отримати енергію компонентів для різних часових вікон. Розіб'ємо наші 100 значень на вікна (по 20 значень) та визначимо енергію для кожного вікна.

Разом отримуємо 5 вікон.

Для кожного вікна енергія визначається як сума квадратів компонентів (формула (7)), $E = \sum \left(detail^2 \right)$:

вікно 1 (перші 20 значень): $E_1 = (-0,404)^2 + (-0,333)^2 + (-0,413)^2 + ... + (-0,439)^2 = 2,968$;

вікно 2 (значення 21—40): $E_2 = (-0,381)^2 + (-0,384)^2 + (-0,434)^2 + ... + (-0,313)^2 = 2,872$;

вікно 3 (значення 41—60): $E_3 = (-0,289)^2 + (-0,450)^2 + (-0,507)^2 + ... + (-0,406)^2 = 3,12$;

вікно 4 (значення 61—80): $E_4 = (-0,423)^2 + (-0,459)^2 + (-0,333)^2 + ... + (-0,443)^2 = 2,946$;

вікно 5 (значення 81—100): $E_5 = (-0,342)^2 + (-0,431)^2 + (-0,377)^2 + ... + (-0,374)^2 = 2,930$.

Визначимо статистичні характеристики вікон.

Середнє значення енергії вікон розраховується за такою формулою:

$$Mean(E) = (2,968 + 2,872 + 3,124 + 2,946 + 2,930)/5 = 2,968.$$

Визначимо стандартне відхилення (0,0894) [13]

$$STD = \sqrt{\frac{\sum (X - \overline{X})^2}{n}}, \tag{8}$$

де X — індивідуальні значення енергетичного вікна, розраховані як сума квадратів на основі часових рядів енергій компонентів; $\overline{X}$ — середнє значення віконної енергії; n — кількість вікон.

Визначаємо порогові значення нестабільності S2 за формулами (7) і (8)

$$E_{low} = mean - stdDev = 2,968 - 0,0894 = 2,879;$$

$$E_{mid} = mean + stdDev = 2,968 + 0,0894 = 3,057;$$

$$E_{high} = mean + 2 \cdot stdDev = 2,968 + 2 \cdot 0,0894 = 3,147;$$

$$E_{total} = 14,840.$$

4. Аналіз поведінки системи (S2)

Проводимо розрахунок кореляції між впевненістю та серйозністю за такою формулою [14]:

$$r_{XY} = \frac{\text{cov}_{XY}}{\sigma_X \sigma_Y} = \frac{\sum (X - \overline{X})(Y - \overline{Y})}{\sqrt{\sum (X - \overline{X})^2 \sum (Y - \overline{Y})^2}}, \tag{9}$$

де X, Y — часовий ряд, кожне значення якого характеризує стан системи в певний момент часу за двома параметрами: впевненості S2 виявлення аномалії у разі оцінювання стану S1 та серйозності, що виникла в S1; $\overline{X}, \overline{Y}$ — середні значення часового ряду; r — коефіцієнт кореляції.

Приклад розрахунку переходу стану системи від Off (немає аномалії) до Red (аномальна поведінка):

Впевненість (x): [0,400, 0,235, 0,213] (перехід від OFF до RED);

Серйозність (y): [0,0, 1,0, 1,0] (перехід від OFF до RED);

$$\mu x = (0,400 + 0,235 + 0,213)/3 = 0,283; \quad \mu y = (0,0 + 1,0 + 1,0)/3 = 0,667;$$

$$x - \mu x = [0,117, -0,048, -0,070]; \quad y - \mu y = [-0,667, 0,333, 0,333].$$

$$\sum ((x - \mu x)(y - \mu y)) = (0,117 \cdot -0,667) + (-0,048 \cdot 0,333) + (-0,070 \cdot 0,333) = -0,078 + (-0,016) +$$

$$+ (-0,023) = -0,117 \sum (x - \mu x)^2 = (0,117)^2 + (-0,048)^2 + (-0,070)^2 = 0,0137 + 0,0023 + 0,0049 = 0,0209;$$

$$\sum (y - \mu y)^2 = (-0,667)^2 + (0,333)^2 + (0,333)^2 = 0,445 + 0,111 + 0,111 = 0,667;$$

$$r = -0,117 / \sqrt{(0,0209 \cdot 0,667)} = -0,117 / \sqrt{0,0139} = -0,117 / 0,118 = -0,992.$$

Підсумкова кореляція для 100 значень: –0,853.

Негативна кореляція (–0,853) між впевненістю системи та серйозністю означає, що зі збільшенням серйозності виявленої аномалії впевненість системи (S2) знижується. Цей результат підтверджує, що система адекватно реагує на аномалії, знижуючи впевненість у разі переходу до критичніших станів.

5. Визначення рівня нестабільності (S2)

Порівнюємо з порогами на вікно

$$\overline{E}_{window} = 14,840/5 = 2,968, \quad 2,879 < 2,968 < 3,057.$$

Середня енергія на вікно (2,968) між нижнім і середнім порогами, що вказує на середній рівень стабільності системи (S2). У ході дослідження вдалося визначити проміжки часу, в які система демонструвала нестабільність. Це важливо для подальшого вдосконалення методу виявлення аномалій (S2).

Висновки

У ході дослідження розроблений метод оцінки стійкості системи виявлення аномалій довів свою ефективність, дозволяючи оцінити стійкість системи виявлення аномалій (S2). В результаті аналізу підтверджено, що використання вейвлет-перетворення DB4 ефективно вирішує проблему розділення нормальної та аномальної поведінки системи, забезпечуючи точне виявлення нестабільних станів. Одним з ключових результатів дослідження стало виявлення значної негативної кореляції (–0,853) між впевненістю системи та серйозністю виявлених аномалій. Це означає, що у разі переходу до критичніших станів рівень довіри системи знижується. Додатковий розрахунок енергії компонентів показав, що (S2) демонструє середній рівень стабільності. Визначені порогові значення дозволили класифікувати часові інтервали нестабільності, що дає можливість у подальшому враховувати їх під час оцінювання надійності системи.

Розроблений метод може бути корисним для перевірки стабільності роботи систем виявлення атак, у фінансовому моніторингу — для виявлення періодів коливань довіри до моделей, в промисловості — для контролю стану обладнання, в медицині — для фільтрації нестабільних ділянок біомедичних сигналів, та в багатьох інших сферах людської діяльності.

Метою подальших досліджень є інтеграція запропонованого методу оцінювання стабільності системи виявлення аномалій із системою ACRAM (S3).

Ця робота виконана за підтримки проекту Horizon Europe TELEMETRY.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

[1] V. Lytvyn, et al., “Fuzzy logic-based methodology for building access control systems based on fuzzy logic,” in *6th International Workshop on Modern Data Science Technologies*, May, 31 — June, 1, 2024. [Electronic resource]. Available: https://ceur-ws.org/Vol-3723/paper7.pdf .

[2] D.-W. Kim, G.-Y. Shin, and M.-M. Han, “Anomaly Detection Based on Discrete Wavelet Transformation for Insider Threat Classification,” *Computer Systems Science and Engineering*, no. 46 (1), pp. 153-164, 2023. https://doi.org/10.32604/csse.2023.034589 .

[3] H. G. Mohamed, et al., “Optimal Wavelet Neural Network-Based Intrusion Detection in Internet of Things Environment,” no. 75 (2), pp. 4467-4483, 2023. https://doi.org/10.32604/cmc.2023.036822 .

[4] M. Wan, Y. Song, Y. Jing, and J. Wang, “Function-Aware Anomaly Detection Based on Wavelet Neural Network for Industrial Control Communication,” *Security and Communication Networks*, Hindawi Limited, 5103270, 2018. https://doi.org/10.1155/2018/5103270 .

[5] R. Purohit, S. Kumar, S. Sayyad, and K. Kotecha, “Time-frequency analysis and autoencoder approach for network traffic anomaly detection,” *MethodsXM*, vol. 14, 103228, 2025. https://doi.org/10.1016/j.mex.2025.103228 .

[6] T. Rajesh, “Image Reconstruction Using Wavelet Transform with Extended Fractional Fourier Transform,” *Blekinge Institute of Technology*, 2014. [Electronic resource]. Available: https://www.diva-portal.org/smash/get/diva2:829956/FULLTEXT01.pdf .

[7] Н. К. Смоленцев, *Основи теорії вейвлетів. Вейвлети в MATLAB*, вид 4-те, с. 102, 2014. [Електронний ресурс]. Режим доступу: https://surl.li/wnwoik .

[8] *Ортогональність*. [Електронний ресурс]. Режим доступу: https://uk.wikipedia.org/wiki/%D0%9E%D1%80%D1%82%D0%BE%D0%B3%D0%BE%D0%BD%D0%B0%D0%BB%D1%8C%D0%BD%D1%96%D1%81%D1%82%D1%8C . Дата звернення 18.03.2025.

[9] I. Daubechies, *Ten Lectures of Wavelets*, pp. 167-204, 1992. [Electronic resource]. Available: https://jqichina.wordpress.com/wp-content/uploads/2012/02/ten-lectures-of-waveletsefbc88e5b08fe6b3a2e58d81e8aeb2efbc891.pdf .

[10] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*. Second edition, p. 41, 1998. [Electronic resource]. Available: https://ru.scribd.com/document/509418212/Discrete-Time-Digital-Signal-Processing-Oppenheim-Schafer-Buck .

[11] O. Rioul, and P. Duhamel, “Fast Algorithms for Discrete and Continuous Wavelet Transforms,” *IEEE Transactions on information theory*, vol. 38, no. 2, p. 4, 1992. https://doi.org/10.1109/18.119724 .

[12] *Energy of a signal*. [Electronic resource]. Available: https://www.gaussianwaves.com/2013/12/power-and-energy-of-a-signal/ . Accessed: 18.03.2025.

[[1]13] *Стандартне відхилення*. [Електронний ресурс]. Режим доступу: https://berg.com.ua/indicators-overlays/stdev . Дата звернення 18.03.2025.

[14] *Кореляція*. [Електронний ресурс]. Режим доступу: https://uk.wikipedia.org/wiki/%D0%9A%D0%BE%D1%80%D0%B5%D0%BB%D1%8F%D1%86%D1%96%D1%8F . Дата звернення 18.03.2025.

Рекомендована кафедрою автоматизації та інтелектуальних інформаційних технологій ВНТУ

Стаття надійшла до редакції 7.07.2025

Гринченко Павло Вікторович — аспірант кафедри системного аналізу та обчислювальної математики, e-mail: phrynchenko@ukr.net ;

Бакурова Анна Володимирівна — д-р екон. наук, професор, професор кафедри системного аналізу та обчислювальної математики, e-mail: abaka@zp.edu.ua .

Національний університет «Запорізька політехніка», Запоріжжя

P. V. Hrynchenko[1]
A. V. Bakurova[1]

Application of Wavelet Analysis for Data Validation in Risk Assessment Systems

[1]National University “Zaporizhzhia Polytechnic”

The object of the study is two information systems. The first system (S1) is a functioning system, the state of which can be normal or abnormal, described by a set of system parameters at a certain point in time. That is, there are time series describing the normal behavior of this system and deviations from it. The second system (S2) is designed to detect abnormal activity in the first system. The outputs of this anomaly detection system are presented by a time series, each of the values of which characterizes the state of the system at a certain point in time by two parameters — the confidence of the second anomaly detection system in assessing the state of the first system and the severity of the anomaly that occurred in the first system.

The study examines in detail the methodology for applying wavelet analysis, which includes the decomposition of the signal into approximating and detailing coefficients, signal reconstruction, calculation of the correlation between confidence and the severity system, as well as the analysis of the energy of the details to assess the instability of the system.

The proposed approach allows to set threshold values for determining instability in the operation of the anomaly detection system and develop metrics for the quality of its operation. Particular attention is paid to assessing the stability of the system's confidence through the energy of details, which characterizes the intensity of high-frequency changes in the signal, and determining the degree of consistency between confidence and severity through correlation analysis.

It is shown how this method integrates with the ACRAM risk assessment methodology, providing preliminary data validation, their filtering, comprehensive assessment and validation of results.

Based on the test data, the results of applying the developed algorithm are demonstrated, including the calculation of the energy of details, the correlation between confidence and severity, as well as setting thresholds to determine the level of system stability. The results obtained show that the anomaly detection system demonstrates an average level of stability, and correlation analysis reveals a significant negative correlation between the confidence of the system and the severity of the detected anomalies, which indicates a decrease in the confidence of the system during the transition to more serious states.

The proposed data validation method is an effective tool for increasing the reliability of risk assessment systems and can be used in various industries that require time series processing and anomaly detection.

Keywords: wavelet analysis, data validation, risk assessment, anomaly detection, component energy, correlation analysis, wavelet.

Hrynchenko Pavlo V. — Post-Graduate Student of the Chair of System Analysis and Computational Mathematics, e-mail: phrynchenko@ukr.net ;

Bakurova Anna V. — Dr. Sc. (Econ.), Professor, Professor of the Chair of System Analysis and Computational Mathematics, e-mail: abaka@zp.edu.ua