
INFORMATION SECURITY

DOI 10.20535/2411-1031.2021.9.2.249887

УДК 004(942:056.53)

ВЯЧЕСЛАВ РЯБЦЕВ,
ДМИТРО ШАРАДКІН,
ЮРІЙ КЛЯТ

ПОРІВНЯЛЬНЕ ДОСЛІДЖЕННЯ АЛГОРИТМІВ ВИЯВЛЕННЯ ТОЧОК ЗМІН В РЕГРЕСІЙНИХ МОДЕЛЯХ ЧАСОВИХ РЯДІВ

Основна мета моніторингу будь яких об'єктів та процесів реального світу – від технічних пристроїв до пандемій – є виявлення внутрішніх змін або змін поведінки об'єкта, що вивчається. Поведінка об'єкта може змінюватися з часом через його реакцію на зовнішні події, а також під впливом законів розвитку самого об'єкта. Проблеми виникають у випадках, коли такі внутрішні зміни не спостерігаються безпосередньо. У такому разі отримати інформацію про внутрішні зміни що відбулися можливо лише за допомогою аналізу часових рядів значень параметрів об'єктів, вимірювання яких можливо здійснити технічно та/або організаційно. Проаналізовано алгоритми, що базуються на описі рядів у вигляді лінійних регресійних моделей та призначених для виявлення змін у таких часових рядах. Припускається, що у разі виникнення змін часовий ряд до і після точки в якій такі зміни відбулися описується моделями, які неможливо вважати ідентичними. Розглядається як відомі алгоритми аналізу ідентичності моделей, так і алгоритми, які раніше не описувалися в літературі. Зокрема, алгоритм, що ґрунтується на аналізі змін значень коефіцієнтів регресійних моделей; на аналізі довірчого інтервалу передбачених за регресійними моделями значень ряду; на аналізі залишків регресії з використання нерівності Гефдінга; алгоритми Чоу порівняння дисперсій залишків моделей; алгоритм порівняння розподілів залишків з використанням критерію Колмогорова-Смирнова. Різноманітність характеру зміни поведінки об'єктів моніторингу, в основі яких лежить непередбачуваність причин, що їх викликали, не дозволяють вибрати єдиний універсальний алгоритм виявлення точок змін і робить порівняльну оцінку різних алгоритмів актуальним інженерним і технічним завданням. Аналіз алгоритмів пропонується виконувати на множині еталонних зразків змін, що найбільш характерні для часових рядів параметрів об'єктів у різних прикладних галузях. Ефективність алгоритмів оцінюється за допомогою статистичного експерименту як з точки зору їх здатності виявити саму зміну, так і за порівнянням кількості помилок хибного виявлення та пропуску реальних змін, що генеруються кожним з алгоритмів. Отримані результати можуть стати основою подальшого дослідження, зокрема при побудові ансамблю алгоритмів виявлення точок змін поведінки об'єктів моніторингу.

Ключові слова: часові ряди, аномалії, виявлення точок зміни, прогнозні алгоритми виявлення змін, регресійні моделі.

Постановка проблеми. В наш час практично будь-який об'єкт реального світу допускає можливість автоматизованого збору інформації про свій стан. За цієї інформацією можливо організувати моніторинг стану об'єкта, часто – в режимі реального часу. Кількість об'єктів які підлягають моніторингу, а також параметрів що при цьому враховуються, постійно збільшується. У зв'язку з цим особливу актуальність набуває задача автоматизації провадження постійного моніторингу та, – що навіть важливіше, – своєчасного виявлення та оповіщення про виникнення нестандартних змін у поведінці об'єкта, за яким ведеться спостереження [1]. Такі нестандартні зміни можуть бути пов'язані з передаварійними та аварійними ситуаціями для технічних пристроїв [2], з передболісними станами при відслідковуванні стану здоров'я пацієнта в медичних системах [3], з виникненням ажіотажних явищ на фінансових ринках [4], з

катаклізмами що насуваються в метеоспостереженнях, з виникненням лісових пожеж при спостереженні за лісовими масивами [5]. Атаки на комп'ютерні мережі зазвичай також викликають різкі зміни параметрів мережного трафіку, а швидке виявлення таких змін здатне істотно підвищити рівень інформаційної безпеки та знизити втрати від шкідливих дій, що завдаються ними [6], [7].

У всіх перелічених та багатьох інших прикладних галузях автоматизація процесу виявлення змін має велике практичне значення для аналітиків та обслуговуючого персоналу. Нестандартна поведінка об'єктів моніторингу зазвичай відбивається у суттєвих змінах в часових рядах значень їх параметрів. В науковій та технічній літературі такі зміни часто називають “аномаліями” [1]. Один різновид аномалій проявляється у вигляді “викидів”, тобто різких і зазвичай нетривалих змін окремих значень параметрів об'єкта. Інший різновид пов'язаний з довгостроковою або навіть постійною зміною значень параметрів. В одних ситуаціях зміни можуть відбуватися раптово, тобто стохастичні властивості часового ряду залишаються постійними протягом певного часу, а потім раптово набувають іншого значення. В інших ситуаціях зміни відбуваються поступово: значення параметрів залишаються більш-менш стабільними протягом певного часу, а після досягнення деякого, зазвичай заздалегідь невідомого, моменту починають змінюватися. У всіх випадках часто вживається узагальнювальний термін “зміна моделі ряду”, а момент часу (або відповідну йому точку часового ряду значень параметра) в якій така зміна відбувається – точкою зміни моделі ряду (надалі – “точкою зміни”). На рис. 1 наведено приклади деяких різновидів змін що найчастіше спостерігаються в багатьох прикладних галузях і можуть бути використані для дослідження відповідних алгоритмів виявлення точок змін [8], [9].

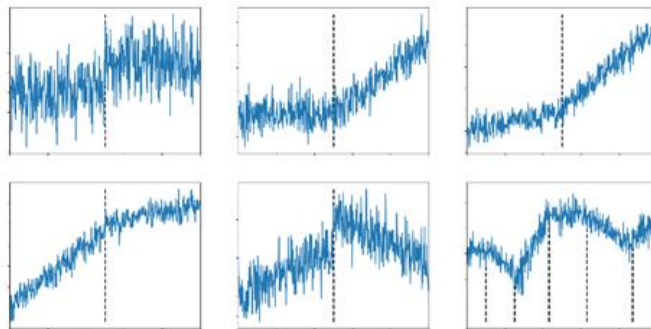


Рисунок 1 – Приклади зміни моделей часових рядів. Верхній ряд, зліва направо: зміна параметра положення; поява тренду; прискорення тренду; нижній ряд, зліва направо: сповільнення тренду; зміна тренду на протилежний; приклад розбиття ряду на сегменти

На теперішній час запропоновано чимало різноманітних методів, направлених на пошук розв'язання задачі виявлення точок зміни у часових рядах. Здебільшого ці методи сконцентровані на виявленні одного з можливих різновидів аномалій і виявляються недостатньо ефективними у ситуаціях, коли різновид аномалій наперед невідомий. Хоча побудова універсального алгоритму швидше за все ніколи не буде здійснена, пошук та аналіз ефективності алгоритмів, здатних виявити точки зміни для максимально широкого спектра аномалій, що зустрічаються в задачах моніторингу об'єктів різноманітної природи, є надзвичайно актуальним.

Аналіз останніх досліджень та публікацій. Існує великий обсяг робіт як з теоретичних, так і з практичних аспектів виявлення точок зміни. Перші роботи у цій галузі сягають 50-х років [3]. Ціль полягала в пошуку зсуву в середньому значенні незалежних і однаково розподілених за гаусовим законом випадкових змінних для задач автоматизації промислового контролю якості. Розвиток цих ідей [10] дозволив узагальнити результати на основі застосування методу максимальної правдоподібності для визначення точки зміни в ситуаціях, коли дані підпорядковуються довільним відомим ймовірнісним розподілам, що належать до

одного і того ж їх параметричного сімейства. В [11] досліджується метод виявлення точок зміни на основі аналізу відносин між функціями щільності розподілу. В [12], [13] вивчаються методи виявлення змін, які можуть ідентифікувати раптові зміни потоку даних за допомогою аналізу гістограм даних при послідовному спостереженні, а тому числі і такі, що не вимагають інформації щодо конкретних законів розподілу.

За останні роки, у зв'язку з розвитком як продуктивності комп'ютерних засобів, так й розширення спектра об'єктів, за якими необхідно здійснювати моніторинг, інтерес до проблеми суттєво зріс. Пропонуються та активно досліджуються нові підходи до її вирішення як на основі статистичних методів, так і з використанням останніх досягнень у галузі машинного навчання та штучних нейронних мереж. У ряді робіт розглядається виявлення змін за допомогою нейромереж [14], зокрема – спеціального їх різновиду, нейромереж типу “автоенкодер” [15].

Теоретичні основи виявлення точок зміни представлені у ряді монографій [6], [16] та дисертацій [5], [17], [18]. Докладну історію розвитку та огляди поточного стану дослідження у цьому напрямі можна знайти у роботах [19], [20], [21].

Аналіз робіт, присвячених застосуванню методів виявлення змін в різних прикладних галузях показує, що одна з основних проблем при застосуванні таких методів полягає в тому, що зазвичай розподіли значень ряду ні до, ні після точки зміни невідомі. Це з одного боку позначається на відповідному зниженні ефективності застосування теоретично обґрунтованих алгоритмів, і з іншого боку змушує аналізувати алгоритми не тільки суто теоретичними методами, але й використовуючи принципи статистичного моделювання [1], [6], [16].

Метою дослідження є аналіз алгоритмів виявлення точок змін у часових рядах, заснованих на використанні прогнозних регресійних моделей. На підставі застосування формальних критеріїв ефективності проводиться порівняння як відомих, так і оригінальних алгоритмів цього класу при виявленні обраного набору різновидів змін, а також аналіз ефективності зазначених алгоритмів в залежності від затримки моменту виникнення зміни відносно моменту початку спостереження.

Виклад основного матеріалу дослідження. Ключова ідея, що лежить в основі методів виявлення точок зміни полягає у порівнянні двох послідовних вибірок значень часового ряду: еталонного вікна та поточного вікна. При цьому еталонне вікно містить фрагмент часового ряду старіший щодо фрагмента поточного вікна. Виявлення змін у потокових даних при застосуванні двовіконних методів будується на основі формалізації відмінностей в описі даних еталонного та поточного вікон.

Дамо математичне формулювання задачі. Нехай $Y = \{y_t\}$ – послідовність n значень часового ряду ($n < \infty$), отриманих (виміряних) у моменти часу t . Виявлення точки зміни можна сформулювати як проблему перевірки та вибору між альтернативами гіпотезами:

- нульова гіпотеза H_0 : “Жодних змін у статистичних характеристиках значень часового ряду Y не відбувається”;
- альтернативна гіпотеза H_A : “Існує момент часу k , $0 < k \leq n$, такий, що статистичні характеристики значень часового ряду $Y' = \{y_1, y_2, \dots, y_{k-1}\}$ істотно відрізняються від статистичних характеристик значень часового ряду $Y'' = \{y_k, y_{k+1}, \dots, y_n\}$ ”.

Крім того, має бути попередньо зафіксовано значення рівня значущості α – прийнятна для даного дослідження ймовірність того, що вірна гіпотеза H_0 буде помилково відкинута. Таким чином, задача зводиться до вибору міри відмінності між даними еталонного і поточного вікон та конструювання відповідного формального критерію вибору гіпотез. Якщо зазначена відмінність за обраним критерієм виявляється більшою за деяке заздалегідь задане порогове значення, то гіпотеза H_0 відхиляється з рівнем значущості α , тобто фіксується факт наявності зміни в поведінці об'єкту. В іншому випадку гіпотеза про відсутність змін H_0 не відхиляється, тобто фіксується факт відсутності зміни в аналізованих даних.

Як критерії відмінності можуть виступати найпростіші статистичні показники (наприклад t-критерій Стьюдента для оцінки відмінності в середніх значеннях, F-критерій Фішера для оцінки відмінності в дисперсіях, відповідні їх непараметричні аналоги та інші), а також складніші показники, що оцінюють відмінності більш комплексно (наприклад, на основі критерію Колмогорова-Смирнова для оцінки відмінності в розподілах) [6], [20], [22]. Ці методи є найбільш глибоко вивченими, однак, як показує практика, у реальних прикладних галузях не завжди здатні забезпечити прийнятний рівень ефективності виявлення змін [1], [16].

Інша група методів для оцінки відмінності використовує моделі, що статистично описують поведінку часового ряду. Основне припущення покладене в основу цих методів полягає в тому, що якщо в об'єкті не відбулися будь-яких внутрішніх змін, модель яка адекватно описує дані часового ряду значень параметрів в еталонному вікні залишиться адекватною і для даних, отриманих у поточному вікні. В іншому випадку, якщо дані поточного вікна неможливо з обраною точністю описати раніше побудованою моделлю, подія інтерпретується як можлива зміна поведінки об'єкта моніторингу.

Як одна з моделей при зазначеному типі детектування змін досить часто використовується статистична регресійна модель, а сама процедура реалізується у вигляді наступних послідовних кроків:

- збір інформації (значень часового ряду) за проміжок часу обраний як довжина еталонного вікна;
- побудова регресійної моделі, що описує поведінку часового ряду;
- за допомогою отриманої на попередньому кроці регресійної моделі побудова прогнозу значень часового ряду, очікуваних у майбутній період (у поточному вікні);
- накопичення даних поточного вікна і перевірка адекватності моделі, побудованої на попередніх кроках для цих даних.

При використанні алгоритмів на основі регресійної моделі, модель часового ряду представляється у вигляді:

$$y = f(x) + \varepsilon_t, \quad (1)$$

де ε_t – випадкові, нормально розподілені значення, що є джерелом випадковості, причому вважається, що математичне очікування $E[\varepsilon_t] = 0$, а дисперсія $D[\varepsilon_t]$ точно невідома, але може бути з достатньою точністю оцінена за вибірковими значеннями.

У найпростішому випадку функція $f(x)$ є функцією від значення часових міток, тобто моментів часу, у яких виконувався вимір чергового значення послідовності, $f(x) = f(t)$. Для випадків еквідистантних рядів $f(t)$ задається послідовністю номерів відліків значень часового ряду $f(x) = f(i), i = 1, 2, \dots$. Для авторегресійних моделей $f(t)$ є функцією від значень ряду в попередні моменти часу $f(t) = f(y_{t-1}, y_{t-2}, \dots)$. Можуть застосовуватися і інші способи визначення функцій $f(t)$.

У даній роботі для опису $f(t)$ буде використовуватися лінійна регресійна модель:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t, \quad (2)$$

яка за наявності точки зміни ряду модифікується наступним чином:

$$y_t = \begin{cases} \beta_0' + \beta_1' x_t + \varepsilon_t, & 1 \leq t < k; \\ \beta_0'' + \beta_1'' x_t + \varepsilon_t, & k \leq t \leq n \end{cases} \quad (3.1)$$

$$(3.2)$$

Нижче розглядаються декілька алгоритмів, що використовують описану загальну схему.

Алгоритм детектування точки зміни на основі аналізу коефіцієнтів регресійної моделі (АКРМ-алгоритм). У цьому алгоритмі в якості параметрів порівняння використовуються коефіцієнти моделі (3.1) та (3.2). Відомо (наприклад, [22]), що для лінійної регресійної моделі ряду, в якій значення параметрів β_0 і β_1 невідомі, але можуть бути оцінені (зокрема – за методом найменших квадратів) як b_0 і b_1 відповідно, статистики

$$t_0 = \frac{\beta_0 - b_0}{S_{b_0}} \text{ та } t_1 = \frac{\beta_1 - b_1}{S_{b_1}}, \quad (4)$$

що відповідають t -розподілу Стьюдента з $(m-2)$ ступенями свободи, m – кількість значень ряду. В даному випадку $m = k - 1$.

Значення S_{b_0} і S_{b_1} визначаються наступним чином:

$$S_{\beta_0}^2 = S^2 \frac{\sum_{i=1}^{k-1} x_i^2}{(k-1) \sum_{i=1}^{k-1} (x_i - \bar{x})^2} = S^2 \left(\frac{1}{k-1} + \frac{\bar{x}^2}{\sum_{i=1}^{k-1} (x_i - \bar{x})^2} \right),$$

$$S_{\beta_1}^2 = \frac{S^2}{\sum_{i=1}^{k-1} (x_i - \bar{x})^2},$$

$$S^2 = \frac{1}{k-3} \sum_{i=1}^{k-1} (\hat{y}_i - y_i)^2,$$

де \hat{y}_i – значення, що отримуються при використанні моделі (2) для незалежної змінної x_i :

$$\hat{y}_i = b_0 + b_1 x_i. \quad (5)$$

Побудована за значеннями ряду Y' модель характеризується парою коефіцієнтів b_0' і b_1' які є точковими оцінками параметрів β_0' і β_1' моделі (3.1) відповідно. Якщо прийняти припущення, що об'єкт не зазнав змін, то це означає, що модель повинна залишитися адекватною і для значень всього ряду Y . Формально це призводить до формулювання наступної системи гіпотез:

$$\begin{cases} H_0 : \beta_0' = \beta_0 \text{ and } \beta_1' = \beta_1; \\ H_A : \beta_0' \neq \beta_0 \text{ or } \beta_1' \neq \beta_1. \end{cases} \quad (6)$$

Отримавши значення коефіцієнтів b_0', b_1' , використовуючи статистики t -розподілу Стьюдента (4) та підставивши в них значення оцінок b_0, b_1 , які визначені за повним набором даних Y , можна отримати значення коефіцієнтів $p_value_{\beta_0}$ і $p_value_{\beta_1}$, що характеризують досягнутий рівень значущості прийняття гіпотези (6) статистичної тотожності коефіцієнтів. На підставі цих значень робиться висновок про прийняття рішення на користь гіпотези H_0 або H_A .

Алгоритм детектування точки зміни на основі аналізу довірчого інтервалу прогнозу (АДИП-алгоритм). При побудові моделі регресії може бути визначено самі значення \hat{Y}'' , тобто передбачені значення даних для поточного вікна, а також їх довірчі інтервали. Припустимо, що за значеннями ряду Y' побудовано регресійну модель (3.1). При прогнозуванні значення залежної змінної у точці $x_j, j \geq k$ підставимо в регресійну модель прогнозу (2) вираз для параметра $b_0' : b_0' = \bar{y}' - b_1' \bar{x}'$. Рівняння регресії набуде вигляду:

$$\hat{y}_j = \bar{y}' + b_1' (x_j - \bar{x}').$$

Використовуючи відомі властивості дисперсії та з урахуванням того, що доступні лише її вибіркові оцінки, отримуємо:

$$S_{\hat{y}_j}^2 = S_{\bar{y}'}^2 + (x_j - \bar{x}')^2 S_{b_1'}^2.$$

Підставляючи значення $S_{b_1'}^2$ і беручи до уваги, що

$$S_{\bar{y}'}^2 = \frac{S'^2}{k-1}$$

отримуємо:

$$S_{\hat{y}_j}^2 = \frac{S'^2}{k-1} + (x_j - \bar{x}')^2 \frac{S'^2}{\sum_{i=1}^{k-1} (x_i - \bar{x}')^2} = S'^2 \left(\frac{1}{k-1} + \frac{(x_j - \bar{x}')^2}{\sum_{i=1}^{k-1} (x_i - \bar{x}')^2} \right),$$

де $k \leq j \leq N$.

Однак, сама випадкова величина \hat{y}_j відома з точністю до ε_j , оцінка дисперсії якої є S^2 . З огляду на цей факт отримуємо для оцінки дисперсії значення y_j у точці x_j :

$$S_{y_j}^2 = S_{\hat{y}_j}^2 + S'^2 = S'^2 \left(\frac{1}{k-1} + \frac{(x_j - \bar{x}')^2}{\sum_{i=1}^{k-1} (x_i - \bar{x}')^2} + 1 \right).$$

Це дозволяє явно виразити межі довірчого інтервалу для прогнозованого значення y_j у точці x_j :

$$\hat{y}_j - t_{\alpha, df} S_{y_j} < y_j < \hat{y}_j + t_{\alpha, df} S_{y_j} \quad (7)$$

де $t_{\alpha, df}$ - α -квантиль стандартного t-розподілу Стьюдента з df ступенями свободи. У разі лінійної регресійної моделі $df = m - 2$.

Таким чином загальна процедура виявлення точки модифікується наступним чином:

– будується регресійна модель за значеннями еталонного вікна Y' ;

– визначається кількість точок еталонного вікна P' значення яких виходять за межі побудованого за формулою (7) довірчого інтервалу, а також їх частка серед усіх значень еталонного вікна $p' = \frac{P'}{k-1}$;

– після отримання значень поточного вікна Y'' підраховується кількість P'' та частка $p'' = \frac{P''}{n-k+1}$ випадків виходу значень точок поточного вікна за межі побудованого для них довірчого інтервалу.

Формально, система гіпотез наявності/відсутності зміни в об'єкті в такому випадку зводиться до перевірки гіпотез про рівність часток у вибірках:

$$\begin{cases} H_0 : p' = p''; \\ H_A : p' \neq p''. \end{cases}$$

критерієм для перевірки якої є статистика

$$z = \frac{p' - p''}{\sqrt{\frac{p'(1-p')}{(k-1)} + \frac{p''(1-p'')}{(n-k+1)}}}.$$

що розподілена за стандартним нормальним законом розподілу. Отримане таким чином значення статистики z дає можливість обчислити для неї значення p_value_p , на підставі якого можна стандартною процедурою прийняти або відхилити зазначену гіпотезу H_0 [22].

Алгоритм детектування точки зміни на основі використання нерівності Гефдінга (АНГ-алгоритм). Нерівність Гефдінга стверджує, що якщо $\xi_1, \xi_2, \dots, \xi_n$ – послідовність незалежних випадково розподілених величин, обмежених зверху та знизу $\xi \in [a, b]$, $\bar{\xi}$ – середнє значення зазначеної послідовності, а $E[\xi]$ – математичне очікування генеральної сукупності з якої отримують значення вибірки, то справедливо наступне співвідношення [23]:

$$P\left(\left|\bar{\xi} - E(\xi)\right| \geq \zeta\right) \leq 2 \cdot \exp\left(\frac{-2n\zeta^2}{(b-a)^2}\right).$$

У рамках задачі, що розглядається побудуємо регресійну модель (3.1) та обчислюємо ряд значень її помилок:

$$\varepsilon_i' = y_i - \hat{y}_i, 1 \leq i < k. \quad (8)$$

Якщо припустити, що побудована модель адекватно визначає поведінку часового ряду на відрізку еталонного вікна, можна вважати, що середнє значення ряду $\bar{\varepsilon}'$ є конзистентною оцінкою математичного очікування значень випадкової величини $\{\varepsilon_i\}$. Якщо поведінка об'єкта не зазнала істотних змін, це значення має залишитися таким самим і при роботі зі значеннями поточного вікна. Позначимо

$$\varepsilon_i'' = y_i - \hat{y}_i, k \leq i < n. \quad (9)$$

У термінах помилки між значеннями поточного вікна та їх передбачень на основі використання моделі, побудованої за значеннями еталонного вікна, нерівність Гефдінга може бути переписана у вигляді:

$$P\left(\left|\bar{\varepsilon}'' - \bar{\varepsilon}'\right| \geq \zeta\right) \leq 2 \cdot \exp\left(\frac{-2n\zeta^2}{(\max\{\varepsilon_i\} - \min\{\varepsilon_i\})^2}\right).$$

Цей вираз задає ймовірність того, що різниця $|\bar{\varepsilon}'' - \bar{\varepsilon}'|$ перевищить значення відмінності, яка реально спостерігаються за умови незмінності моделі. Якщо ця ймовірність менша від прийнятого значення значущості α , гіпотеза не H_0 відхиляється.

Алгоритми виявлення структурних змін Чоу (АСІЧ1 та АСІЧ2). Тести Чоу [24] виходять із припущення, що у разі відсутності точки зміни модель часового ряду, побудована за значеннями Y не буде суттєво покращувати точність у порівнянні з моделями, побудованими за рядами значень Y' і Y'' . У якості параметру для оцінки точності використовуються сума квадратів залишків відповідних моделей:

$$SSR = \sum_{i=1}^n (\hat{y}_i - y_i)^2.$$

$$SSR' = \sum_{i=1}^{k-1} (\hat{y}_i' - y_i)^2.$$

$$SSR'' = \sum_{i=k}^n (\hat{y}_i'' - y_i)^2.$$

Статистика першого тесту Чоу порівнює відношення помилок моделі (2) ряду відносно суми помилок моделей (3.1) та (3.2):

$$Ch1 = \frac{(SSR - (SSR' + SSR'')) \cdot (n-4)}{(SSR' + SSR'') \cdot 2}.$$

Для визначення статистики другого тесту Чоу будується модель ряду еталонного вікна Y' та на основі цієї моделі прогнозується значення ряду для поточного вікна. Після надходження значень поточного вікна будується модель за всіма наявними даними та розраховується для неї показник $SSR^* = \sum_{i=1}^n (\hat{y}_i - y_i)^2$. Чим більша різниця між помилками, отриманими при використанні цих моделей на значеннях поточного вікна, тим більша різниця між даними у вікнах.

$$Ch2 = \frac{(SSR^* - SSR') \cdot (k-3)}{SSR' \cdot (n-k+1)}.$$

Обидві зазначені статистики підпорядковуються закону розподілу Фішера з кількістю ступенів свободи 2 та $(n-4)$ для $Ch1$ та $(k-3)$ і $(n-k+1)$ для $Ch2$. Це дає можливість отримати для них значення p_value_{Ch1} та/або p_value_{Ch2} , на підставі яких можна стандартною процедурою прийняти чи відхилити гіпотезу H_0 .

Алгоритм детектування точки зміни з використанням критерію Колмогорова-Смирнова. (АКС-алгоритм). Після того, як модель побудована на даних еталонного вікна, спрогнозовано значення для поточного вікна та отримано реальні дані ряду Y , можуть бути обчислені значення ряду помилок передбачення (8) та (9). У разі, якщо змін в об'єкті не відбулися, закони розподілу, яким підпорядковуються відповідні випадкові величини мають бути тотожні. Для перевірки цього пропонується використовувати критерій однорідності Смирнова (інша назва – двовибірковий критерій однорідності Колмогорова-Смирнова) [22].

Експериментальне дослідження. Дослідження, що проводилося з використанням процедури статистичного моделювання, мало за мету отримати порівняльну характеристику ефективності різних алгоритмів, що ґрунтуються на використанні лінійних регресійних моделей опису ряду при виявленні точок зміни. Дослідження розбивалося на частини – на статистичне моделювання роботи алгоритмів та оцінку ефективності їх з допомогою різних критеріїв визначення ефективності, і на з'ясування залежності ефективності алгоритмів від положення точки зміни відносно початкової точки поточного вікна.

Для досягнення мети дослідження генерувалися дані, що описують різновиди змін моделей часового ряду, а саме:

1. Стаціонарний, нормально розподілений часовий ряд.
2. Часовий ряд, що моделює постійний тренд даних що зростає.
3. Часовий ряд, що моделює постійний тренд даних що спадає.
4. Часовий ряд, що моделює різку, одномоментну зміну середнього значення.
5. Часовий ряд, що моделює стаціонарну поведінку ряду до точки зміни та появу тренду що зростає після цієї точки.
6. Часовий ряд, що моделює наявність деякого тренду що зростає до точки зміни та збільшення швидкості зростання тренду після цієї точки.
7. Часовий ряд, що моделює наявність спадаючого тренду до точки зміни та уповільнення швидкості падіння тренда після цієї точки.
8. Часовий ряд, що моделює розворот тренду після точки зміни.

У перших трьох рядах точки змін даних свідомо були відсутні, отже ці приклади використовувалися для аналізу генерованих алгоритмами помилок I роду. В інших прикладах точка зміни була присутня, й вони використовувалися для аналізу виявлення помилок II роду.

Оцінки ефективності алгоритмів, що використовувалися в дослідженні. Після проведення експериментів для кожного з алгоритмів підраховувалися чотири показники:

- TP-кількість згенерованих сигналів про наявність точки зміни за умови, що така зміна насправді мала місце у наданих для аналізу даних;
- TF-кількість згенерованих сигналів про відсутність точки зміни за умови, що така зміна у наданих даних в дійсності була відсутня;
- FP-кількість згенерованих сигналів про наявність точки зміни за умови, що в наданих для аналізу даних такі зміни були відсутні (помилка I роду);
- FN-кількість згенерованих сигналів про відсутність точки зміни за умови, що зміна мала місце у наданих для аналізу даних (помилка II роду).

Ідеальним був би алгоритм, який міг би одночасно мінімізувати помилки як I так і II роду про те, це теоретично неможливо [22]. Тож для аналізу використовувалися узагальнювальні метрики, які є похідними від перелічених вище. Такими метриками є:

- точність (accuracy) – відношення всіх правильно розпізнаних точок змін до загальної кількості проведених тестів:

$$acc = \frac{TP + TN}{TP + TN + FP + FN};$$

– влучність (precision) – показує, яка частка випадків, що були розпізнані як такі, що містять точки змін, насправді такими й були:

$$pr = \frac{TP}{TP + FP};$$

– повнота (recall) – показує, яку частку серед усіх точок змін, що насправді містяться в даних було виявлено:

$$rec = \frac{TP}{TP + FN}.$$

Аналіз результатів експериментів. Перша серія експерименту проводилася з метою з'ясування ефективності роботи алгоритмів у найбільш сприятливій ситуації, коли зміна відбувалася в першій точці поточного вікна. Для наведених зразків рядів даних проводилося по 10000 експериментів, у кожному з яких фіксувалося рішення, прийняте кожним з описаних алгоритмів. Значення довжини еталонного і поточного вікна приймалося рівним 80, значення $\alpha = 0.05$. Результати, що дозволяють оцінити ефективність алгоритмів зведені у табл. 1 та табл. 2.

У табл. 1 наведені результати показують, що більшість алгоритмів зберігають достатній рівень точності рішення з точки зору мінімізації помилок I і II роду. Виключення складає алгоритм АКС, який є одним з найкращих по відношенню до помилок II роду, проте показує занадто високий рівень помилок I роду. Алгоритм АНГ показує високу залежність від типу зміни з огляду на кількість генерованих помилок II роду. З іншого боку, алгоритм АСІЧ1 показує найвищий рівень ефективності серед алгоритмів, що досліджувалися. У табл. 2 наведені узагальнювальні показники ефективності за всіма обраними зразками змін ряду та обраними метриками якості.

Наступна серія експериментів виявляла залежність ефективності алгоритмів від зміщення точки зміни відносно першої точки поточного вікна. Природно очікувати, що якість будь-якого алгоритму буде погіршуватися зі збільшенням зазначеної затримки. Тому кожен експеримент передбачав використання трьох послідовних вікон, причому вважалося, що експеримент вдалий, якщо факт зміни, який мав місце в середньому з цих вікон, був виявлений або при аналізі даних першого та другого, або при аналізі даних другого та третього вікон. Хоча у другому випадку час затримки виявлення точки зміни зростає, водночас з'являється можливість уникнення помилки II роду. У багатьох прикладних галузях саме цей фактор (зниження ймовірності пропуску точки зміни) є критичним. Результати цієї серії експериментів у вигляді значень метрики *rec*, наведені у табл. 3.

Таблиця 1 – Частка помилкових рішень (найкращий випадок)

Номер зразка ряду	Алгоритми					
	АКРМ	АДП	АНГ	АСІЧ1	АСІЧ2	АКС
	<i>Частка помилок I типу.</i>					
1	0.0782	0.0829	0.0496	0.0495	0.1150	0.3501
2	0.0767	0.0803	0.0495	0.0495	0.0109	0.3414
3	0.0832	0.0853	0.0536	0.0526	0.1270	0.3482
	<i>Частка помилок типу II.</i>					
4	0.1421	0.1350	0.1004	0.0204	0.2825	0.0101
5	0.0209	0.0753	0.1968	0.0043	0.1506	0.0151
6	0.0755	0.1273	0.3479	0.0012	0.2424	0.1026
7	0.0930	0.1535	0.3289	0.0032	0.2624	0.0981
8	0.0097	0.0310	0.1809	0.0005	0.0672	0.0285

Таблиця 2 – Порівняння алгоритмів (найкращий випадок)

Метрики якості	Алгоритми					
	АКРМ	АДП	АНГ	АСІЧ1	АСІЧ2	АКС
pr	0.95138	0.94742	0.96180	0.97041	0.99129	0.82029
acc	0.92759	0.90368	0.83655	0.97741	0.86998	0.83824
rec	0.93176	0.89558	0.76902	0.99418	0.79898	0.94912

Таблиця 3 – Порівняння алгоритмів за метрикою Res

Зсув відносно початку вікна	Алгоритми					
	АКРМ	АДИП	АНХ	АСИЧ1	АСИЧ2	АКС
1 точка	0.91902	0.88330	0.73518	0.97976	0.76036	0.95212
25%	0.88570	0.79716	0.60946	0.96885	0.57706	0.94282
50%	0.91856	0.83282	0.72772	0.98416	0.65432	0.96540
75%	0.88602	0.83970	0.76748	0.97036	0.71658	0.94142
n-1 точка	0.91962	0.88728	0.79094	0.98968	0.78014	0.97400
Усереднене значення	0.90578	0.84805	0.72616	0.98050	0.69769	0.95515

Висновки. Досліджені у роботі алгоритми показали досить високий рівень ефективності та здатні зайняти свою нішу серед інших алгоритмів моніторингу. Отримані результати дозволяють оцінити їх ефективність в залежності від типів змін, які виникають у часовому ряді, а також в усередненому вигляді, по всій множині змін. Разом з тим, відкритим залишається питання підвищення швидкості виявлення точки змін, що особливо актуально для тих прикладних галузей, для яких характерно високошвидкісні потоки інформації, що надходить. Предметом подальшого дослідження може стати спільне використання регресійних моделей для опису часового ряду параметрів та методів накопичувального виявлення відхилень, що використовують менші значення зсуву поточного вікна відносно еталонного. Іншим напрямом розвитку отриманих у роботі результатів передбачається вивчення можливості одночасного застосування низки методів створення ансамблю алгоритмів [8]. Отримані в роботі результати можуть бути підставою для відбору алгоритмів, що включаються до ансамблю, по-перше, на основі їх ефективності для різних типів змін ряду, а по-друге, таких, для яких результати роботи виявляться найменш корельованими між собою, що дозволить знизити обчислювальну складність реалізації ансамблю алгоритмів. Очікується, що за рахунок його використання вдасться підвищити точність прийнятих рішень, а використання у всіх алгоритмах загальної для них базової регресійної моделі – не вимагатиме істотного зростання обчислювальної складності програмної реалізації.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] D. G. Pascual, *Artificial Intelligence Tools: Decision Support Systems in Condition Monitoring and Diagnosis*, Boca Raton, Florida, USA: CRC Press, 2015.
- [2] E. S. Page, “A test for a change in a parameter occurring at an unknown point”, *Biometrika*, vol. 42, no. 3/4, pp. 523-527, 1955, doi: <https://doi.org/10.2307/2333401>.
- [3] S. Liu, A. Wright, and M. Hauskrecht, “Change-Point Detection Method for Clinical Decision Support System Rule Monitoring”, *Artif Intell Med Conf Artif Intell Med*, pp. 126-135, 2017, doi: https://doi.org/10.1007/978-3-319-59758-4_14.
- [4] S. J. Taylor, and B. Letham, “Forecasting at Scale”, *The American Statistician*, Taylor & Francis J., vol. 72 (1), pp. 37-45, 2018, doi: <https://doi.org/10.1080/00031305.2017.1380080>.
- [5] D.-H. Tran, “Change detection in streaming data”, *Thesis of dissertation of the Requirements for the degree of Doktor-Ingenieur*, Technische Universität Ilmenau, Germany, 2013.

- [6] A. G. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential Analysis: Hypothesis Testing and Change-point Detection*, Monographs on Statistics & Applied Probability, vol. 136, Publisher: Chapman & Hall/CRC, Taylor and Francis Group, 2015.
- [7] B. Kurt, Ç. Yıldız, T. Y. Ceritli, B. Sankur, and A. T. Cemgil, “A Bayesian change point model for detecting SIP-based DDoS attacks”, *Digital Signal Processing*, vol. 77, pp. 48-62, 2018, doi: <https://doi.org/10.1016/j.dsp.2017.10.009>.
- [8] В. В. Соколов, О. М. Шаповал, та Д. М. Шарадкін “Ансамбль алгоритмів виявлення аномалій в часових рядах та його використання до задач моніторингу стану систем в реальному часі”, *Збірник наукових праць ВІТІ*, 2020, № 3, с. 82-93.
- [9] Д. М. Шарадкін, “Використання критерію виявлення змін поведінки об’єкта на основі аналізу коефіцієнта автокореляції в задачах забезпечення інформаційної безпеки”, *Information Technology and Security*, vol. 5, iss.1 (8), pp. 42-54, January-June 2017, doi: <https://doi.org/10.20535/2411-1031.2017.5.1.120555>.
- [10] D. V. Hinkley, “Inference about the change-point in a sequence of random variables” *Biometrika*, vol. 57, no. 1, pp. 1-17, 1970, doi: <https://doi.org/10.2307/2334932>.
- [11] S. Liu, M. Yamada, N. Collier, and M. Sugiyama, “Change-point detection in time-series data by relative density-ratio estimation”, *Neural Networks*, vol. 43, pp. 72-83, 2013, doi: <https://doi.org/10.1016/j.neunet.2013.01.012>.
- [12] G. Boracchi, C. Cervellera, and D. Macciò, “Uniform histograms for change detection in multivariate data”, *International Joint Conference on Neural Networks (IJCNN)*, Anchorage, pp. 1732-1739, 2017, doi: <https://doi.org/10.1109/IJCNN.2017.7966060>.
- [13] R. Sebastiao, and J. Gama, “A study on change detection methods”, in *Proc. 4th Portuguese Conf. on Artificial Intelligence*, Aveiro, Portugal, 2009, pp. 353-364.
- [14] M. Hushchyn, K. Arzymatov, and D. Derkach, “Online Neural Networks for Change-Point Detection”, arXiv: 2010.01388, 2020. [Online]. Available: <https://arxiv.org/abs/2010.01388>. Accessed on: Aug. 27, 2021.
- [15] T. de Ryck, M. de Vos, and A. Bertrand, “Change Point Detection in Time Series Data Using Autoencoders With a Time-Invariant Representation”, in *IEEE Transactions on Signal Processing*, vol. 69, pp. 3513-3524, 2021, doi: <https://doi.org/10.1109/TSP.2021.3087031>.
- [16] K. G. Mehrotra, C. K. Mohan, and H. M. Huang, *Anomaly Detection. Principles and Algorithms*. Springer International Publishing AG 2017, doi: <https://doi.org/10.1007/978-3-319-675268>.
- [17] S. Ratnasingam, “Sequential change-point detection in linear regression and linear quantile regression models under high dimensionality”, *A Dissertation Submitted to the Graduate College of Bowling Green State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy*, August 2020.
- [18] М. В. Житлухин, “Последовательные методы проверки статистических гипотез и обнаружения разладки”, *Диссертация на соискание ученой степени кандидата физико-математических наук*, Математический институт им. В. А. Стеклова, Москва, 2013.
- [19] X. Xu, H. Liu, and M. Yao, Recent Progress of Anomaly Detection. *Hindawi Complexity*, vol. 2019, Article ID 2686378, 11 p., doi: <https://doi.org/10.1155/2019/2686378>.
- [20] G. J. J. van den Burg, and C. K. I. Williams, “An Evaluation of Change Point Detection Algorithms”, arXiv: 2003.06222v2 [stat.ML] 25 May 2020, 33 p. [Online] Available: <https://arxiv.org/abs/2003.06222v2>. Accessed on: Aug. 27, 2021.
- [21] E.-S. Apostol, C.-O. Truica, F. Pop, and C. Esposito, “Change Point Enhanced Anomaly Detection for IoT Time Series Data”, *Water (Special Issue Smart Water Solutions with Big Data)*, vol. 13, no. 12, 20 p., 2021, doi: <https://doi.org/10.3390/w13121633>.
- [22] С. А. Айвазян, В. М. Бухштабер, И. С. Енюков, и Л. Д. Мешалкин, *Прикладная статистика. Классификации и снижение размерности*, Москва : Финансы и статистика, 1989.
- [23] Ш. Шалев-Шварц, и Ш. Бен-Давид, *Идеи машинного обучения. От теории к алгоритмам*, Москва: ДМК Пресс, 2019.

- [24] Л. А. Ландман, и А. В. Фаддеенков, “Сравнительный анализ различных методов оценки структурной стабильности объектов на основе их временных рядов”, *Вестник АГТУ. Серия: Управление, вычислительная техника и информатика*, № 1. С. 119-128, 2019, doi: <https://doi.org/10.24143/2072-9502-2019-1-119-128>.

Стаття надійшла до редакції 10.09.2021.

REFERENCES

- [1] D. G. Pascual, *Artificial Intelligence Tools: Decision Support Systems in Condition Monitoring and Diagnosis*, Boca Raton, Florida, USA: CRC Press, 2015.
- [2] E. S. Page, “A test for a change in a parameter occurring at an unknown point”, *Biometrika*, vol. 42, no. 3/4, pp. 523-527, 1955, doi: <https://doi.org/10.2307/2333401>.
- [3] S. Liu, A. Wright, and M. Hauskrecht, “Change-Point Detection Method for Clinical Decision Support System Rule Monitoring”, *Artif Intell Med Conf Artif Intell Med*, pp. 126-135, 2017, doi: https://doi.org/10.1007/978-3-319-59758-4_14.
- [4] S. J. Taylor, and B. Letham, “Forecasting at Scale”, *The American Statistician, Taylor & Francis J.*, vol. 72 (1), pp. 37-45, 2018, doi: <https://doi.org/10.1080/00031305.2017.1380080>.
- [5] D.-H. Tran, “Change detection in streaming data”, *Thesis of dissertation of the Requirements for the degree of Doktor-Ingenieur*, Technische Universität Ilmenau, Germany, 2013.
- [6] A. G. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential Analysis: Hypothesis Testing and Changepoint Detection*, Monographs on Statistics & Applied Probability, vol. 136, Publisher: Chapman & Hall/CRC, Taylor and Francis Group, 2015.
- [7] B. Kurt, Ç. Yıldız, T. Y. Ceritli, B. Sankur, and A. T. Cemgil, “A Bayesian change point model for detecting SIP-based DDoS attacks”, *Digital Signal Processing*, vol. 77, pp. 48-62, 2018, doi: <https://doi.org/10.1016/j.dsp.2017.10.009>.
- [8] V. V. Sokolov, O. M. Shapoval, and D. M. Sharadkin “Ensemble of algorithms for detecting anomalies in time series and its use for problems of monitoring the state of systems in real time”, *VITI's collection of scientific papers*, 2020, no. 3, p. 82-93, 2020.
- [9] D. M. Sharadkin, “Change-point detection test based on the analysis of the time series' autocorrelation and its application for information security”, *Information Technology and Security*, vol. 5, iss. 1 (8), pp. 42-54, January-June 2017, doi: <https://doi.org/10.20535/2411-1031.2017.5.1.120555>.
- [10] D. V. Hinkley, “Inference about the change-point in a sequence of random variables” *Biometrika*, vol. 57, no. 1, pp. 1-17, 1970, doi: <https://doi.org/10.2307/2334932>.
- [11] S. Liu, M. Yamada, N. Collier, and M. Sugiyama, “Change-point detection in time-series data by relative density-ratio estimation”, *Neural Networks*, vol. 43, pp. 72-83, 2013, doi: <https://doi.org/10.1016/j.neunet.2013.01.012>.
- [12] G. Boracchi, C. Cervellera, and D. Macciò, “Uniform histograms for change detection in multivariate data”, *International Joint Conference on Neural Networks (IJCNN)*, Anchorage, pp. 1732-1739, 2017, doi: <https://doi.org/10.1109/IJCNN.2017.7966060>.
- [13] R. Sebastiao, and J. Gama, “A study on change detection methods”, in *Proc. 4th Portuguese Conf. on Artificial Intelligence*, Aveiro, Portugal, 2009, pp. 353-364.
- [14] M. Hushchyn, K. Arzymatov, and D. Derkach, “Online Neural Networks for Change-Point Detection”, arXiv: 2010.01388, 2020. [Online]. Available: <https://arxiv.org/abs/2010.01388>. Accessed on: August 27, 2021.
- [15] T. de Ryck, M. de Vos, and A. Bertrand, “Change Point Detection in Time Series Data Using Autoencoders With a Time-Invariant Representation”, in *IEEE Transactions on Signal Processing*, vol. 69, pp. 3513-3524, 2021, doi: <https://doi.org/10.1109/TSP.2021.3087031>.
- [16] K. G. Mehrotra, C. K. Mohan, and H. M. Huang, *Anomaly Detection. Principles and Algorithms*. Springer International Publishing AG 2017, doi: <https://doi.org/10.1007/978-3-319-675268>.

- [17] S. Ratnasingam, “Sequential change-point detection in linear regression and linear quantile regression models under high dimensionality”, *A Dissertation Submitted to the Graduate College of Bowling Green State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy*, August 2020.
- [18] M. V. Zhitlukhin, “Sequential Methods for Testing Statistical Hypotheses and Detecting Disorders”, *Dissertation for the degree of candidate of physical and mathematical sciences*, V. A. Steklov’s Mathematical Institute., Moscow, 2013.
- [19] X. Xu, H. Liu, and M. Yao, Recent Progress of Anomaly Detection. *Hindawi Complexity*, vol. 2019, Article ID 2686378, 11 p., doi: <https://doi.org/10.1155/2019/2686378>.
- [20] G. J. J. van den Burg, and C. K. I. Williams, “An Evaluation of Change Point Detection Algorithms”, arXiv: 2003.06222v2 [stat.ML] 25 May 2020, 33 p. [Online] Available: <https://arxiv.org/abs/2003.06222v2>. Accessed on: August 27, 2021.
- [21] E.-S. Apostol, C.-O. Truica, F. Pop, and C. Esposito, “Change Point Enhanced Anomaly Detection for IoT Time Series Data”, *Water (Special Issue Smart Water Solutions with Big Data)*, vol. 13, no. 12, 20 p., 2021, doi: <https://doi.org/10.3390/w13121633>.
- [22] S. A. Ayvazyan, V. M. Buchstaber, I. S. Enyukov, and L. D. Meshalkin, *Applied Statistics. Classifications and dimensionality reduction*, Moskow: Finance and statistics, 1989.
- [23] Sh. Shalev-Schwartz, and Sh. Ben-David, *Machine Learning Ideas. From theory to algorithms*, M: DMK Press, 2019.
- [24] L. A. Landman and A. V. Faddeenkov, “Comparative analysis of various methods for assessing the structural stability of objects based on their time series”, *Bulletin of the ASTU. Series: Management, Computer Engineering and Informatics*, no. 1, pp. 119-128, 2019, doi: <https://doi.org/10.24143/2072-9502-2019-1-119-128>.

VIACHESLAV RIABTSEV,
DMYTRO SHARADKIN,
YURII KLIAT

COMPARATIVE ANALYSIS OF ALGORITHMS FOR CHANGE POINTS DETECTION IN REGRESSION MODELS OF TIME SERIES

Changes detection in the behavior of the object is one of the main goals of any real-world objects’ monitoring. The behavior of an object can change over time due to its reaction to external events, as well as under the influence of the internal development laws of the object. Problems occur in cases when internal changes couldn't be observed directly. In such cases, it is possible to obtain information about the internal changes only through analysis of the time series observed parameters, the measurement of which could be performed technically and/or organizationally. Several changes detection algorithms in time series, which are based on the linear regression models are discussed. It is assumed that if the event of changes occurs, the time series before and the time series after the change point are described by models which cannot be considered identical. Some known algorithms for analyzing the identity of models and algorithms that have not previously been deposited in the literature are considered. In particular, an algorithm for change points detection based on changes in the values of the regression models coefficients; based on analyzing the confidence interval of the predicted values of the series; based on analyzing regression residuals based on the use of Höfdding's inequality; Chow algorithms for comparing variances of residuals; an algorithm for comparing the distributions of residuals using the Kolmogorov-Smirnov's two-samples test. The variety of objects in the real world, types of changes in their behavior, which are based on the unpredictability of the reasons that caused them, do not allow choosing a single change points detection method and makes a comparative assessment of various algorithms an urgent engineering problem. The paper proposes an algorithms’ analysis depending on the types of changes that are most typical for parameters time series in various domains. The effectiveness of the algorithms is assessed by the method of the

statistical experiment by their ability to detect a change, as well as by comparing the number of false detection errors and skipping real changes. The obtained results can be used for further research of algorithms, in particular, in the case of constructing an ensemble of algorithms for identifying change points of the behavior of monitored objects.

Key words: time series, anomalies, change points detection, change detection predictive algorithms, regression models.

Рябцев Вячеслав Віталійович, кандидат технічних наук, доцент, доцент кафедри кібербезпеки і застосування інформаційних систем і технологій, Інститут спеціального зв'язку та захисту інформації Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського”, Київ, Україна, ORCID 0000-0001-8331-0132, viacheslav.riabtsev@gmail.com.

Шарадкін Дмитро Михайлович, кандидат технічних наук, доцент, доцент кафедри кібербезпеки і застосування інформаційних систем і технологій, Інститут спеціального зв'язку та захисту інформації Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського”, Київ, Україна, ORCID 0000-0001-6407-8040, dm.sharadkin@gmail.com.

Клят Юрій Олександрович, кандидат технічних наук, заступник начальника Військової академії, Одеса, Україна, ORCID 0000-0002-8267-3748, Klyatt@ukr.net.

Riabtsev Viacheslav, candidate of engineering sciences, associate professor, associate professor at the cybersecurity and application of information systems and technology academic department, Institute of special communication and information security of National technical university of Ukraine “Igor Sikorsky Kyiv polytechnic institute”, Kyiv, Ukraine.

Sharadkin Dmytro, candidate of engineering sciences, associate professor, associate professor at the cybersecurity and application of information systems and technology academic department, Institute of special communication and information security of National technical university of Ukraine “Igor Sikorsky Kyiv polytechnic institute”, Kyiv, Ukraine.

Kliat Yurii, candidate of engineering sciences, deputy chief of the Military academy, Odesa, Ukraine.