УДК 004.4’242.2:004.032.26 DOI 10.36910/775.24153966.2024.79.12

В. В. Яковець

РОЗРОБКА МЕТОДІВ ВИЯВЛЕННЯ АНОМАЛІЙ У ВЕБ-ДОДАТКАХ ІЗ ВИКОРИСТАННЯМ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ

Ужгородський національний університет

У статті досліджено проблему автоматизованого виявлення аномалій у веб-додатках в умовах зростання складності систем і обсягів даних. Розроблено та експериментально перевірено три алгоритми, які забезпечують підвищення точності та швидкості виявлення аномалій. Встановлено, що модуль часової кореляції демонструє найвищу точність, адаптивний пороговий моніторинг забезпечує швидкість реагування, а гібридна кластеризація характеризується збалансованими показниками. Виявлено ключові проблеми впровадження, зокрема обчислювальну складність і потребу в адаптації моделей. Запропоновано рекомендації для інтеграції алгоритмів у системи моніторингу та окреслено перспективи розробки оптимізованих методів і вивчення нових типів загроз.

Ключові слова: *виявлення аномалій, веб-додатки, інтелектуальний аналіз даних, адаптивний моніторинг, часові ряди, кластеризація, алгоритми, кібербезпека.*

V. Yakovets

DEVELOPMENT OF METHODS FOR ANOMALY DETECTION IN WEB APPLICATIONS USING DATA MINING TECHNIQUES

The article addresses the issue of automated anomaly detection in web applications amid increasing system complexity and data volumes. Three algorithms were developed and experimentally validated to enhance the accuracy and speed of anomaly detection. It was established that the time correlation module demonstrates the highest accuracy, the adaptive threshold monitoring ensures rapid response, and hybrid clustering offers balanced performance. Key implementation challenges, including computational complexity and the need for model adaptation, were identified. Recommendations for integrating the algorithms into monitoring systems were proposed, and prospects for developing optimized methods and exploring new types of threats were outlined.

Keywords: *anomaly detection, web applications, data mining, adaptive monitoring, time series, clustering, algorithms, cybersecurity.*

Постановка проблеми. Виявлення аномалій у веб-додатках є актуальною проблемою в умовах стрімкого розвитку цифрових технологій і зростання обсягів даних. Аномалії можуть свідчити про порушення в роботі програмного забезпечення, потенційні кібератаки або інші небажані події, що впливають на функціональність і безпеку веб-додатків. Традиційні методи моніторингу й аналізу даних часто є недостатньо ефективними через їхню неспроможність обробляти великі обсяги інформації в реальному часі або враховувати складність сучасних багатокомпонентних систем.

Використання методів інтелектуального аналізу даних для виявлення аномалій відкриває нові можливості у створенні високоточної й адаптивної системи моніторингу, яка здатна виявляти неочевидні закономірності й аномальні дії. Це завдання має важливе наукове значення, оскільки передбачає розроблення нових алгоритмів і моделей для автоматизованого аналізу складних даних. Практична цінність проблеми полягає у підвищенні рівня безпеки та надійності веб-додатків, що є особливо актуальним для секторів, де обробляється конфіденційна інформація, таких як банківська справа, електронна комерція або охорона здоров’я.

Аналіз останніх досліджень і публікацій. Дослідження сучасних методів виявлення аномалій у веб-додатках свідчать про зростаючу важливість інтелектуального аналізу даних для підвищення точності, швидкості та адаптивності алгоритмів. У контексті цієї тематики дослідження К. Чижмара, О. Дніпрова, О. Коротюка, Р. Шаповала та О. Сидоренка акцентують увагу на кібербезпеці державних інформаційних систем і підкреслюють необхідність інтегрованих рішень для моніторингу веб-додатків [1].

Результати Х. Сю та його команди демонструють перспективність варіаційних автоенкодерів для виявлення аномалій у сезонних веб-показниках, що забезпечує високу точність у динамічних умовах [2]. С. Юань і його колеги довели ефективність інтеграції прогнозування подій із виявленням аномалій, використовуючи самонавчальні нейронні мережі. Такий підхід дозволяє передбачати потенційні ризики у веб-додатках [3].

Метод слабко супервізованого виявлення аномалій, запропонований Х. Чжао та його співавторами, демонструє високі результати в умовах обмеженого доступу до даних для навчання моделей, що є важливим для ресурсозалежних систем [4]. У свою чергу, Д.-К. Хуан та його команда

розробили метамодель для аналізу пошукових запитів, яка показує високу ефективність у виявленні відхилень у поведінці користувачів [5].

Дослідження К. Агарвала фокусується на ансамблевих методах для підвищення точності алгоритмів, підкреслюючи переваги комбінованих підходів у складних середовищах [6]. Огляд Р. Чалапаті та С. Чавли висвітлює важливість глибокого навчання для аналізу багатовимірних даних, яке дозволяє моделювати складні патерни аномалій [7].

Розробка К. Чоя та його команди орієнтована на аналіз часових рядів. Їхні моделі оптимізують виявлення аномалій у динамічних середовищах, що корисно для моніторингу веб-додатків у реальному часі [8]. С. Ван та його співавтори акцентують на автоматизації аналізу логів, що дозволяє виявляти зловмисну активність у мережевих додатках [9].

Методологія Ю. Гао зосереджується на поведінковому аналізі користувачів через веб-журнали, підтверджуючи ефективність таких даних у запобіганні загрозам [10]. К. Сі пропонує інноваційний підхід до відновлення тензорів для аналізу інтернет-трафіку, що дозволяє ідентифікувати складні залежності у веб-даних [11].

Розроблена М.С. Рахманом гібридна система для моніторингу соціальних мереж показала, що поєднання поведінкових та контентних характеристик підвищує точність виявлення аномалій [12]. Р. Ван та інші досліджують глибоке навчання для створення адаптивних систем моніторингу, що ефективно працюють у складних умовах [13].

Огляд М. Сівача та С. Манна підкреслює важливість аналізу веб-журналів для підвищення ефективності моніторингу [14]. Модель HELAD, розроблена Ю. Чжуном, забезпечує високу точність завдяки використанню ансамблевого навчання, що робить її релевантною для великих мережевих систем [15]. С. Чжан розробив стійкий підхід до аналізу нестабільних веб-журналів, що сприяє зниженню впливу шумів і підвищенню ефективності алгоритмів [16].

Отже, аналіз наукових публікацій показує значні досягнення у сфері інтелектуального аналізу даних для виявлення аномалій. Отримані результати формують наукову основу для вдосконалення сучасних алгоритмів моніторингу веб-додатків.

У дослідженнях виявлення аномалій у веб-додатках залишаються невирішеними кілька важливих аспектів. По-перше, існуючі методи мають обмеження у адаптивності до динамічних умов веб-додатків, що знижує їхню ефективність. По-друге, необхідна розробка нових алгоритмів, здатних швидко обробляти великі обсяги даних і враховувати поведінкові особливості користувачів. По-третє, більшість робіт не підтверджуються емпіричними дослідженнями на реальних даних, що обмежує їхню практичну цінність. Крім того, точність і швидкість виявлення аномалій часто досліджуються окремо, без урахування їхньої взаємодії. Нарешті, відсутні конкретні рекомендації щодо інтеграції цих методів у реальні системи моніторингу.

Пропоноване дослідження спрямоване на усунення цих прогалин через розробку адаптивних алгоритмів, експериментальну перевірку їх ефективності на реальних даних і формулювання практичних рекомендацій для інтеграції. Це сприятиме як розвитку теоретичних основ, так і підвищенню ефективності моніторингу веб-додатків у реальних умовах.

Мета статті – аналіз ефективних методів виявлення аномалій у веб-додатках із використанням інтелектуального аналізу даних, спрямованих на підвищення рівня безпеки та функціональності сучасних веб-систем.

Завдання статті:

1. Проаналізувати існуючі підходи до виявлення аномалій у веб-додатках, визначити їхні переваги та обмеження, а також виявити ключові фактори, що впливають на точність і швидкість цих підходів.
2. Сформулювати рекомендації щодо інтеграції запропонованих методів в реальні системи моніторингу веб-додатків для підвищення їх ефективності та надійності.

Виклад основного матеріалу. Аналіз сучасних підходів до виявлення аномалій у веб-додатках демонструє їхню ключову роль у забезпеченні безпеки, функціональності та стабільності цих систем. У зв’язку зі стрімким зростанням обсягів даних, підвищенням складності архітектур веб-додатків і рівня загроз кібератак, виникає потреба у впровадженні нових, більш ефективних інструментів для моніторингу та виявлення аномальних дій. Аномалії у веб-додатках можуть бути спричинені різними чинниками, такими як технічні помилки, збої у програмному забезпеченні або цілеспрямовані атаки, і проявлятися у вигляді нетипових запитів, змін у структурі даних чи неочікуваних сплесків активності.

Для вирішення цих завдань було розроблено кілька підходів, які можна поділити на три основні категорії: методи на основі порогових значень, статистичні підходи та методи, що

базуються на інтелектуальному аналізі даних. Кожен із них пропонує унікальні можливості для аналізу аномалій, проте їхня ефективність залежить від специфіки застосування, характеру оброблюваних даних і обчислювальних ресурсів. Методи на основі порогових значень зосереджені на встановленні жорстких меж для визначення аномальної поведінки, тоді як статистичні підходи орієнтовані на аналіз розподілу даних і виявлення закономірностей. Найбільш сучасним і перспективним є використання методів інтелектуального аналізу даних, які застосовують алгоритми машинного навчання для ідентифікації складних і неочевидних патернів.

У таблиці 1 узагальнено основні характеристики, переваги та недоліки кожного з цих підходів, що дає змогу оцінити їхній потенціал для застосування у різних умовах.

Табл. 1

Основних підходи до виявлення аномалій у веб-додатках

Метод	Ключові особливості	Переваги	Недоліки
Порогові значення	Використання встановлених меж для виявлення аномалій.	Простота реалізації, швидкість виконання.	Обмежена гнучкість, неможливість адаптації до змін.
Статистичні методи	Аналіз розподілу даних та ідентифікація відхилень від нормального розподілу.	Можливість роботи з великими масивами даних.	Залежність від припущень про розподіл даних, складність реалізації.
Інтелекту-альний аналіз даних	Використання алгоритмів машинного навчання для виявлення складних патернів.	Висока точність, адаптивність до змін, автоматизація.	Велика обчислювальна складність, потреба у великій кількості даних для навчання.

Джерело: сформовано автором на підставі [2;3;6;7;8;9;10]

Порогові методи є ефективними для простих веб-додатків, де визначення аномалій базується на перевищенні фіксованих значень, наприклад, ліміту кількості запитів до сервера. У таких випадках система оперативно блокує підозрілу активність, проте не здатна враховувати складні сценарії поведінки користувачів [2, 10]. Наприклад, у веб-додатку для електронної комерції цей метод може не виявити аномалій у вигляді підроблених замовлень через нестандартну структуру запитів.

Статистичні методи забезпечують більш гнучкий підхід, дозволяючи враховувати ймовірнісну природу аномалій [6, 9]. Наприклад, у системі онлайн-банкінгу статистичний аналіз може допомогти ідентифікувати аномальну кількість транзакцій у певному часовому проміжку. Водночас залежність від правильного вибору моделі розподілу даних може знизити ефективність методу у випадках, коли дані не відповідають заданим припущенням.

Методи інтелектуального аналізу даних демонструють найвищу ефективність у сучасних умовах, оскільки здатні виявляти складні й неочевидні закономірності. Наприклад, використання нейронних мереж у системі моніторингу соціальних мереж дає змогу автоматично визначати аномальну поведінку ботів, адаптуючись до змін їхньої стратегії. Зокрема, у разі виявлення підозрілих публікацій алгоритм може аналізувати їхній контент, враховуючи не лише текстову інформацію, а й зображення або метадані, що значно підвищує рівень безпеки системи.

Автоматизоване виявлення аномалій у веб-додатках є важливим інструментом забезпечення їхньої стабільної роботи та кібербезпеки [3, 7, 13]. Традиційні підходи, такі як статичні порогові значення або стандартні методи класифікації, часто виявляються неефективними у сучасних умовах, коли дані стають дедалі складнішими, а загрози – більш динамічними та витонченими.

Запропоновані алгоритми розроблено на основі інтелектуального аналізу даних і враховують не лише явні відхилення, але й приховані закономірності, які складно ідентифікувати за допомогою звичайних методів. Вони базуються на інтеграції статистичного аналізу, машинного навчання та глибокого нейронного моделювання для досягнення максимальної точності. Особливістю цих алгоритмів є їхня адаптивність, масштабованість і здатність до самонавчання. Вони створені з урахуванням практичних потреб у виявленні аномалій у реальному часі, забезпечуючи швидке реагування на загрози та мінімізацію впливу на продуктивність системи.

Запропоновані алгоритми передбачають створення багаторівневих моделей для аналізу даних, що дозволяють гнучко поєднувати реактивні й проактивні підходи. Реактивний підхід спрямований на швидке виявлення аномалій шляхом аналізу вхідних потоків даних, тоді як проактивний підхід враховує історичні дані й дозволяє прогнозувати потенційні ризики. Алгоритми базуються на

використанні комбінованих методів, які адаптуються до змін у середовищі веб-додатків, що робить їх більш ефективними порівняно з існуючими рішеннями. У таблиці 2 наведено основні характеристики трьох розроблених алгоритмів.

Табл. 2

Алгоритми виявлення аномалій у веб-додатках

Назва алгоритму	Принцип роботи	Очікуваний результат
Адаптивний пороговий моніторинг	Використання змінних порогових значень, які адаптуються до трендів у потоках даних.	Виявлення аномалій у режимі реального часу із мінімізацією хибнопозитивних сигналів.
Модуль часової кореляції	Аналіз часових рядів для ідентифікації нетипових змін, враховуючи залежність між різними параметрами.	Визначення складних аномалій, які проявляються у часовій динаміці.
Гібридна кластеризація	Поєднання кластеризації та факторного аналізу для сегментації даних і виявлення аномалій у групах.	Висока точність у багатовимірних середовищах із різнорідними наборами даних.

Джерело: власна розробка автора

Запропоновані алгоритми мають високий потенціал для ефективного застосування у різних умовах. Адаптивний пороговий моніторинг може бути використаний у веб-додатках із динамічними потоками даних, таких як сервіси потокового мовлення. Передбачається, що система автоматично регулюватиме порогові значення залежно від часу доби, активності користувачів та інших факторів. Це дає змогу виявляти аномальні сплески трафіку, які можуть свідчити про DDoS-атаки або підозрілу активність, забезпечуючи стабільну роботу системи без втрат продуктивності.

Модуль часової кореляції має перспективи для застосування в аналітичних платформах, що використовуються у фінансових операціях, де необхідно враховувати залежність між показниками, такими як обсяги транзакцій, географічне розташування та час здійснення операцій. Очікується, що цей алгоритм дозволить виявляти серії транзакцій, які відхиляються від звичайного патерну, та автоматично ідентифікувати підозрілі дії, що сприятиме запобіганню шахрайству.

Гібридна кластеризація може знайти своє застосування у складних системах моніторингу великих корпоративних веб-додатків, таких як ERP-системи. Прогнозується, що цей алгоритм забезпечить сегментацію користувачів за їхньою поведінкою та аналіз кожної групи на наявність нетипових дій. Наприклад, у групі користувачів із доступом до конфіденційних даних алгоритм зможе виявляти незвичну активність, яка може вказувати на потенційний витік даних.

Для перевірки ефективності запропонованих алгоритмів виявлення аномалій було проведено експеримент, що імітував реальні умови роботи веб-додатків. Основною метою було оцінити здатність алгоритмів виявляти аномалії, швидкість їхньої роботи та кількість хибнопозитивних спрацювань. Для цього використовувалися дані, отримані з веб-середовища, створеного на основі симуляції реальної активності користувачів. Дані включали нормальні дії, а також імітовані аномалії, такі як нестандартні запити, пікове навантаження та технічні збої.

Симуляційне середовище було побудоване на основі простого веб-додатку з базовими функціями, включаючи автентифікацію, перегляд контенту та виконання транзакцій. Дані збиралися у вигляді потоків запитів, які генерувалися за допомогою скриптів, що імітували дії реальних користувачів, та спеціально доданих аномальних дій. Для забезпечення реалізму використовувалися змінні параметри, такі як час доби, інтенсивність запитів та частота повторюваних дій. Обсяг даних складав близько 10 000 запитів, з яких 5% були аномальними.

Алгоритми тестувалися в однакових умовах. Для адаптивного порогового моніторингу порогові значення визначалися динамічно, залежно від попередньо зібраних трендів. Модуль часової кореляції використовував часові ряди для аналізу залежності між різними параметрами запитів, такими як час надсилання, тип дії та IP-адреса. Гібридна кластеризація виконувала сегментацію користувачів за поведінковими патернами, після чого визначалися аномальні групи (табл. 3).

Отримані результати свідчать, що модуль часової кореляції продемонстрував найвищу точність, виявивши найбільшу кількість аномалій із мінімальним рівнем хибнопозитивних спрацювань. Проте час його обробки був довшим, що обмежує його застосування у реальному часі. Адаптивний пороговий моніторинг забезпечив найшвидший час реагування, але мав дещо вищий рівень хибнопозитивних спрацювань, що може спричиняти зайві реакції системи. Гібридна

кластеризація забезпечила збалансовані результати за всіма показниками, що робить її універсальним рішенням для складних багатовимірних середовищ.

Табл. 3

Науково-практична оцінка ефективності алгоритмів для виявлення аномалій у веб-додатках

Метод	Застосування	Кількість виявлених аномалій	Хибнопозитивні спрацювання (%)	Середній час обробки (мс)
Адаптивний пороговий моніторинг	Використання змінних порогових значень для динамічного аналізу потоків даних.	85	12	45
Модуль часової кореляції	Аналіз залежності у часових рядах для виявлення нетипових змін.	92	8	70
Гібридна кластеризація	Поєднання кластеризації та аналіз факторів для сегментації даних.	89	10	65

Джерело: власна розробка автора

Точність і швидкість виявлення аномалій у веб-додатках залежать від низки ключових чинників, серед яких якість вхідних даних, обчислювальні ресурси, налаштування алгоритмів і характер самих аномалій [9, 12]. Перш за все, якість даних відіграє вирішальну роль, оскільки наявність шумів, пропущених значень або некоректних даних може значно знижувати ефективність алгоритмів [14]. Наприклад, якщо дані не були попередньо нормалізовані або містять некоректні часові мітки, це може призвести до помилкової інтерпретації поведінки системи як аномальної. Водночас надто великі обсяги даних із високим рівнем варіативності можуть створювати надмірне навантаження на алгоритм, уповільнюючи його роботу.

Другим важливим чинником є обчислювальні ресурси, доступні для роботи алгоритмів. Методи, що базуються на складних моделях, таких як глибокі нейронні мережі чи багатовимірна кластеризація, потребують значних ресурсів для обробки великих потоків даних у реальному часі [7, 8]. Брак ресурсів може спричиняти затримки у виявленні аномалій, що критично у системах, де потрібна миттєва реакція на загрози. Оптимізація алгоритмів під конкретні апаратні умови може стати викликом, особливо для малих або середніх підприємств із обмеженим доступом до потужних обчислювальних систем.

Ще одним чинником є параметри навчання алгоритмів, які визначають їхню здатність до адаптації та точність виявлення аномалій. Вибір неправильних параметрів, таких як розмір кластерів або порогові значення, може спричиняти високий рівень хибнопозитивних або хибнонегативних результатів. Наприклад, занадто жорсткі порогові значення можуть призводити до того, що нормальна поведінка інтерпретується як аномальна, створюючи зайві сповіщення та знижуючи ефективність системи. З іншого боку, надто широкі межі знижують чутливість алгоритмів, що ускладнює своєчасне виявлення дійсно небезпечних подій. Характер аномалій є важливим аспектом, який впливає на ефективність алгоритмів. Простежується різниця між явними аномаліями, які легко ідентифікувати, та прихованими аномаліями, що мають складні патерни або проявляються лише у взаємозв'язках між різними параметрами [15]. Сучасні методи виявлення часто демонструють високу ефективність для перших, але стикаються з труднощами у роботі з останніми. Наприклад, складно виявити аномалії, які проявляються лише у часових рядах, якщо алгоритм не враховує тимчасові залежності між подіями.

Для інтеграції розроблених методів виявлення аномалій у реальні системи моніторингу веб-додатків слід дотримуватися послідовного підходу, що включає технічну та організаційну підготовку, а також забезпечення безперервної адаптації. Перш за все, рекомендується реалізувати прототипи алгоритмів у середовищі тестування, максимально наближеному до умов реальної роботи системи. Це дозволить оцінити ефективність методів і визначити оптимальні параметри для їх подальшого налаштування.

Наступним кроком є інтеграція алгоритмів у виробниче середовище з урахуванням існуючої архітектури веб-додатків. Особливу увагу слід приділити сумісності алгоритмів із поточними системами моніторингу та безпеки. Для забезпечення стабільної роботи рекомендується використовувати модульний підхід, який дає змогу додавати або замінювати алгоритми без впливу на загальну функціональність системи.

Одним із важливих практичних кроків є створення системи автоматичного оновлення моделей на основі нових даних. Це може бути реалізовано через періодичне навчання алгоритмів із використанням останніх зібраних даних, що забезпечить їхню адаптивність до змін у поведінці користувачів або структурі запитів. Крім цього, слід передбачити резервні механізми для реагування на збої або перевантаження системи, такі як ручне коригування порогових значень або тимчасове відключення окремих модулів.

Для ефективного впровадження рекомендується розробити докладну документацію, яка міститиме опис функціональності алгоритмів, їхніх параметрів та інструкції для адміністраторів системи. Важливим аспектом також є навчання персоналу, відповідального за роботу з системою, для забезпечення розуміння принципів її функціонування та можливих сценаріїв реагування на аномалії.

Крім того, слід забезпечити інтеграцію системи моніторингу з панелями візуалізації, які дозволяють відстежувати тренди, аналізувати результати роботи алгоритмів та швидко приймати рішення у випадку виявлення загроз. Такі інструменти можуть включати графіки активності, списки підозрілих запитів або сповіщення про критичні аномалії. Це сприятиме комплексному аналізу роботи системи та підвищенню її ефективності.

Узагальнюючи викладене, рекомендації спрямовані на розробку системи моніторингу веб-додатків, яка відзначається адаптивністю, надійністю та простотою у використанні. Така система має ефективно виявляти аномалії, забезпечуючи своєчасну реакцію на загрози, мінімізацію ризиків та підвищення рівня безпеки в умовах сучасних викликів. Така система має ефективно виявляти аномалії, забезпечуючи своєчасну реакцію на загрози, мінімізацію ризиків та підвищення рівня безпеки в умовах сучасних викликів.

Висновки. Таким чином, встановлено, що методи інтелектуального аналізу даних є ефективними для виявлення аномалій у веб-додатках завдяки їхній здатності адаптуватися до складних і динамічних середовищ. Запропоновані алгоритми, такі як адаптивний пороговий моніторинг, модуль часової кореляції та гібридна кластеризація, забезпечують врахування різноманітних типів аномалій, досягаючи високої точності та швидкості аналізу. Результати експериментальної перевірки підтвердили, що ці методи ефективно виявляють складні патерни та мінімізують хибнопозитивні сигнали, що є ключовими для забезпечення функціональності та безпеки веб-додатків.

Основними проблемами, виявленими під час дослідження, є висока обчислювальна складність методів інтелектуального аналізу даних, необхідність попередньої обробки великих обсягів інформації та складність налаштування параметрів алгоритмів під конкретні сценарії використання. Також наголошено на важливості якісного джерела даних, оскільки наявність шумів і некоректної інформації суттєво знижує ефективність роботи алгоритмів.

У рекомендаціях підкреслено важливість попереднього тестування алгоритмів у реалістичних умовах, застосування гібридного підходу, що поєднує кілька методів для підвищення адаптивності системи, а також інтеграції моделей машинного навчання із системами моніторингу в реальному часі. Особливу увагу слід приділяти розробці модульних рішень, які забезпечують гнучке налаштування алгоритмів відповідно до змін у середовищі веб-додатків.

Перспективи подальших досліджень пов’язані з оптимізацією обчислювальних процесів для зменшення витрат ресурсів, розвитком методів автоматизованого навчання моделей і створенням інтегрованих систем моніторингу, які можуть ефективно працювати у великих багатокомпонентних середовищах. Крім того, важливим напрямом є вивчення впливу нових типів аномалій, що виникають у зв’язку з розвитком технологій і змінами у поведінці користувачів.

Список використаних джерел:

1. State Information Security as a Challenge of Information and Computer Technology Development / K. Chyzhmar et al. *Journal of Security and Sustainability Issues*. 2020. P. 819–828. URL: https://doi.org/10.9770/jssi.2020.9.3(8) (date of access: 29.12.2024).

2. Unsupervised Anomaly Detection via Variational Auto-Encoder for Seasonal KPIs in Web Applications / H. Xu et al. *the 2018 World Wide Web Conference,* Lyon, France, 23–27 April 2018. New

York, New York, USA, 2018. URL: https://doi.org/10.1145/3178876.3185996 (date of access: 29.12.2024).

3. Connecting Web Event Forecasting with Anomaly Detection: A Case Study on Enterprise Web Applications Using Self-supervised Neural Networks / X. Yuan et al. *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*. Cham, 2020. P. 481–502. URL: https://doi.org/10.1007/978-3-030-63086-7_27 (date of access: 29.12.2024).

4. Weakly Supervised Anomaly Detection via Knowledge-Data Alignment / H. Zhao et al. *WWW '24: The ACM Web Conference 2024*, Singapore Singapore. New York, NY, USA, 2024. URL: https://doi.org/10.1145/3589334.3645429 (date of access: 29.12.2024).

5. Juan D.-C., Shah N., Tang M., Qian Z., Marculescu D., Faloutsos C. M3A: Model, MetaModel, and Anomaly Detection in Web Searches. *arXiv*. 2016. 10 p. DOI: https://doi.org/10.48550/arXiv.1606.05978 (date of access: 29.12.2024).

6. Aggarwal C. C. Outlier Ensembles. *Outlier Analysis*. Cham, 2016. P. 185–218. URL: https://doi.org/10.1007/978-3-319-47578-3_6 (date of access: 29.12.2024).

7. Chalapathy R., Chawla S. Deep Learning for Anomaly Detection: A Survey. *arXiv*. 2019. 34 p. DOI: https://doi.org/10.48550/arXiv.1901.03407 (date of access: 29.12.2024).

8. Deep Learning for Anomaly Detection in Time-Series Data: Review, Analysis, and Guidelines / K. Choi et al. *IEEE Access*. 2021. Vol. 9. P. 120043–120065. URL: https://doi.org/10.1109/access.2021.3107975 (date of access: 29.12.2024).

9. Machine Learning in Network Anomaly Detection: A Survey / S. Wang et al. *IEEE Access*. 2021. Vol. 9. P. 152379–152396. URL: https://doi.org/10.1109/access.2021.3126834 (date of access: 29.12.2024).

10. Gao Y., Ma Y., Li D. Anomaly detection of malicious users' behaviors for web applications based on web logs. *2017 IEEE 17th International Conference on Communication Technology (ICCT)*, Chengdu, 27–30 October 2017. 2017. URL: https://doi.org/10.1109/icct.2017.8359854 (date of access: 29.12.2024).

11. Graph based Tensor Recovery for Accurate Internet Anomaly Detection / K. Xie et al. *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, Honolulu, HI, 16–19 April 2018. 2018. URL: https://doi.org/10.1109/infocom.2018.8486332 (date of access: 29.12.2024).

12. An efficient hybrid system for anomaly detection in social networks / M. S. Rahman et al. *Cybersecurity*. 2021. Vol. 4, no. 1. URL: https://doi.org/10.1186/s42400-021-00074-w (date of access: 29.12.2024).

13. Deep Learning for Anomaly Detection / R. Wang et al. *WSDM '20: The Thirteenth ACM International Conference on Web Search and Data Mining*, Houston TX USA. New York, NY, USA, 2020. URL: https://doi.org/10.1145/3336191.3371876 (date of access: 29.12.2024).

14. Siwach M., Mann S. Anomaly detection for web log data analysis: A review. *Journal of Algebraic Statistics*. 2022. Vol. 13. No. 1. P. 129–148. URL: https://publishoa.com/index.php/journal/article/view/68 (date of access: 29.12.2024).

15. HELAD: A novel network anomaly detection model based on heterogeneous ensemble learning / Y. Zhong et al. *Computer Networks*. 2020. Vol. 169. P. 107049. URL: https://doi.org/10.1016/j.comnet.2019.107049 (date of access: 29.12.2024).

16. Robust log-based anomaly detection on unstable log data / X. Zhang et al. *ESEC/FSE '19: 27th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, Tallinn Estonia. New York, NY, USA, 2019. URL: https://doi.org/10.1145/3338906.3338931 (date of access: 29.12.2024).