УДК 621.391.8+004.832
DOI: 10.31673/2412-9070.2025.046209
О. С. ВЕЛІГУРСЬКИЙ, аспірант,
ORCID: 0009-0007-8852-7948
Державний університет інформаційно-комунікаційних технологій, Київ

АНАЛІЗ СУЧАСНИХ МЕТОДІВ ВИЯВЛЕННЯ ТА УСУНЕННЯ АНОМАЛІЙ В ІНФОКОМУНІКАЦІЙНИХ МЕРЕЖАХ В УМОВАХ ВИСОКОГО НАВАНТАЖЕННЯ

У статті здійснюється комплексний аналіз сучасних методів виявлення та усунення аномалій в інфокомунікаційних мережах, з урахуванням актуальних проблем, пов'язаних із динамічним масштабуванням трафіку, високою інтенсивністю передавання даних та необхідністю забезпечення стійкої роботи мережної інфраструктури в режимі реального часу. Постійне зростання навантаження на інфокомунікаційні системи, через активне використання потокового відео, хмарних обчислень та мобільних сервісів, призводить до збільшення частоти появи аномальних подій. Такі події можуть спричинити втрату пакетів, підвищення затримок або повну втрату зв'язку. Це актуалізує потребу у впровадженні високоефективних систем виявлення порушень, здатних адаптуватися до змін у мережному середовищі. У традиційних підходах часто використовуються фіксовані порогові значення, сигнатури або евристичні правила, що істотно обмежує їх ефективність у динамічних умовах. Такі системи не здатні оперативно реагувати на невідомі загрози або аномальні патерни трафіку, що не були передбачені на етапі налаштування. У роботі розглянуто широкий спектр сучасних методів: від класичних статистичних моделей до алгоритмів машинного навчання. Зокрема, акцент зроблено на методах кластеризації, байєсівських мережах, деревоподібних моделях Decision Tree та Random Forest, а також підходах на основі часових рядів ARIMA та LSTM. Оцінка методів проводиться з позицій відповідності сучасним вимогам: здатності до масштабування, роботи в реальному часі, інтеграції з системами автоматичного управління та мінімізації помилкових спрацьовувань. У процесі аналізу виявлено, що більшість існуючих рішень демонструють недостатню здатність до генералізації, високу чутливість до налаштувань і складність адаптації до змінного навантаження.

Обґрунтувано перспективу використання гібридних систем, що комбінують кілька різних методів, включаючи Autoencoders, GANs, VAE та алгоритми Isolation Forest. Такий підхід дозволяє суттєво підвищити адаптивність і точність виявлення аномалій навіть у складних і непередбачуваних умовах. Результати* роботи *є практично значущими для подальшої розробки інтелектуальних систем моніторингу трафіку, що враховують описані проблеми в умовах критичного навантаження.

Ключові слова: інфокомунікаційні мережі; аномалії; адаптивність; прогнозування; масштабованість; навантаження; патерни; машинне навчання; аналіз трафіку; гібридні моделі.

Вступ

Постановка проблеми. Інфокомунікаційні мережі є критично важливими компонентами сучасної цифрової інфраструктури, забезпечуючи обмін інформацією між мільйонами користувачів та пристроїв. Постійне зростання обсягу переданих даних, зокрема через поширення відеостримінгових сервісів, хмарних технологій та мобільних додатків, призводить до суттєвого підвищення навантаження на мережі. Наявні методи розпізнавання мережних відхилень здебільшого ґрунтуються на статистичних підходах, сигнатурному аналізі, простих евристиках або фіксованих порогових значеннях. Вони добре справляються з типовими шаблонами порушень, проте втрачають ефективність в умовах швидкозмінного та непередбачуваного середовища. Технології розпізнавання, що застосовуються на практиці не завжди виявляють

нові або приховані типи аномалій, також демонструють високу кількість хибнопозитивних спрацьовувань і часто не масштабуються до великих та розподілених мережних середовищ.

У цьому контексті особливо актуальним стає критичний аналіз сильних і слабких сторін існуючих підходів, зокрема, методів на основі кластеризації, ентропійного аналізу, дерев рішень, байєсівських моделей та нейронних мереж. Таке дослідження не лише дозволить ідентифікувати обмеження та вразливості існуючих методів, але й створить підґрунтя для формулювання вимог до нових поколінь інтелектуальних систем виявлення та усунення аномалій, які мають забезпечувати високу адаптивність, точність та оперативну ро-боту в умовах реального часу.

Аналіз останніх досліджень. Аномалії в інфокомунікаційних мережах традиційно розглядаються як нетипові прояви у поведінці системи, що мають відхилення від встановленої або очікуваної моделі нормального функціонування. Так, автори Varun Chandola та Vipin Kumar (2009) визначають аномалії як спостереження, що не відповідають жодній із домінантних моделей даних, сформованих на основі історичного контексту [1]. У той час як Patcha і Park (2007) у сфері мережної безпеки трактують аномалії як відхилення у шаблонах трафіку, які можуть свідчити про вторгнення або зловмисну активність [2]. Натомість автори статті про виявлення аномалій у мережах акцентують увагу на часовій природі, описуючи їх як відхилення у часових рядах параметрів, що можуть вказувати як на технічні несправності, так і на не-типову поведінку користувачів [3]. Усі ці визначення відображають важливі аспекти, від контенту і поведінки до часових характеристик і контексту. У цьому дослідженні аномалія визначається як стійке або короткочасне, але статистично або структурно значуще відхилення параметрів мережного трафіку або поведінкових патернів від встановленої або динамічно адаптованої моделі нормального функціонування системи. Це допоможе довести необхідність дослідження адаптивних моделей, здатних працювати в умовах реального часу та пікового навантаження.

Постановка задачі. Метою даної статті є аналіз сучасних способів ідентифікації та нейтралізації аномальних подій в інфокомунікаційних мережах в умовах високого навантаження. Наукова праця спрямована на аналіз існуючих недоліків і обмежень технологій, що використовуються, а також на формулювання вимог до перспективних рішень, здатних забезпечити високу точність, адаптивність і стабільну роботу у реальному часі в умовах змінного мережного середовища.

Для досягнення поставленої мети у вирішуються наступні задачі:

- Провести огляд і класифікацію існуючих методів виявлення та усунення аномалій у сфері інфокомунікаційних технологій.
- Визначити критерії ефективності для таких методів у контексті високонавантажених мереж, таких як: затримки, пропускна здатність, хибнопозитивні та хибнонегативні спрацьовування, масштабованість, точність та адаптивність.
- Порівняти популярні моделі за обраними критеріями на основі: правил, сигнатур, статистичних моделей, методів кластеризації та байєсівських мереж.
- Обґрунтувати напрями майбутніх досліджень, спрямованих на подолання виявлених недоліків та підвищення ефективності виявлення та реагування на аномалії.

Основна частина

У сучасних інфокомунікаційних мережах своєчасна оцінка аномальних подій та ефективне реагування на них є критично важливими завданнями, з огляду на зростаючу складність інфраструктури, обсяги трафіку та зростаючі вимоги до надійності. Методи виявлення ненормальних подій поділяються на кілька ключових категорій, кожна з яких має власні переваги, недоліки та області застосування.

Таблиця 1

Класифікація основних методів виявлення аномалій в ІКМ

Категорія	Опис	Переваги	Недоліки	Приклади
Статистичні методи	Виявлення відхилень на основі ймовірнісних моделей нормального трафіку.	Простота реалізації, добре працюють у стабільних умовах.	Погано пристосовані до змін трафіку, висока частота хибних результатів.	Z-score, гістограми, коваріаційний аналіз
Сигнатурні методи	Пошук шаблонів, що відповідають відомим аномаліям або атакам.	Висока точність для відомих загроз.	Не виявляють нові або модифіковані загрози.	Snort, Suricata
Порогові методи	Застосування жорстких меж (порогів) до метрик трафіку.	Швидкодія, низькі вимоги до обчислювальних ресурсів.	Негнучкість, висока чутливість до помилкових спрацьовувань.	NetFlow Threshold Alerts
Методи кластеризації	Групування схожих аномалій визначаються як ті, що не входять до жодного кластеру.	Можуть працювати без розмічених даних.	Вразливі до вибору гіпер-параметрів, складність масштабування.	K-means, DBSCAN
Методи на основі графів	Аналіз мережної поведінки за допомогою топологічних моделей або зв'язків між вузлами.	Ефективні для аналізу складних структур взаємодій.	Висока обчислювальна складність, важкість побудови моделей.	Graph Neural Networks (GNN)

Методи машинного навчання	Включають як традиційні алгоритми (SVM, Decision Trees), так і глибинне навчання.	Висока адаптивність, здатність до самонавчання.	Потребують великих обсягів даних, складність представлення результатів.	Random Forest, LSTM, Autoencoders
Гібридні методи	Комбінація кількох підходів для підвищення точності.	Підвищена стійкість до хибних рішень.	Збільшення складності системи, потреба у балансуванні компонентів.	SIEM-системи з ML-модулями

Попри широке розмаїття методів, більшість з них мають спільні недоліки, особливо в умовах високого навантаження:

- Відсутність адаптивності: традиційні методи часто не враховують динаміку мережного трафіку.
- Низька точність у непередбачуваних ситуаціях: сигнатурні та порогові методи ефективні лише для відомих шаблонів.
- Високий рівень хибнопозитивних рішень: особливо у статистичних моделях, що базуються на фіксованих порогах.
- Недостатня масштабованість: багато алгоритмів не придатні для обробки великих обсягів трафіку у реальному часі.
- Складність інтерпретації результатів: особливо актуально для моделей на основі нейронних мереж.

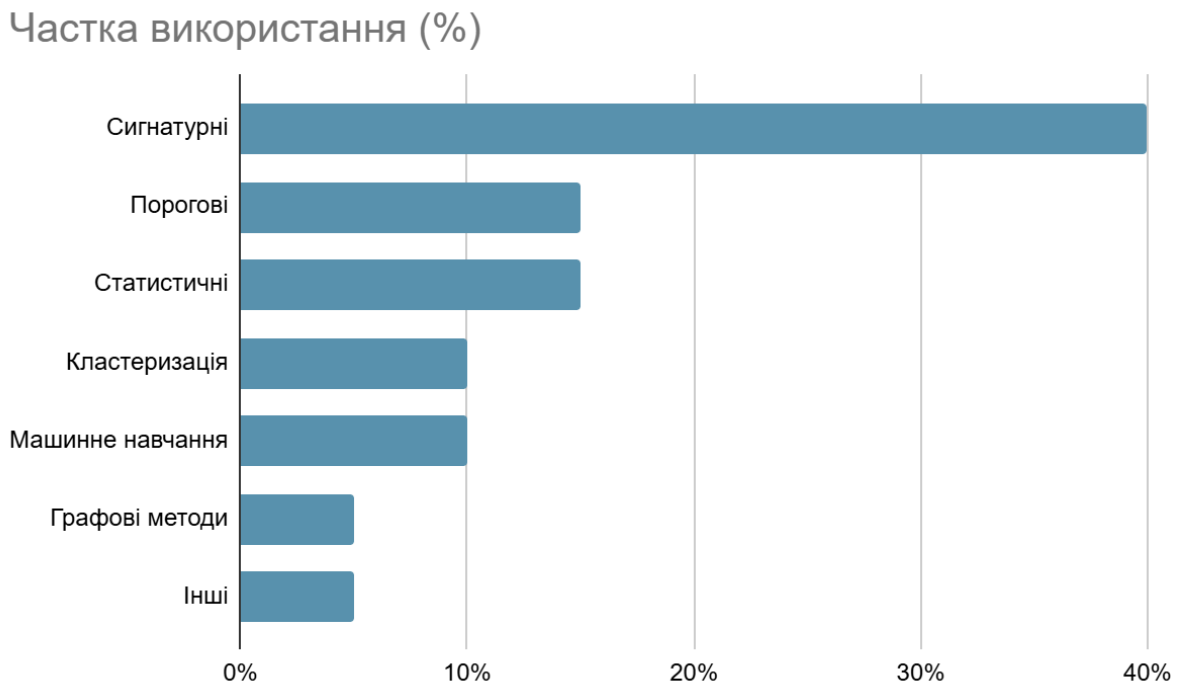

Рис. 1. Приблизне розповсюдження методів виявлення аномалій у сучасних системах [4]

У результаті проведеного огляду можна зробити висновок, що незважаючи на наявність великої кількості підходів, жоден з них не є універсальним. Для оцінки ефективності методів виявлення та усунення аномалій в інфокомунікаційних мережах необхідно враховувати кілька

важливих критеріїв, що прямо впливають на їх здатність функціонувати в умовах високої інтенсивності роботи. Оскільки високонавантажені мережі характеризуються великими обсягами переданих даних, змінними умовами та постійними загрозами безпеки, то відповідні методи повинні задовольняти специфічні вимоги до продуктивності та точності. Основними критеріями ефективності є: затримка, пропускна здатність, кількість хибнопозитивних та хибнонегативних подій, масштабованість, точність та адаптивність.

Затримка є одним з основних параметрів для методів виявлення аномалій. У високонавантажених мережах кожна секунда затримки може серйозно вплинути на ефективність управління ресурсами та час реагування на загрози. Високий рівень затримки може призвести до пропуску небажаної активності або навіть до негативного впливу на роботу інших процесів у мережі. Метод фіксації поведінкових патернів повинен мати мінімальну затримку, особливо у випадках екстремальних пікових навантажень.

Пропускна здатність визначає, скільки даних можна обробити за одиницю часу. Для методів виявлення аномалій в ІКМ важливо, щоб обробка даних не обмежувала загальні пропускні якості мережі. Занадто велике навантаження на обробку даних може знизити ефективність мережі в цілому, тобто алгоритм має обробляти великі обсяги трафіку без помітного зниження пропускної здатності.

Хибнопозитивні та хибнонегативні спрацьовування - це випадки, при яких система помилково ідентифікує нормальний трафік, як відхилення, та коли система не виявляє реальні відмінності поміж трафіку, відповідно. У високонавантажених мережах важливо підтримувати баланс між цими двома типами. Високий рівень хибнопозитивних викликів може призвести до перевантаження системи та додаткових витрат часу на обробку хибних сигналів. Високий рівень хибнонегативних реакцій ставить під загрозу безпеку мережі, оскільки аномалії можуть залишитися непоміченими.

Масштабованість є критично важливою для ефективної роботи методів виявлення аномалій у великих та динамічних мережах, таких як 5G або IoT-системи, де кількість підключених пристроїв може бути величезною, а обсяг трафіку змінюється дуже швидко. Методи, що не мають достатньої масштабованості, можуть призвести до значного зниження продуктивності або не здатності ефективно працювати з новими, більш складними мережами. Рішення повинне ефективно масштабуватися на більші обсяги трафіку і кількість підключених пристроїв.

Точність встановлення нехарактерної активності та здатність адаптуватися до змін у мережних умовах (наприклад, підвищення навантаження, зміни в архітектурі мережі, нові загрози) також є важливими критеріями. Адаптивність дозволяє системам ефективно реагувати на зміни без необхідності переналаштування або повторного навчання.

Таблиця 2

Ключові критерії ефективності для методів виявлення аномалій в умовах високонавантажених мереж

Критерій	Опис	Вимоги для ефективності	Приклади методів
Затримка	Час, необхідний для виявлення аномалії після її виникнення.	Мінімальна затримка у реальному часі для швидкої реакції на загрози.	Алгоритми з низьким часом відгуку, адаптивні моделі.

Пропускна здатність	Обсяг даних, який може бути оброблений без зниження продуктивності.	Висока пропускна здатність при великих навантаженнях.	Статистичні методи, прості алгоритми.
Хибнопозитивні / хибнонегативні спрацьовування	Кількість неправильних спрацьовувань системи.	Баланс між FP та FN для мінімізації витрат часу та збереження безпеки.	Алгоритми на основі ML з адаптивними порогами.
Масштабованість	Здатність методів обробляти великі обсяги даних і працювати з великою кількістю вузлів.	Масштабованість для обробки великої кількості підключених пристроїв.	Розподілені системи, алгоритми машинного навчання.
Точність та адаптивність	Здатність правильно виявляти аномалії і адаптуватися до змін у мережі.	Висока точність і здатність адаптуватися до нових загроз та умов.	Глибинне навчання, гібридні моделі.

Визначені критерії ефективності, такі як час реакції, пропускна здатність, рівень хибнопозитивних і хибнонегативних спрацьовувань, масштабованість та адаптивність, дозволяють здійснити змістовне порівняння різних підходів до виявлення аномалій. Серед найпоширеніших методів, що використовуються в інфокомунікаційних мережах, виділяють підходи на основі правил, сигнатур, статистичного аналізу, кластеризації та байєсівських мереж. Кожен із них має свої переваги та обмеження, які особливо проявляються в умовах високого навантаження.

Таблиця 3

Порівняння методів виявлення аномалій за ключовими критеріями ефективності [5]

Критерій	**На основі правил**	**Сигнатурний підхід**	**Статистичні моделі**	**Кластеризація**	**Байєсівські мережі**	**(Autoencoders, CNN)**	**LSTM**
Час реакції	Високий	Високий	Середній	Середній	Низький	Низький	Низький
Хибнопозитивні спрацьовування	Високі	Низькі	Середні	Високі	Низькі	Низькі	Низькі

Хибнонегативні спрацьовування	Низькі	Високі	Середні	Високі	Середні	Низькі	Низькі
Адаптивність до нових аномалій	Низька	Низька	Середня	Висока	Висока	Висока	Висока
Масштабованість	Висока	Середня	Висока	Середня	Середня	Висока	Середня
Залежність від набутих даних	Низька	Середня	Висока	Висока	Висока	Висока	Висока
Складність впровадження	Низька	Низька	Середня	Висока	Висока	Висока	Висока

Можна зробити висновок, що методи на основі правил та сигнатур швидко реагують на загрози, однак надто залежать від заздалегідь відомих шаблонів, тому не дуже підходять для виявлення нових або модифікованих загроз. Вони можуть бути використані як перший рівень фільтрації, але не як самостійна система захисту.

Статистичні методи добре працюють при моделюванні звичайної поведінки трафіку. Вони здатні виявляти деякі типи відхилень, але часто потребують регулярного налаштування параметрів [3].

Методи кластеризації, як правило, не вимагають наявності міток для тренування, що є перевагою в умовах нестачі даних. Проте ці методи мають вищу затримку й потребують складної інтерпретації результатів [6].

Байєсівські мережі демонструють високу точність і добре підходять для складного контекстного аналізу. Проте вони менш ефективні у реальному часі через високу обчислювальну складність [7].

Методи глибинного навчання, зокрема автокодувальники та згорткові нейронні мережі (CNN) забезпечують високу точність виявлення аномалій навіть у випадках складних багатовимірних патернів. Завдяки здатності навчатися без нагляду або з частковою наглядовою інформацією, ці методи ефективно працюють в умовах змінного трафіку. Проте їх використання потребує значних обчислювальних ресурсів, що ускладнює розгортання у реальному часі на периферійних пристроях [8].

Рекурентні нейронні мережі типу LSTM показують високу ефективність при роботі з часовими рядами, особливо для виявлення нестандартної поведінки, що проявляється у вигляді нестандартних змін трафіку. LSTM здатні зберігати контекст довготривалих залежностей, що критично важливо для телекомунікаційних мереж. Разом з тим, через складну структуру та високі вимоги до навчання, їх впровадження потребує додаткових оптимізаційних рішень та ресурсів [9].

Порівняння показує, що жоден із традиційних методів не здатен комплексно задовольнити вимоги до точності, адаптивності, масштабованості та швидкодії, які необхідні для стабільної роботи інфокомунікаційних мереж в умовах високого навантаження. Це вимагає пошуку нових архітектур, які можуть об'єднувати переваги різних підходів у межах єдиної системи. У цьому контексті перспективним напрямом є використання сучасних моделей машинного

навчання, зокрема Generative Adversarial Networks (GANs), Variational Autoencoders (VAE) та Isolation Forest, які відкривають можливості для побудови гібридних, самонавчальних систем розпізнавання аномалій. Саме на інтеграції таких технологій пропонується зосередити подальші дослідження.

Висновки

Майбутні дослідження, спрямовані на подолання недоліків у системах розпізнавання аномалій в інфокомунікаційних мережах, будуть зосереджені на комбінуванні новітніх підходів у сфері машинного навчання, таких як Generative Adversarial Networks (GANs), Variational Autoencoders (VAE) та Isolation Forest, щоб подолати основні обмеження традиційних методів. Ключовими проблемами існуючих систем є: залежність від заздалегідь визначених шаблонів, що обмежує їх здатність виявляти нові або невідомі загрози, масштабованість в умовах високого навантаження та проблема адаптивності до нових або модифікованих загроз. GANs здатні генерувати нові зразки даних, які наближаються до реальних мережних відхилень, що дозволяє навчати моделі на синтетичних даних, покращуючи їх здатність до ідентифікації незнайомих атак. Завдяки цьому, цей підхід може бути корисним для створення нових навчальних даних, що дозволить підвищити точність моделей виявлення аномалій. Інтеграція GANs із Autoencoders може допомогти розв'язати проблему хибнопозитивних спрацьовувань, оскільки це поєднання дає змогу виявляти та реконструювати аномальні патерни з більшою гнучкістю. Зокрема, застосування VAE для стискання та реконструкції вхідних даних дасть можливість отримати гнучкішу модель, що не залежить від конкретних шаблонів і може адаптуватися до змінного трафіку. Isolation Forest, який добре працює при виявленні відхилень у великих наборах даних, може бути використаний разом з іншими методами для виявлення аномалій в умовах високого навантаження. У поєднанні з іншими алгоритмами, такими як LSTM, для прогнозування майбутніх значень параметрів мережі Isolation Forest дозволить зменшити час реакції та підвищити ефективність аналізу нетипової поведінки у реальному часі. Впровадження Reinforcement Learning (RL) у поєднанні з VAE або GANs може дозволити системам самостійно оптимізувати параметри для виявлення порушень у роботі системи, адаптуючись до нових умов без необхідності ручного втручання. Цей підхід може бути особливо корисним у складних та динамічних мережах, де загрози можуть з'являтися неочікувано.

Така інтеграція різних методів, таких як GANs, VAE, Isolation Forest, LSTM та Reinforcement Learning створить потужну, гнучку систему для виявлення та усунення аномалій, здатну ефективно працювати в умовах високого навантаження та швидко адаптуватися до нових типів атак. Крім того, поєднання цих методів дозволить досягти необхідної масштабованості, зменшити кількість хибнопозитивних спрацьовувань та підвищити точність системи.

Список літератури

1. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), Article 15, pp. 1–72. https://doi.org/10.1145/1541880.1541882

2. Patcha, A., & Park, J. M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. Computer Networks, 51(12), pp. 3448–3470. https://doi.org/10.1016/j.comnet.2007.02.001

3. Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, pp. 19–31. https://doi.org/10.1016/j.jnca.2015.11.016

4. Fahim, M., & Sillitti, A. (2019). Anomaly Detection, Analysis and Prediction Techniques in IoT Environment: A systematic literature review, pp. 1-8. https://ieeexplore.ieee.org/document/8885651

5. Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. IEEE Symposium on Security and Privacy, pp. 305–316. https://doi.org/10.1109/SP.2010.25

6. Jain, A. K. (2010). Data clustering: 50 years beyond K-means. Pattern Recognition Letters, 31(8), pp. 651-666. https://doi.org/10.1016/j.patrec.2009.09.011

7. Heckerman, D., Geiger, D., & Chickering, D. M. (1995). Learning Bayesian networks: The combination of knowledge and statistical data. Machine Learning, 20(3), pp. 197–243. https://doi.org/10.1007/BF00994053

8. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press, p. 800. https://deeplearningbook.org/

9. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation , 9(8), pp. 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

O. Velihurskyi

ANALYSIS OF MODERN ANOMALY DETECTION AND CORRECTION METHODS IN COMMUNICATION NETWORKS UNDER HIGH-LOAD CONDITIONS

This paper presents a comprehensive analysis of modern methods for anomaly detection and mitigation in communication networks, taking into account current challenges related to dynamic traffic scaling, high data transmission intensity, and the need to ensure stable network infrastructure performance in real-time conditions. The increasing intensity of data transmission in communication systems driven by the widespread use of streaming video, cloud computing, and mobile services leads to a higher frequency of anomalous events. This events can result in packet loss, increased latency, or complete service shutdown. This highlights the need for highly efficient detection systems capable of adapting to changes in the network environment. Traditional approaches often rely on fixed thresholds, signatures, or heuristic rules, which significantly limit their effectiveness under dynamic conditions. These systems cannot respond promptly to unknown threats or anomalous traffic patterns that were not anticipated during system configuration.

The study examines a broad range of contemporary methods, from classical statistical models to machine learning algorithms. Particular emphasis is placed on clustering techniques (e.g., k-means), Bayesian networks, tree-based models (Decision Tree, Random Forest), and time series approaches such as ARIMA and LSTM. The evaluation of these methods is based on their compliance with modern requirements, including scalability, real-time processing, integration with automated control systems, and minimization of false positives.

The analysis revealed that most existing solutions show limited generalization capabilities, high sensitivity to parameter tuning, and difficulties in adapting to variable load conditions. The novelty of this research lies in the comprehensive assessment for adopting hybrid systems that combine several diverse methods, including Autoencoders, GANs, VAEs, and Isolation Forest algorithms. This approach significantly enhances the adaptability and accuracy of anomaly detection even in complex and unpredictable environments. The results of this research are practically relevant for the further development of intelligent traffic monitoring systems that account for the described issues under critical load conditions.

Keywords: communication networks; anomalies; adaptivity; forecasting; scalability; network load; patters; machine learning; traffic analysis; hybrid models.