УДК 004.8:004.032.26: 621.317.08
DOI https://doi.org/10.32782/2663-5941/2025.6.2/53

Фуріхата Д.В.
Державний університет «Житомирська політехніка»

ГІБРИДНІ МЕТОДИ ВИЯВЛЕННЯ АНОМАЛІЙ У ДАНИХ РОЗУМНИХ ЛІЧИЛЬНИКІВ

У статті розглянуті гібридні методи виявлення аномалій у даних розумних лічильників, які набувають особливої актуальності в контексті розвитку цифрових технологій, інтелектуальних енергомереж, систем Інтернету речей та смарт-лічильників. Це пов'язане з постійним зростанням кількості даних, що передаються смарт-пристроями через технології IoT. Традиційні підходи до виявлення аномалій часто виявляються неефективними при обробці складних багатовимірних та динамічних даних, тоді як гібридні методи дозволяють поєднати переваги різних підходів для підвищення точності детекції та зменшення помилкових спрацьовувань.

Досліджено основні напрямки застосування гібридних архітектур, зокрема LSTM-автокодерів у різних модифікаціях. Розглянуто переваги федеративного навчання для розподілених систем, що забезпечує баланс між ефективністю виявлення аномалій та збереженням приватності даних споживачів. Проаналізовано ансамблеві методи, які комбінують множинні базові моделі для прийняття фінальних рішень та демонструють суттєве покращення продуктивності порівняно з окремими алгоритмами.

Особливу увагу приділено важливості врахування контекстуальних ознак та темпоральних характеристик, включаючи календарний контекст, сезонність та погодні умови. Розглянуто методи динамічного порогування на основі імовірнісних критеріїв, що дозволяють адаптувати системи до різних умов роботи. Проаналізовано як контрольовані підходи для виявлення крадіжок електроенергії з використанням згорткових нейронних мереж, так і неконтрольовані методи на основі автокодерів для детекції невідомих типів аномалій.

Визначено перспективні напрямки подальших досліджень, включаючи розробку більш ефективних методів обробки пропущених даних, покращення інтерпретованості моделей глибокого навчання та створення адаптивних архітектур для різних предметних областей.

Ключові слова: *виявлення аномалій, розумні лічильники, LSTM-автокодер, федеративне навчання, гібридні методи, інтелектуальні енергомережі, глибоке навчання.*

Постановка проблеми. У сучасному цифровому світі обсяги даних зростають експоненційно, що створює нові виклики для забезпечення безпеки інформаційних систем, якості продукції та надійності технологічних процесів. Виявлення аномалій відіграє критичну роль у широкому спектрі галузей – від кібербезпеки та фінансового моніторингу до промислової діагностики та аналітики даних розумних лічильників. Традиційні методи виявлення аномалій, засновані на статистичних підходах або класичних алгоритмах машинного навчання, часто виявляються недостатньо ефективними при роботі зі складними, багатовимірними та динамічними даними.

Водночас, окремі методи, зокрема глибоке навчання, висувають високі вимоги до обчислювальних ресурсів. З огляду на це, гібридні методи, що поєднують переваги різних підходів до виявлення аномалій, стають перспективним напрямком розвитку цієї галузі. Інтеграція статистичних методів з алгоритмами машинного навчання, комбінування класичних та глибоких нейронних мереж, а також поєднання різних методів дозволяє підвищити точність виявлення аномалій, знизити кількість помилок та адаптувати системи до специфічних особливостей предметної області. Це особливо актуально при передачі показів з розумних IoT пристроїв, що включають пропуски даних, нерегулярну передачу інформації, шуми та інші характерні проблеми.

Аналіз останніх досліджень і публікацій. Дослідження [6] представило систему розумного обліку, здатну виявляти аномалії за допомогою двонаправленого LSTM-Autoencoder (автокодер). Двонаправлена архітектура дозволяє моделі аналізувати часові послідовності як у прямому, так і у зворотному напрямку, що покращує здатність виявляти відхилення від нормальної поведінки.

Науковці [8] запропонували підхід виявлення аномалій з використанням LSTM-автокодера в контексті федеративного навчання. Це дозволяє навчати модель на розподілених даних з різних джерел, що особливо важливо для застосування у розумних електромережах. Федеративне навчання вирішує проблеми ізоляції даних та приватності, дозволяючи кожній підстанції локально тренувати моделі та обмінюватися лише агрегованими параметрами моделі.

Дослідники [10] представили метод неконтрольованого виявлення аномалій з використанням LSTM-автокодерів у поєднанні зі статистичною фільтрацією даних. За цим методом попередня обробка даних допомагає підвищити якість виявлення аномалій, відфільтровуючи шум та несуттєві відхилення перед подачею даних у нейронну мережу. Автори запропонували алгоритм навчання, що базується на імовірнісному критерії згідно з центральною граничною теоремою, який дозволяє оцінювати ймовірність того, що точка даних походить від нормального розподілу.

Вчені [2] розробили метод виявлення аномалій у даних розумних лічильників з використанням варіаційних рекурентних автокодерів з механізмом уваги, який дозволяє моделі фокусуватися на найбільш релевантних частинах часових послідовностей. Критичною інновацією їхнього підходу є попереднє виявлення пропущених значень та глобальних аномалій для зменшення їх внеску під час тренування моделі на забруднених даних.

Усі ці дослідження зосереджені на використанні автокодерів на базі LSTM-методів для виявлення аномалій у даних розумних лічильників, але з різними модифікаціями та акцентами – від двонаправленої обробки та федеративного навчання до статистичної фільтрації та механізмів уваги. Тому метою даної статті є узагальнення гібридних методів виявлення аномалій у розумних лічильниках.

Постановка завдання. Метою статті є детальний огляд гібридних методів виявлення аномалій у даних розумних лічильників.

Виклад основного матеріалу.

1. Основні підходи гібридних методів на основі LSTM-Autoencoder. Застосування LSTM-автокодера для виявлення аномалій у системах розумного обліку представляє один з напрямів сучасних досліджень. Науковці [9] запропонували фреймворк для ідентифікації аномалій у промислових даних, зібраних від віддалених терміналів на підстанціях інтелектуальних електромереж. Їхня система виявлення аномалій базується на архітектурі Long Short-Term Memory та автокодерів із застосуванням підходів Mean Standard Deviation та Median Absolute Deviation для детекції аномалій.

Двонаправлені LSTM архітектури демонструють переваги у контексті обробки часових послідовностей. Тому у [6] було використано двонаправлений LSTM-автокодер для пошуку аномальних точок даних, де модель обчислює помилку реконструкції через автокодер, навчений на неаномальних даних, а викиди класифікуються як аномалії за допомогою попередньо визначеного порогу. Двонаправлена архітектура дозволяє моделі враховувати як попередній, так і наступний контекст для кожної точки часового ряду, що критично важливо для виявлення контекстуальних аномалій.

Варіаційні автокодери представляють альтернативний підхід до моделювання невизначеності. Дослідники [2] розробили неконтрольований метод виявлення аномалій на основі варіаційного рекурентного автокодера з механізмом уваги. Автори провели кількісне порівняння з базовим підходом та чотирма іншими неконтрольованими методами, демонструючи ефективність запропонованого методу на реальному кейсі виявлення аномалій температури подачі води від промислової теплової установки.

2. Федеративне навчання для розподілених систем. Специфіка систем розумного обліку часто вимагає розподіленого підходу до обробки даних через питання приватності та безпеки. Автори [8] досліджували використання федеративного навчання з LSTM- автокодером для виявлення аномалій у системах інтелектуальних електромереж. Їхнє дослідження демонструє метод розробки неконтрольованої системи виявлення аномалій на основі федеративного навчання з використанням синтетичного набору даних, що базується на поведінці реальної системи. Для більш точної ідентифікації автори досліджували продуктивність інтеграції LSTM-автокодером з однокласовим методом опорних векторів та Isolation Forest, порівнюючи результати з пороговим підходом до виявлення аномалій.

Науковці [9] підкреслюють, що в традиційних системах велика кількість енергетичних даних від підстанцій має мігрувати до центрального сховища, що може призвести до зловживання даними, маніпуляцій або витоку приватної інформації. Федеративне навчання вирішує проблеми ізоляції даних та приватності даних, дозволяючи кожній підстанції локально тренувати моделі та обмінюватися лише агрегованими параметрами моделі.

3. Ансамблеві методи. Єдина модель може мати обмеження у виявленні різноманітних типів аномалій, тому ансамблеві підходи також розглядаються дослідниками. У [4] було запропоновано техніку пакетування ознак, яка розглядає лише підмножину ознак одночасно, та застосували трансформацію на основі вкладеної ротації, обчислену за допомогою аналізу головних компонент. Автори використали п'ять базових архітектур: автокодер, згортковий автокодер, LSTM, LSTM- автокодер та LSTM варіаційний автокодер у повністю неконтрольованому підході. Напівконтрольований підхід, що використовує логістичну регресію для комбінування передбачень множинних моделей, виявився найефективнішим варіантом, значно перевершивши всі базові методи та неконтрольовані ансамблі, що підтверджує доцільність використання ансамблевих технік для розв'язання задачі виявлення аномалій у багатовимірних часових рядах. Для подальшого покращення продуктивності прогнозування Iliopoulos та ін. [4] запропонували ансамблеву техніку, яка комбінує множинні базові моделі для прийняття фінального рішення. Додатково був запропонований напівконтрольований підхід з використанням логістичного регресора для комбінування виходів базових моделей. Запропонована методологія була застосована до набору даних Skoltech Anomaly Benchmark, який містить дані часових рядів, пов'язані з потоком води в замкненому контурі. Експериментальні результати показали, що запропонована ансамблева техніка перевершує базові алгоритми, демонструючи покращення продуктивності виявлення аномалій.

Інші науковці [5] підкреслюють важливість використання ансамблевих підходів для підвищення продуктивності системи. Автори запропонували ансамбль множинних CNN та LSTM моделей разом з ансамблем кількох трансформерних та графових нейронних мереж. Дослідження використовує машинне та глибоке навчання для підготовки набору даних, попередньої обробки та ідентифікації аномалій.

4. Контекстуальні ознаки та темпоральні характеристики. Врахування контекстуальної інформації є критично важливим для точного виявлення аномалій у системах розумного обліку. Дослідження показують, що інтеграція календарних та темпоральних ознак значно покращує здатність моделей розрізняти справжні аномалії від природних варіацій у споживанні. Календарний контекст включає день тижня, час доби, святкові дні та сезонність, що безпосередньо впливає на патерни енергоспоживання різних типів споживачів.

Багатовимірні підходи, які комбінують додаткові дані, такі як календарна та погодна інформація з енергоспоживанням, демонструють кращі результати порівняно з одновимірними методами, що фокусуються виключно на енергоспоживанні як вхідній змінній [5].

5. Динамічне порогування та адаптивні пороги. Встановлення оптимального порогу для класифікації точок даних як аномальних залишається однією з найбільш складних задач у неконтрольованому навчанні. У [6] використовували попередньо визначений поріг для відділення аномалій від неаномальних даних на основі помилки реконструкції. Однак статичні пороги мають фундаментальне обмеження через нездатність адаптуватися до зміни характеристик даних з часом.

Автори [10] запропонували імовірнісний критерій на основі центральної граничної теореми, який дозволяє автоматичне маркування даних. Цей підхід дозволяє системі динамічно оцінювати, чи походить точка даних від нормального розподілу, що забезпечує більш гнучкий механізм прийняття рішень порівняно зі статичними порогами. Автори підкреслюють, що їхній алгоритм може бути налаштований для балансування між рівнем помилкових спрацьовувань та точністю, що є критично важливим для практичного застосування в промислових системах.

6. Методи навчання з вчителем. Дослідження, присвячене розробці ефективної системи виявлення крадіжок електроенергії, демонструє використання комбінації згорткових нейронних мереж та моделей довготривалої короткочасної пам'яті для аналізу даних споживання електроенергії та виявлення аномальних патернів, що можуть свідчити про шахрайство. З появою розумних лічильників у інтелектуальних мережах з'явилася можливість збирати великі обсяги даних про споживання, що можна використовувати для автоматичного виявлення шахрайства за допомогою алгоритмів машинного навчання [3].

Для подолання проблем незбалансованості класів застосовується техніка аугментації даних LoRAS, оскільки випадків чесного споживання значно більше, ніж випадків крадіжок. Запропонована модель на основі LSTM досягає точності понад 90% на реальних даних Державної енергетичної корпорації Китаю [12].

Гібридна система виявлення крадіжок електроенергії на основі згорткових нейронних мереж у розумних електромережах використовує патерни споживання. Інноваційний підхід поєднує CNN для автоматичного вилучення ознак з традицій-

ними алгоритмами машинного навчання для класифікації користувачів як чесних або шахрайських. Для вирішення проблеми незбалансованості класів використовується техніка генеративно-змагальних мереж. Найкращі результати досягнуто моделлю CNN з логістичною регресією [12].

Комплексний фреймворк для виявлення аномалій у енергосистемах поєднує дані з кіберпростору та фізичної інфраструктури. Система використовує мережі довготривалої короткочасної пам'яті та класифікатори Random Forest для високоточного виявлення кіберзагроз у розумних енергомережах. Модель LSTM показала високу продуктивність завдяки здатності відстежувати часові залежності в послідовних системних даних [11].

Гібридна система виявлення крадіжок електроенергії в розумній електромережі на основі глибокого навчання використовує дані споживання. Методика поєднує метод опорних векторів та алгоритм оптимізації рою частинок для точного виявлення шахрайських споживачів у мережі з урахуванням раптових змін споживання. Метод опорних векторів виявляє невідповідні кореляції раптових змін споживання, тоді як алгоритм оптимізації рою частинок оптимізує відповідні параметри моделі. Розроблена система також включає календарний контекст для планування споживання енергії та виявлення крадіжок [3].

7. Методи навчання без вчителя. Система виявлення крадіжок електроенергії в розумних мережах з використанням федеративного навчання та глибоких нейронних мереж пропонує інноваційний підхід, який використовує детектори аномалій на основі автокодерів замість традиційних контрольованих методів машинного навчання. Ключова перевага такого підходу полягає в тому, що система навчається тільки на нормальних даних споживання електроенергії і може виявляти будь-які відхилення від звичайного патерну, включаючи нові, раніше невідомі типи атак [1]. Важливою особливістю запропонованої системи є використання федеративного навчання, яке дозволяє захистити приватність споживачів. Замість передачі особистих даних про споживання електроенергії компанії, споживачі навчають локальні моделі на своїх пристроях і передають лише параметри навчених моделей на сервер агрегації. Результати показали, що федеративний підхід досягає порівнянної ефективності з централізованим навчанням, але при цьому зберігає конфіденційність даних споживачів [1].

8. Інтеграційні архітектури та практичні аспекти. Деякі науковці [8] підкреслюють важливість розробки неконтрольованої системи з використанням синтетичного набору даних, що базується на поведінці реальної системи. Такий підхід дозволяє тестувати різні компоненти системи в контрольованих умовах перед розгортанням на реальних даних.

Комбінація LSTM- автокодерів з традиційними методами машинного навчання демонструє синергетичний ефект. Інтеграція LSTM- автокодерів з однокласовим методом опорних векторів та Isolation Forest дозволяє використовувати переваги кожного підходу. LSTM-автокодери ефективно вивчають темпоральні залежності та створюють compressed representations, тоді як OC-SVM та Isolation Forest забезпечують додаткові механізми детекції на основі цих представлень.

Висновки. Гібридні методи виявлення аномалій демонструють значні переваги порівняно з окремими підходами, поєднуючи сильні сторони різних архітектур та алгоритмів. Федеративне навчання забезпечує баланс між ефективністю виявлення та збереженням приватності даних. Ансамблеві підходи підвищують точність та знижують рівень помилкових спрацьовувань. Інтеграція контекстуальних ознак та динамічного порогування дозволяє адаптувати системи до специфічних особливостей предметної області та змінних умов роботи.

Перспективи подальших досліджень включають розробку більш ефективних методів обробки пропущених даних, покращення інтерпретованості моделей глибокого навчання, створення адаптивних архітектур, що автоматично налаштовуються до характеристик конкретного набору даних, та розширення застосування гібридних методів на інші галузі, де виявлення аномалій є критично важливим.

Список літератури:

1. Ali Alshehri, Mahmoud M. Badr, Mohamed Baza and Hani Alshahrani. Deep Anomaly Detection Framework Utilizing Federated Learning for Electricity Theft Zero-Day Cyberattacks. *Sensors*. 2024. 24 (10). 3236. https://doi.org/10.3390/s24103236

2. Dai W., Liu X., Heller A., Nielsen P. S. Smart Meter Data Anomaly Detection Using Variational Recurrent Autoencoders with Attention. In F. Sanfilippo, O.-C. Granmo, S. Y. Yayilgan, & I. S. Bajwa (Eds.), *Intelligent Technologies and Applications - 4th International Conference, INTAP 2021*, Revised Selected Papers. Springer Science and Business Media Deutschland GmbH. 2022. Pp. 311–324. https://doi.org/10.1007/978-3-031- 10525-8_25.

3. Ida Evangeline S., Darwin S., Peter Anandkumar P., M. Chithambara Thanu. Anomaly detection in smart grid using a trace-based graph deep learning model. *Electrical Engineering.* 2024. 106. Pp. 5851–5867. https://doi.org/10.1007/s00202-024-02327-6

4. Iliopoulos A., Seferis G., Diou C. Varlamis I. Detection of Anomalies in Multivariate Time Series Using Ensemble Techniques. *2023 IEEE Ninth International Conference on Big Data Computing Service and Applications (BigDataService),* Athens, Greece. 2023. Pp. 1–8. https://doi.org/10.1109/BigDataService58306.2023.00007.

5. Iqbal A., Amin R., Alsubaei F.S., Alzahrani A. Anomaly detection in multivariate time series data using deep ensemble models. *PLOS ONE.* 2024. 19(6). Pp. e0303890. https://doi.org/10.1371/journal.pone.0303890

6. Lee S., Jin H., Vecchietti L. F., Hong J., Park D., Kim J. Smart Metering System Capable of Anomaly Detection by Bi-directional LSTM Autoencoder. *2022 IEEE International Conference on Consumer Electronics* (ICCE), Las Vegas, NV, USA. 2022. Pp. 1–6, https://doi.org/10.1109/ICCE53296.2022.9730398.

7. Liu Y., Zhu L., Ding L., Huang Z., Sui H., Wang S., Song Yu. Selective ensemble method for anomaly detection based on parallel learning. *Scientific Reports*. 2024. Vol. 14, 1420. https://doi.org/10.1038/s41598-024-51849-3

8. Mohammadreza Mohammadi, Rakesh Shrestha, Sima Sinaei, Alberto Salcines, David Pampliega, Raul Clemente, and Ana Lourdes Sanz. Anomaly Detection Using LSTM-Autoencoder in Smart Grid: A Federated Learning Approach. In *Proceedings of the 2023 7th International Conference on Cloud and Big Data Computing (ICCBDC '23).* Association for Computing Machinery, New York, NY, USA, 2023. Pp. 48–54. https://doi.org/10.1145/3616131.3616138.

9. Rakesh Shrestha, Mohammadreza Mohammadi, Sima Sinaei, Alberto Salcines, David Pampliega, Raul Clemente, Ana Lourdes Sanz, Ehsan Nowroozi, Anders Lindgren, Anomaly detection based on LSTM and autoencoders using federated learning in smart electric grid, *Journal of Parallel and Distributed Computing.* 2024. Vol. 193. 104951. https://doi.org/10.1016/j.jpdc.2024.104951.

10. Sepehr Maleki, Sasan Maleki, Nicholas R. Jennings. Unsupervised anomaly detection with LSTM autoencoders using statistical data-filtering. *Applied Soft Computing.* 2021. Vol. 108. 107443. https://doi.org/10.1016/j.asoc.2021.107443.

11. Wu Y., Zang Z., Zou X., Luo W., Bai N., Xiang Yi, Li W., Dong W. Graph attention and Kolmogorov–Arnold network based smart grids intrusion detection. *Scientific Reports*. 2025. Vol. 15. 88054. https://doi.org/10.1038/s41598-025-88054-9

12. Zhuang W., Jiang W., Xia M., Liu J. Dynamic Generative Residual Graph Convolutional Neural Networks for Electricity Theft Detection. *IEEE Access*. 2024. Vol. 12. Pp. 42737–42750. https://doi.org/10.1109/ACCESS.2024.3379201.

Furikata D.V. HYBRID METHODS FOR DETECTING ANOMALIES IN SMART METER DATA

The article discusses hybrid methods for detecting anomalies in smart meter data, which are becoming particularly relevant in the context of the development of digital technologies, smart grids, Internet of Things systems, and smart meters. This is due to the constant growth in the amount of data transmitted by smart devices via IoT technologies. Traditional approaches to anomaly detection are often ineffective when processing complex multidimensional and dynamic data, while hybrid methods allow combining the advantages of different approaches to improve detection accuracy and reduce false positives. The main areas of application of hybrid architectures, in particular LSTM autoencoders in various modifications. The advantages of federated learning for distributed systems are considered, which provides a balance between the effectiveness of anomaly detection and the preservation of consumer data privacy. Ensemble methods that combine multiple base models for final decision-making and demonstrate significant performance improvements over individual algorithms are analyzed. Particular attention is paid to the importance of considering contextual features and temporal characteristics, including calendar context, seasonality, and weather conditions. Dynamic thresholding methods based on probabilistic criteria are considered, allowing systems to adapt to changing operating conditions. Both supervised approaches for detecting electricity theft using convolutional neural networks and unsupervised methods based on autoencoders for detecting unknown types of anomalies are analyzed. Promising areas for further research are identified, including the development of more efficient methods for processing missing data, improving the interpretability of deep learning models, and creating adaptive architectures for different subject areas.

Key words: *anomaly detection, smart meters, LSTM autoencoder, federated learning, hybrid methods, smart grids, deep learning.*

Дата надходження статті: 20.11.2025
Дата прийняття статті: 05.12.2025
Опубліковано: 30.12.2025