УДК 004.056.53:004.75:004.032.26 **DOI:** 10.31673/2412-9070.2026.017404
І. А. БУЧЕНКО[1], ст. викл.;
ORCID: 0009-0003-9012-9632
А. В. ЛЕМЕШКО[2], PhD, доцент;
ORCID: 0000-0001-8003-3168
Н. О. ЛАЩЕВСЬКА[1], канд. техн. наук, доцент,
ORCID: 0000-0003-2148-115X
[1]Державний університет інформаційно-комунікаційних технологій, Київ
[2]Державний торговельно-економічний університет, Київ

АДАПТИВНИЙ ПІДХІД ДО ОЦІНЮВАННЯ РИЗИКІВ КІБЕРБЕЗПЕКИ В РОЗПОДІЛЕНИХ ІНФОРМАЦІЙНИХ СИСТЕМАХ НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ

У статті досліджено актуальні проблеми оцінювання ризиків кібербезпеки в умовах функціонування сучасних розподілених інформаційних систем (РІС). Проаналізовано ключові обмеження традиційних підходів, які базуються на експертних, імовірнісних чи статичних методах. Встановлено, що такі методи не забезпечують достатньої адаптивності та точності в динамічному середовищі, що характеризується високою складністю, гетерогенністю джерел даних та децентралізацією інфраструктури. Сучасні РІС, що обробляють великі обсяги потокових даних у режимі реального часу (логи, телеметрія, транзакції), вимагають негайного реагування на загрози, що робить класичні методи малоефективними. Додатковими викликами є гетерогенність даних, відсутність уніфікованих форматів та складність інтеграції з інструментами інвентаризації ІТ-активів.

Обґрунтовано необхідність розробки нового адаптивного підходу до оцінювання кіберризиків, що базується на інтелектуальному аналізі даних, кореляційному моделюванні та використанні глибоких нейронних мереж. Метою дослідження є підвищення ефективності цього процесу шляхом розробки адаптивного методу на основі нейромережевого аналізу. Для досягнення мети було розроблено профіль ключових факторів ризику (КФР) та відповідних контролів безпеки. Цей профіль включає такі динамічні контрольні ознаки, як середній час відповіді, частота аномалій, інтенсивність трафіку та топологічні зміни в мережі.

У роботі описано практичну реалізацію методу, що включає етапи уніфікації, часової синхронізації та агрегації гетерогенних потоків даних. На основі даних ІТ-інфраструктури сформовано навчальні вибірки, на яких побудовано та протестовано комплекс нейромережевих моделей (зокрема, рекурентних мереж та автокодерів) для оцінювання рівня ризику. Особливу увагу приділено вирішенню проблеми "concept drift" — зміни статистичних характеристик даних — шляхом впровадження механізмів онлайн-навчання (online learning) та ковзних вікон. Запропонований підхід дозволяє моделям постійно оновлювати параметри без повного перенавчання, реагуючи на нові типи загроз.

Наукова новизна полягає в розробці комплексу нейромережевих моделей, що синтезує метрико-орієнтований та стандарт-орієнтований підходи, та вдосконаленні методу формування профілю КФР. Практична значущість полягає в тому, що запропонований підхід забезпечує більш точну, масштабовану та автоматизовану оцінку кіберризиків. Це дозволяє перейти від реактивного до проактивного управління безпекою, прогнозуючи збої та виявляючи аномалії до їхньої реалізації.

Ключові слова: комп'ютерна мережа; машинне навчання; обробка потокових даних; виявлення аномалій; concept drift; відмовостійкість; інтелектуальний аналіз даних; інформаційна безпека; кореляційне моделювання.

Постановка проблеми

Стрімка цифрова трансформація призводить до експоненційного зростання обсягів даних, що генеруються та обробляються у сучасних інформаційних системах. Особливого значення набувають потокові дані – безперервні, постійно оновлювані потоки інформації, що надходять з великою швидкістю з численних джерел, таких як журнали серверів, телеметрія, дані IoT-сенсорів, фінансові транзакції та активність користувачів. Обробка таких даних у режимі реального часу є критично важливою для підтримки безперервної роботи, а також для оперативного виявлення аномалій та забезпечення кібербезпеки.

Ці процеси відбуваються в середовищі розподілених інформаційних систем (РІС), які характеризуються використанням мікросервісних архітектур, контейнеризації та географічної розподіленості вузлів. Така архітектура забезпечує гнучкість і масштабованість, однак водночас створює безпрецедентні виклики для систем безпеки.

Проблематика дослідження полягає в тому, що існуючі підходи до оцінювання ризиків кібербезпеки, які базуються на експертних, імовірнісних чи статичних методах, є малоефективними для складних, динамічних і масштабованих РІС. Вони або не враховують реальний стан інфраструктури та динаміку потокових даних, або потребують надмірних обчислювальних ресурсів, або базуються на суб'єктивних оцінках, що призводить до низької точності прогнозування та запізнілого реагування.

Ситуація додатково ускладнюється гетерогенністю даних, відсутністю уніфікованих форматів та слабкою інтеграцією засобів безпеки з інструментами інвентаризації IT-активів. Крім того, самі аналітичні моделі швидко застарівають через явище concept drift (зміна статистичних характеристик потоку), що призводить до деградації точності. Таким чином, виникає гостра науково-практична проблема розробки нового, адаптивного підходу до оцінювання кіберризиків, здатного функціонувати в режимі реального часу в умовах гетерогенного, динамічного середовища РІС.

Аналіз останніх досліджень і публікацій

Аналіз фахової літератури та технологічних рішень свідчить про активний пошук інструментів для роботи з потоковими даними та забезпечення безпеки в РІС. Сучасні інфраструктури обробки потоків переважно будуються на платформах Apache Kafka, Apache Flink та Apache Spark Streaming. Kafka зарекомендувала себе як надійна publish-subscribe платформа з високою пропускною здатністю, однак вона має обмеження в сценаріях зі складною обчислювальною логікою. Flink надає потужні засоби обробки зі збереженням стану (stateful processing) та підтримкою "event time", але вимагає значних ресурсів та має високий поріг входу. Spark Streaming, працюючи за принципом "міні-пакетів", добре інтегрується з екосистемою Spark, але менш придатний для завдань із жорсткими вимогами реального часу. Для оркестрації цих сервісів стандартом де-факто стало використання Kubernetes та Docker, що забезпечує масштабування та відмовостійкість.

Водночас науковці в галузі інформаційної безпеки [1] та інтелектуального аналізу даних [2] вказують на обмеженість класичних сигнатурних методів. У відповідь активно розвиваються адаптивні методи машинного навчання [4]. Дослідження останніх років зосереджені на застосуванні глибокого навчання для виявлення вторгнень та аналізу ризиків. Зокрема, у роботі [6] автори пропонують гібридну модель, що поєднує згорткові нейронні мережі (CNN) для просторової екстракції ознак з мережевих пакетів та рекурентні мережі (LSTM) для аналізу часових залежностей. Такий підхід показує високу точність у виявленні складних, багатоетапних атак.

Ключовою невирішеною проблемою залишається concept drift – неминуча зміна статистичних характеристик потоку даних з часом, що призводить до деградації статично навчених моделей. Для вирішення цієї проблеми у праці [7] пропонується огляд та аналіз методів детекції дрейфу, таких як DDM (Drift Detection Method) та ADWIN (Adaptive Windowing), які дозволяють системі виявити момент, коли модель починає помилятися, та ініціювати її перенавчання.

Окремим напрямом, що набуває актуальності для PIC, є федеративне навчання (Federated Learning). Дослідження [8] демонструє підхід, за якого нейромережеві моделі навчаються на локальних вузлах (наприклад, у різних філіях компанії або на edge-пристроях), не передаючи сирі дані до центрального сервера. Обмінюючись лише вагами (параметрами) моделі, вдається побудувати точну глобальну модель оцінки ризиків, одночасно забезпечуючи високий рівень конфіденційності даних.

Проте, невирішеною частиною загальної проблеми залишається відсутність комплексного методу, який би поєднував інфраструктурні можливості потокових платформ (Kafka, Kubernetes) з адаптивними нейромережевими моделями (LSTM, Autoencoders) та сучасними механізмами протидії "concept drift" саме для кількісного оцінювання ризиків кібербезпеки в PIC. Більшість існуючих рішень або фокусуються на інфраструктурі, або пропонують статичні моделі аналізу. Відсутній цілісний підхід, що синтезує метрико-орієнтований та стандарт-орієнтований аналіз [3] і дозволяє системі адаптуватися до нових, раніше невідомих типів загроз в режимі реального часу.

Формулювання мети статті

Метою написання цієї статті є підвищення ефективності процесу оцінювання ризиків кібербезпеки в умовах динамічного середовища сучасних масштабованих розподілених інформаційних систем шляхом розробки адаптивного методу на основі нейромережевого аналізу гетерогенних даних.

Виклад основного матеріалу

В ході дослідження було розроблено комплексний адаптивний метод оцінювання ризиків, що інтегрує аналіз потокових даних та нейромережеве моделювання.

Архітектура адаптивної системи оцінювання ризиків. Запропонована архітектура є багаторівневою та призначена для функціонування поверх існуючих інфраструктурних рішень, таких як Kubernetes та Apache Kafka. Вона описує логічний потік даних через чотири основні рівні:

1. рівень збору даних (Data Collection Layer) відповідає за підключення до різнорідних джерел потокових даних (логи мікросервісів, метрики мережевого трафіку, дані телеметрії, API-запити) та їх передачу до брокера повідомлень (наприклад, Apache Kafka) для забезпечення надійної та асинхронної доставки;
2. рівень попередньої обробки та агрегації (Pre-processing & Aggregation Layer) споживає дані з брокера. Його ключові завдання – уніфікація даних (приведення до єдиного формату JSON або Avro) та часова синхронізація (вирішення проблем запізнілих подій). Тут відбувається агрегація сирих подій у логічні структури (часові вікна, транзакційні сесії), що формують «контекст події»;
3. рівень адаптивного моделювання (Adaptive Modeling Layer) – ядро системи. Цей рівень отримує агреговані дані та застосовує до них комплекс нейромережевих моделей для розрахунку динамічного профілю ризику. Тут же реалізовані механізми онлайн-навчання та детекції "concept drift";
4. рівень прийняття рішень (Decision & Action Layer) отримує кількісну оцінку ризику. На основі заданих політик, цей рівень може ініціювати автоматичні дії: відправка сповіщення (alert) до SIEM-системи, ініціювання блокування IP-адреси, або передача команди системі оркестрації (Kubernetes) для ізоляції підозрілого контейнера.

Формування та агрегація профілю ризику. На відміну від статичних оцінок, запропонований профіль ризику є динамічним вектором ознак, що оновлюється в реальному часі. Він формується на рівні агрегації та включає контрольні ознаки (KRFs), що характеризують поточний стан системи. До них належать метрики продуктивності, такі як середній час відповіді сервісів (service response time) та відсоток помилок (error rate), оскільки аномальне зростання цих показників може свідчити про DoS-атаку або збій компонента. Також враховуються метрики поведінки, такі як частота аномальних запитів, що може вказувати на SQL-ін'єкції або кількість повторюваних спроб доступу (brute-force), співвідношення трафіку (upload/download). Важливу роль відіграють і метрики топології, адже раптова поява нового, невідомого вузла або зміна звичних маршрутів трафіку є суттєвим фактором ризику, що може вказувати на атаку

"людина посередині" (Man-in-the-Middle) або несанкціоноване підключення. Крім того, звертають увагу і на метрики інтенсивності, які характеризуються загальним обсягом трафіку та кількістю одночасних сесій. Модель при цьому реагує не на абсолютні значення, а на динаміку змін цих параметрів, розраховуючи відхилення від "нормальної" базової лінії (baseline).

Комплекс нейромережевих моделей для оцінювання. Для аналізу цього багатовимірного профілю ризику використовується комбінація двох типів нейронних мереж, що виконують різні завдання: рекурентні нейронні мережі (LSTM/GRU) та автокодери (Autoencoders). Оскільки дані профілю ризику є часовими послідовностями (time series), застосування стандартних мереж прямого поширення є неефективним, тому для їх аналізу застосовуються рекурентні нейронні мережі (LSTM/GRU). Моделі з довгою короткочасною пам'яттю (LSTM) або керованими рекурентними блоками (GRU) здатні аналізувати послідовності подій у часі та виявляти складні патерни атак, які складаються з кількох, на перший погляд не пов'язаних, кроків. Складний патерн атаки виглядає так: [низька активність] -> [серія невдалих логінів] -> [успішний логін з нетипової IP-адреси] -> [аномальне зростання вихідного трафіку]. Для статичної моделі кожна подія окремо може не бути ризиком, але їх послідовність є чіткою ознакою вторгнення.

Паралельно для виявлення аномалій (anomaly detection), особливо нових, раніше невідомих загроз (zero-day attacks), використовуються автокодери (Autoencoders). Цей тип мереж навчається на виключно легітимних, "нормальних" даних з метою ефективного стискання вхідного вектора ознак у прихований простір (encoding) і потім максимально точно відновлювати його (decoding). Коли навчена модель отримує на вхід аномальний вектор даних, який вона ніколи не бачила, вона не може його коректно відновити. Це призводить до високої помилки реконструкції (reconstruction error), що є прямим числовим сигналом про аномалію та підвищення рівня ризику.

Механізм адаптації до "Concept Drift" (рис. 1). Ключовою особливістю методу є його адаптивність, оскільки класичні статичні моделі в умовах РІС швидко деградують через постійні зміни (оновлення сервісів, зміна навантаження) "нормальної" поведінки системи. Запропонований метод використовує гібридний підхід. По-перше, це онлайн-навчання (Online Learning), де модель використовує механізм поступового донавчання (інкрементного навчання) на невеликих блоках (mini-batches) нових даних, оновлюючи свої ваги без повної зупинки та перенавчання. По-друге, паралельно працює детектор дрейфу (Drift Detection) (наприклад, ADWIN [7]), який моніторить статистичний розподіл помилок моделі. Якщо цей розподіл починає суттєво змінюватися, це свідчить про радикальну зміну в поведінці системи (concept drift) і сигналізує про необхідність повного перенавчання моделі на новому, свіжому наборі даних.

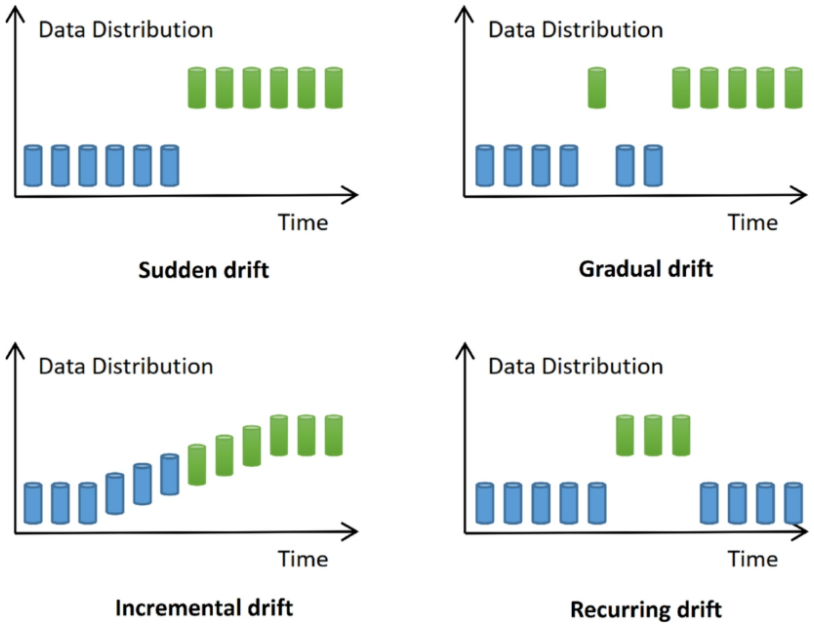

Рис. 1. Типи дрейфу концепцій з використанням методів паралельного виявлення та шумозаглушеннях [9]

Інтеграція з системами оркестрації та відмостійкості. Розроблений метод не лише оцінює ризики, але й слугує основою для проактивного забезпечення безперебійної роботи. Результатом роботи моделі є кількісна оцінка ризику R (1) для кожного сервісу або вузла. Ця оцінка інтегрується з системами оркестрації (Kubernetes) через його API.

$$R \in [0,1] \quad (1)$$

Ця оцінка R знаходиться в діапазоні [0,1] (1), що є фундаментальною вимогою для інтерпретації та автоматизації. Такий діапазон використовується тому, що він по суті являє собою нормалізовану шкалу або ймовірність. Значення «0» позначає повну відсутність виявленого ризику (ідеальний, "нормальний" стан), тоді як «1» позначає максимальний виявлений ризик або стовідсоткову впевненість моделі в тому, що відбувається атака чи критичний збій. Багато нейромережевих моделей, особливо ті, що вирішують задачі класифікації, використовують на вихідному шарі сигмоїдну функцію активації, яка природним чином відображає будь-яке вхідне значення у діапазон [0,1], що ідеально підходить для інтерпретації як ймовірності ризику. Ця нормалізація є критично важливою, оскільки вона дозволяє встановлювати чіткі, зрозумілі та універсальні пороги для реагування що було б неможливо з використанням необробленої, необмеженої шкали помилок.

Наприклад, якщо при отриманні сигналу про високий ризик $R > 0{,}85$ для певного мікросервісу, система може автоматично ініціювати примусову перевірку працездатності health-check для даного сервісу, застосувати тимчасове правило NetworkPolicy для ізоляції підозрілого сервісу (pod) від решти інфраструктури, запобігаючи поширенню атаки, або запустити процедуру поступового перезапуску rolling restart, переходячи від реактивного моніторингу до проактивного самовідновлення.

Висновки

У ході проведеного дослідження було здійснено комплексний аналіз проблем обробки потокових даних у РІС та розроблено адаптивний метод оцінювання ризиків кібербезпеки на основі нейромережевого аналізу.

Доведено, що традиційні статичні та експертні методи оцінювання ризиків є малоефективними для сучасних РІС через високу динаміку, гетерогенність даних та неминуче явище concept drift.

Запропоновано та обґрунтовано комплексний адаптивний метод, що включає чотирирівневу архітектуру, динамічне формування профілю ризику (KRFs) та використання комбінованого нейромережевого підходу. Застосування рекурентних мереж (LSTM) дозволяє аналізувати часові послідовності загроз, тоді як автокодери забезпечують ефективне виявлення раніше невідомих аномалій (zero-day).

Наукова новизна роботи полягає в синтезі цих моделей з механізмами онлайн-навчання та детекції дрейфу (ADWIN/DDM), що дозволяє системі не лише навчатися, але й адаптуватися до змін "нормальної" поведінки інфраструктури, підтримуючи високу точність у довгостроковій перспективі.

Практична значущість полягає в тому, що розроблений метод дозволяє перейти від реактивного до проактивного управління безпекою. Його інтеграція з системами оркестрації (Kubernetes) надає механізм для автоматичного реагування на загрози (ізоляція сервісів, перезапуск), що критично важливо для забезпечення відмовостійкості в сферах кібербезпеки, фінансових транзакцій, IoT та моніторингу критичної інфраструктури.

Перспективи подальших досліджень полягають у розширенні сфери застосування розроблених методів. Зокрема, це стосується побудови повністю автономних самонавчальних систем безпеки (Self-Learning Security Systems), управління інформаційними потоками у кіберфізичних системах та IoT-інфраструктурах, інтелектуальних транспортних системах та енергетичних мережах. Подальші дослідження доцільно спрямувати на поєднання адаптивного потокового аналізу із методами розподіленого прийняття рішень, технологіями блокчейн (для забезпечення прозорості аналізу), а також дослідити застосування федеративного навчання (Federa-

ted Learning) [8] для побудови глобальних моделей ризику без необхідності централізації та передачі конфіденційних даних з розподілених вузлів.

Список літератури

1. Інформаційна безпека України [Електронний ресурс] : наук. журнал / Ін-т інформації, безпеки і права НАПрН України. – 2024. – № 3. – Режим доступу: https://ippi.org.ua/sites/default/files/2024-3.pdf. – (дата звернення: 02.11.2025).

2. Бобровник С. В. Інтелектуальний аналіз даних у кібербезпеці [Електронний ресурс] / С. В. Бобровник, О. М. Романенко. – Режим доступу: https://www.academia.edu/44557470/. – (дата звернення: 02.11.2025).

3. Штучний інтелект і аналітика даних у НУЛП [Електронний ресурс] / Львівський IT Кластер. – Режим доступу: https://itcluster.lviv.ua/ai_nulp/. – (дата звернення: 02.11.2025).

4. Адаптивні системи автоматичного управління [Електронний ресурс] : темат. вип.: інформаційна безпека та машинне навчання. – Режим доступу: https://asac.kpi.ua/. – (дата звернення: 02.11.2025).

5. Kubernetes. Production-Grade Container Orchestration [Electronic resource]. – Mode of access: https://kubernetes.io/. – (date of access: 02.11.2025).

6. Al-Qatf, M. A Hybrid Deep Learning Model (CNN-LSTM) for Real-Time Intrusion Detection in IoT Networks [Electronic resource] / M. Al-Qatf, Y. Lasheng, M. Al-Habib // IEEE Access. – 2023. – Vol. 11. – P. 101340-101351. – Mode of access: https://ieeexplore.ieee.org/document/10246835. – (date of access: 02.11.2025).

7. Khamassi, I. Addressing Concept Drift in Cybersecurity: A Review of Adaptive Learning Techniques [Electronic resource] / I. Khamassi, S. Gherissi, M. Hamdi // Journal of Network and Computer Applications. – 2024. – Vol. 227. – Art. 103932. – Mode of access: https://www.sciencedirect.com/science/article/pii/S108480452300486X. – (date of access: 02.11.2025).

8. Nguyen, T. D. Federated Learning for Cybersecurity Risk Assessment in Distributed Systems [Electronic resource] / T. D. Nguyen, Q. D. Tran, S. R. Maskey // IEEE Transactions on Information Forensics and Security. – 2022. – Vol. 17. – P. 3178-3191. – Mode of access: https://ieeexplore.ieee.org/document/9863417. – (date of access: 02.11.2025).

9. *An enhanced concept drift detection and adaptation framework using optimized deep learning models for streaming data [Electronic resource] / Mahdi Al-Sarem, Larbi-Listan L. M. K. O. Ang, Mohammed F. A. Mohammed [ma ін.] // Applied Intelligence. – 2024. – Vol. 54. – P. 15004–15024. – Mode of access: https://link.springer.com/article/10.1007/s10489-024-05988-9. – (date of access: 02.11.2025).*

I. Buchenko, A. Lemeshko, N. Lashchevska

AN ADAPTIVE APPROACH TO CYBERSECURITY RISK ASSESSMENT IN DISTRIBUTED INFORMATION SYSTEMS BASED ON NEURAL NETWORKS

This article investigates the current challenges of cybersecurity risk assessment within modern distributed information systems (DIS). It analyzes the key limitations of traditional approaches based on expert, probabilistic, or static methods. It is established that such methods lack sufficient adaptability and accuracy in a dynamic environment characterized by high complexity, heterogeneity of data sources, and infrastructure decentralization. Modern DIS, which process large volumes of real-time streaming data (logs, telemetry, transactions), require immediate responses to threats, rendering classic methods ineffective. Additional challenges include data heterogeneity, the absence of unified formats, and the complexity of integration with IT asset inventory tools.

The necessity of developing a new adaptive approach to cyber risk assessment is substantiated, based on intelligent data analysis, correlation modeling, and the use of deep neural networks. The aim of the research is to increase the efficiency of this process by developing an adaptive method based on neural network analysis. To achieve this aim, a profile of key risk factors (KRFs) and corres-

ponding security controls was developed. This profile includes dynamic control features such as average response time, anomaly frequency, traffic intensity, and topological network changes.

The paper describes the practical implementation of the method, which includes stages of unification, time synchronization, and aggregation of heterogeneous data streams. Based on IT infrastructure data, training datasets were formed, upon which a complex of neural network models (including recurrent neural networks and autoencoders) was built and tested for risk level evaluation. Special attention is given to solving the "concept drift" problem—the change in data statistical characterristics—by implementing online learning mechanisms and sliding windows. The proposed approach allows models to continuously update their parameters without full retraining, reacting to new types of threats.

The scientific novelty lies in the development of a complex of neural network models that synthesizes metric-oriented and standard-oriented approaches, and in the improvement of the KRF profile formation method. The practical significance is that the proposed approach provides a more accurate, scalable, and automated assessment of cyber risks. This enables a shift from reactive to proactive security management by predicting failures and detecting anomalies before they occur.

Keywords: computer network; machine learning; stream data processing; anomaly detection; concept drift; fault tolerance; intelligent data analysis; information security; correlation modeling.

Надійшла до редакції: 21.10.2025
Прийнята до друку: 10.12.2025
Опубліковано: 27.02.2026