УДК 004.056:004.75
DOI https://doi.org/10.32782/IT/2025-1-9

Сергій ДОРОЖИНСЬКИЙ
PhD, доцент кафедри кібербезпеки, інформаційних технологій та економіки, декан факультету інформаційних технологій та соціально-гуманітарних наук, Київський університет інтелектуальної власності та права Національного університету «Одеська юридична академія», Харківське шосе, 210, м. Київ, Україна, 02121
***ORCID:** 0000-0002-5395-6423*
Scopus author ID: 57222147991

Віталій ТУПКАЛО
доктор технічних наук, професор, професор кафедри кібербезпеки, інформаційних технологій та економіки, Київський університет інтелектуальної власності та права Національного університету «Одеська юридична академія», Харківське шосе, 210, м. Київ, Україна, 02121
***ORCID:** 0000-0002-6594-530X*

Юрій ЩЕРБИНА
кандидат технічних наук, доцент кафедри інформаційних технологій, Національний університет «Одеська юридична академія», Фонтанська дорога, 23, м. Одеса, Україна, 65009
***ORCID:** 0000-0003-3885-6747*

Бібліографічний опис статті: Дорожинський, С., Тупкало, В., Щербина, Ю. (2025). Критичні аспекти обробки великих даних для забезпечення їх захисту в розподілених мережах. *Information Technology: Computer Science, Software Engineering and Cyber Security*, 60–67, doi: https://doi.org/10.32782/IT/2025-1-9

КРИТИЧНІ АСПЕКТИ ОБРОБКИ ВЕЛИКИХ ДАНИХ ДЛЯ ЗАБЕЗПЕЧЕННЯ ЇХ ЗАХИСТУ В РОЗПОДІЛЕНИХ МЕРЕЖАХ

У статті детально розглянуто критичні аспекти обробки великих даних у розподілених мережах, які становлять сучасний виклик для забезпечення кібербезпеки.

***Метою даної роботи** є проведення глибокого аналізу теоретичних і практичних підходів до вирішення проблем конфіденційності, цілісності та доступності даних у багатокористувацьких середовищах із високою складністю взаємодії. Особлива увага приділена класифікації загроз, що виникають під час обробки великих обсягів інформації, із врахуванням специфіки архітектур розподілених систем, таких як хмарні технології, однорангові мережі та інфраструктури з високим рівнем децентралізації. Також проаналізовано основні вразливості, що виникають через динамічний характер обробки даних у таких середовищах, зокрема: асинхронність доступу, обмеженість обчислювальних ресурсів, складність синхронізації операцій і ризики перехоплення інформації під час передавання.*

***Наукова новизна.** Обговорено механізми забезпечення безпеки на рівні мережевої взаємодії, включаючи використання блокчейн-технологій для аудиту подій і трасування транзакцій.*

***Висновки.** Особливу увагу приділено методам захисту великих даних, зокрема криптографічним підходам, таким як гомоморфне шифрування (Rahman, 2022, с. 8–11), протоколи збереження конфіденційності на основі обчислювальної складності та механізми розподіленого управління ключами. Додатково розглянуто методи виявлення аномалій у поведінкових патернах систем на основі машинного навчання (Syed, 2021, с. 4–6) та алгоритмів глибокого навчання, які дозволяють ефективно ідентифікувати потенційні загрози в режимі реального часу.*

***Ключові слова:** великі дані, розподілені мережі, кібербезпека, захист інформації, криптографія, гомоморфне шифрування, виявлення аномалій, машинне навчання, блокчейн, управління ризиками, конфіденційність, цілісність даних, хмарні технології.*

Serhii DOROZHYNSKYI
PhD, Associate Professor at the Department of Cybersecurity, Information Technology and Economics, Dean of the Faculty of Information Technology and Social Sciences and Humanities, Kyiv University of Intellectual Property and Law of the National University "Odesa Law Academy", 210, Kharkivske highway, Kyiv, Ukraine, 02121, dorozhun1706@gmail.com
***ORCID:** 0000-0002-5395-6423*
Scopus author ID: 57222147991

Vitaliy TUPKALO
Doctor of Technical Sciences, Professor, Professor at the Department of Cybersecurity, Information Technologies and Economics, Kyiv University of Intellectual Property and Law of the National University "Odesa Law Academy", 210, Kharkivske highway, Kyiv, Ukraine, 02121,
***ORCID:** 0000-0002-6594-530X*

Yuriy SHCHERBYNA
Candidate of Technical Sciences, Associate Professor at the Department of Information Technologies, National University «Odesa Law Academy», 23, Fontanska doroha, Odesa, Ukraine, 65009, shcherbyna@onua.edu.ua
***ORCID:** 0000-0003-3885-6747*

To cite this article: Dorozhynskyi, S., Tupkalo, V., Shcherbyna, Y. (2025). Krytychni aspekty obrobky velykykh danykh dlia zabezpechennia yikh zakhystu v rozpodilenykh merezhakh [Critical aspects of big data processing to ensure its protection in distributed networks]. *Information Technology: Computer Science, Software Engineering and Cyber Security*, 60–67, doi: https://doi.org/10.32782/IT/2025-1-9

CRITICAL ASPECTS OF BIG DATA PROCESSING TO ENSURE ITS PROTECTION IN DISTRIBUTED NETWORKS

The article examines in detail the critical aspects of big data processing in distributed networks, which pose a modern challenge for ensuring cybersecurity.

***The purpose** of this work is to conduct an in-depth analysis of theoretical and practical approaches to solving the problems of confidentiality, integrity, and availability of data in multi-user environments with high interaction complexity. Particular attention is paid to the classification of threats that arise when processing large amounts of information, taking into account the specifics of distributed systems architectures, such as cloud technologies, peer-to-peer networks, and infrastructures with a high level of decentralization. The main vulnerabilities arising from the dynamic nature of data processing in such environments were also analyzed, including: asynchronous access, limited computing resources, complexity of synchronization of operations, and risks of interception of information during transmission.*

***Scientific novelty.** Mechanisms for ensuring security at the level of network interaction are discussed, including the use of blockchain technologies for event auditing and transaction tracing.*

***Conclusions.** Particular attention is paid to methods for protecting big data, in particular cryptographic approaches such as homomorphic encryption (Rahman, 2022, p. 8–11), computational complexity-based confidentiality protocols, and distributed key management mechanisms. Additionally, methods for detecting anomalies in behavioral patterns of systems based on machine learning (Syed, 2021, p. 4–6) and deep learning algorithms are considered, which allow for effective identification of potential threats in real time.*

***Key words:** big data, distributed networks, cybersecurity, information protection, cryptography, homomorphic encryption, anomaly detection, machine learning, blockchain, risk management, privacy, data integrity, cloud technologies.*

Вступ. У сучасну епоху цифрової трансформації великі дані стали ключовим ресурсом для організацій у різних галузях. Розподілені мережі, такі як хмарні обчислення, однорангові мережі та децентралізовані платформи, забезпечують масштабованість, доступність і ефективність обробки цих даних. Однак, зростання обсягів інформації та її критичності супроводжується появою нових загроз для її безпеки, що зумовлено динамічністю розподілених середовищ, асиметрією обчислювальних потужностей та складністю інтеграції різних компонентів системи.

Особливу небезпеку становлять кібератаки, спрямовані на компрометацію конфіденційності, цілісності та доступності даних. Уразливості мережевої інфраструктури, ненадійні протоколи передачі інформації, недоліки у системах доступу до даних створюють сприятливі умови для реалізації атак. Зокрема, зловмисники використовують специфіку розподілених мереж для атак відмови в обслуговуванні (DDoS), крадіжки даних та їх несанкціонованої модифікації. Крім того, проблема ускладнюється вимогами до швидкості обробки даних у режимі реального часу, що накладає обмеження на можливість використання традиційних методів захисту, які є надто ресурсоємними. У цьому контексті критично важливим є розроблення ефективних, масштабованих і адаптивних механізмів забезпечення безпеки, які враховують специфіку великих даних та архітектур розподілених мереж.

Мета дослідження полягає у визначенні, аналізі та розробці ефективних методів і підходів до забезпечення захисту великих даних

у розподілених мережах, які враховують специфіку таких середовищ, зокрема їхню динамічність, асинхронність і масштабованість.

Для досягнення цієї мети поставлено такі основні **завдання:**

1. Провести огляд та систематизацію сучасних загроз, що виникають у процесі обробки великих даних у розподілених мережах, з акцентом на проблеми конфіденційності, цілісності та доступності інформації.

2. Дослідити існуючі криптографічні методи захисту, їхню ефективність і можливості адаптації до умов розподілених середовищ.

3. Розробити та проаналізувати алгоритми виявлення загроз із використанням технологій машинного навчання, зокрема глибокого навчання, для ідентифікації аномалій у реальному часі.

4. Оцінити потенціал блокчейн-технологій для забезпечення безпеки великих даних у розподілених мережах, зокрема для аудиту подій і трасування транзакцій.

Один із ключових викликів у забезпеченні захисту великих даних у розподілених мережах полягає у необхідності балансування між ефективністю обробки інформації та її захистом. Багатокористувацькі середовища часто характеризуються високою динамічністю доступу до даних, що призводить до складності управління правами доступу, забезпечення прозорості операцій та запобігання несанкціонованому втручанню. Крім того, характер обробки великих даних вимагає інтеграції різноманітних технологій, таких як контейнеризація, оркестрація обчислень та автоматизація робочих процесів, що збільшує кількість можливих точок уразливості.

Особливого значення набувають проблеми забезпечення цілісності даних, коли зміна навіть невеликої частини інформації може призвести до значних порушень у функціонуванні системи або до появи недостовірних результатів. Водночас необхідність швидкої обробки даних у реальному часі ускладнює впровадження комплексних перевірок безпеки на кожному етапі обчислень.

Сучасні виклики також стосуються уніфікації протоколів і стандартів безпеки (Sun, 2023, с. 8) в умовах глобалізації інформаційного простору. Розподілені мережі часто охоплюють декілька юрисдикцій, де можуть діяти різні закони та регуляторні вимоги, що ускладнює створення універсальних підходів до захисту даних. Питання сумісності систем і забезпечення їхнього безперебійного функціонування у випадку загроз або атак є ще одним важливим аспектом.

Ще одним викликом є розвиток децентралізованих платформ, де обробка даних відбувається без чіткого централізованого контролю. У таких системах зростає ризик виникнення суперечностей між учасниками, недовіри до результатів обробки даних та поширення маніпуляцій у транзакціях. Ці виклики підкреслюють необхідність розробки не лише технічних рішень, але й нормативних та організаційних підходів, що забезпечують комплексний захист великих даних у розподілених середовищах.

Загрози великим даним у розподілених мережах можна класифікувати за трьома основними аспектами: конфіденційність, цілісність та доступність, що формують базову тріаду кібербезпеки – модель CIA. Кожен з цих аспектів має свої специфічні характеристики загроз та їх наслідки.

Конфіденційність визначає забезпечення того, що доступ до даних надається лише авторизованим суб'єктам. Формалізація порушення конфіденційності може бути описана формулою:

$$B(C) = P(U_i \in U_a),$$

де U_i – суб'єкт доступу, U_a – множина авторизованих суб'єктів, P – ймовірність несанкціонованого доступу. Основними типами атак на конфіденційність є перехоплення інформації та аналіз метаданих, зокрема через атаки типу «людина посередині» та атаки на зашифровані канали передачі даних.

Цілісність забезпечує недоторканність і точність даних під час їх обробки, зберігання чи передачі. Порушення цілісності можна виразити формулою:

$$B(I) = \exists\, x_i, x_j \in D,\ x_i \neq x_j,\ h(x_i) \neq h(j),$$

де D – множина даних, $h(x)$ – хеш-функція, яка використовується для верифікації даних. Загрози цілісності включають несанкціоновану модифікацію даних та впровадження фальшивих даних (data injection). Зокрема, атаки типу «перезапис даних» впливають на критичні системи, знижуючи їхню надійність.

Доступність передбачає забезпечення можливості легального доступу до даних і ресурсів у будь-який необхідний час. Порушення доступності описується формулою:

$$B(A) = \frac{D_t}{T_t} > \delta,$$

де δ – допустимий рівень простою, D_t – час недоступності системи, T_t – загальний час роботи системи.

Класифікація загроз за джерелами та методами атак.

Загрози для великих даних у розподілених мережах можна класифікувати за різними ознаками, зокрема за джерелами атак та методами їх здійснення (Liao, 2024, с. 5). Визначення цих категорій дозволяє створити більш точні моделі захисту та реагування на загрози. Основними джерелами загроз є внутрішні та зовнішні атаки, а також несанкціоноване використання ресурсів. Методи атак можна поділити на атакувальні та експлуатаційні, що використовуються для порушення кожного з аспектів безпеки: конфіденційності, цілісності та доступності.

Така класифікація дає змогу врахувати не лише типи атак, але й їхні джерела, що дозволяє створити більш гнучкі та ефективні стратегії захисту. Наприклад, для боротьби з внутрішніми загрозами необхідно застосовувати строгий контроль доступу до даних і моніторинг користувацької активності. Для зовнішніх загроз, зокрема атаки типу DDoS, необхідно використовувати методи фільтрації трафіку та захисту на рівні мережі. Окрім того, розробка методів швидкої ідентифікації фальшивих сертифікатів або атаки на верифікаційні механізми стане важливим кроком у забезпеченні безпеки в умовах постійного розвитку атакувальних технік.

Для ефективного захисту великих даних у розподілених мережах необхідно впроваджувати комплексні методи, орієнтовані на забезпечення конфіденційності, цілісності та доступності. Ці методи мають враховувати специфіку розподілених середовищ, таких як децентралізований доступ, велика кількість точок входу та високі обсяги даних, які підлягають обробці.

Захист конфіденційності спрямований на збереження інформації від несанкціонованого доступу. Основні методи (Park, 2022, с. 9):

1. Шифрування даних, де для захисту конфіденційності використовуються симетричні (AES) та асиметричні алгоритми (RSA, ECC). Процес шифрування можна описати формулою:

$$C = E_k(P),$$

де C – шифротекст, P – відкритий текст, E_k – функція шифрування з ключем k.

2. Анонімізація даних, де методи, такі як диференційна конфіденційність, додають шум до даних, щоб зберегти приватність користувачів.

3. Управління доступом, де впровадження багаторівневих політик, наприклад, Role-Based Access Control.

Цілісність даних гарантує, що інформація залишається незмінною та достовірною. Основні методи:

- Хешування – використовуються алгоритми, такі як SHA-256, для перевірки змін у даних.
- Цифрові підписи – забезпечують аутентифікацію джерела даних, використовуючи асиметричне шифрування.
- Контроль версій – зберігання попередніх версій даних для забезпечення можливості відновлення після модифікації.

Таблиця 1

Класифікація загроз за джерелами та методами атак

Категорія загрози	Джерело загрози	Методи атак	Цілі атаки	Приклад атаки
Внутрішні загрози	Співробітники або користувачі системи	1. Незаконний доступ до даних 2. Модифікація даних без авторизації	Конфіденційність Цілісність	Інсайдерські атаки або Витік даних
Зовнішні загрози	Зовнішні зловмисники або організації	Атаки типу «людина посередині» або розподілені атаки відмови в обслуговуванні	Конфіденційність Цілісність Доступність	DDoS-атаки Перехоплення даних
Загрози з боку технологій	Вразливості програмного та апаратного забезпечення	Експлуатація вразливостей ОС або програмного забезпечення; Використання недоліків у мережевих протоколах	Конфіденційність Цілісність Доступність	Злом баз даних; Мережеві атаки
Варіативні загрози	Атаки на механізми верифікації	Фальшиві сертифікати або підробка метаданих; Атаки на системи ідентифікації	Конфіденційність Цілісність	Підробка цифрових підписів; Використання фальшивих сертифікатів
Загрози, спричинені зовнішніми факторами	Непередбачувані обставини	Природні катастрофи; Втрати енергоенергії	Доступність	Втрата доступу до даних через стихійні лиха

• Доступність даних гарантує безперервний доступ до інформації, навіть у випадку загроз. Основні методи:

• Розподілене зберігання – використання систем типу RAID або децентралізованих мереж, що гарантують надмірність даних:

• Захист від атак типу DDoS – застосування механізмів фільтрації трафіку, наприклад, через мережі доставки контенту (CDN).

• Автоматизація резервування – регулярне створення резервних копій для швидкого відновлення в разі втрат.

Ці методи мають взаємодіяти в єдиній системі, щоб забезпечити повний спектр захисту даних. Наприклад, шифрування гарантує конфіденційність, але для верифікації даних потрібні цифрові підписи або хешування.

Розподілені мережі, що обробляють великі обсяги даних, характеризуються низкою специфічних особливостей, які суттєво впливають на впровадження та ефективність методів захисту інформації. До таких особливостей належать географічно розподілена структура, висока динаміка топології мережі, великий обсяг оброблюваних даних, а також середовище з великою кількістю користувачів.

Однією з ключових особливостей є **децентралізація** (Atman, 2022, с.7), що створює додаткові виклики для управління доступом та моніторингу безпеки. У розподілених мережах немає єдиного центру обробки даних, що ускладнює використання традиційних централізованих систем захисту. У таких умовах необхідне застосування методів, які дозволяють динамічно адаптувати політики доступу залежно від зміни топології мережі. Наприклад, рольові та атрибутивні системи контролю доступу повинні забезпечувати швидку реакцію на зміну ролей користувачів або стану ресурсів у реальному часі. Також важливим є забезпечення ефективності захисту в умовах масштабованості. Обробка великих даних вимагає обчислювальних ресурсів, які значно перевищують можливості традиційних підходів до кібербезпеки. Тому сучасні методи захисту повинні враховувати обмеження обчислювальної потужності, пропускної здатності мережі та затримки в обробці. Це призводить до зростання попиту на легковагові криптографічні алгоритми, методи компресії даних та оптимізації потоків трафіку, які мінімізують вплив на продуктивність системи.

Іншою особливістю є підвищена вразливість до атак через багатокористувацьке середовище (Lenard, 2023, с. 8–11). Взаємодія численних користувачів та пристроїв у розподіленій мережі створює ризики для конфіденційності та цілісності даних. Ці ризики можуть бути зменшені за допомогою застосування багаторівневого шифрування, обов'язкової багатофакторної автентифікації та інтеграції систем виявлення вторгнень із машинним навчанням для аналізу аномалій у поведінці користувачів. Важливим викликом є забезпечення безперебійної доступності даних у мережах із розподіленим зберіганням. У таких системах будь-яке пошкодження чи відмова окремих вузлів не повинна впливати на функціонування всієї мережі. Цього досягають за допомогою надмірності даних, регулярного резервування, а також впровадження механізмів автоперевірки та самовідновлення даних.

Розподілені обчислення є фундаментальною основою для роботи з великими даними, оскільки вони забезпечують паралельну обробку інформації, оптимізацію використання ресурсів і масштабованість систем. Проте така архітектура накладає значні вимоги та створює нові виклики для побудови систем захисту даних. Вони характеризуються багаточисельністю вузлів, які можуть бути географічно розподілені. Це потребує впровадження децентралізованих систем захисту, здатних функціонувати без централізованого керування. Зокрема, у таких системах використовується блокчейн, що забезпечує надійну верифікацію транзакцій, або криптографічні алгоритми, які дозволяють обмінюватися ключами без необхідності централізованих довірених сторін.

Другим критичним аспектом є масштабованість архітектури захисту. Системи захисту мають бути здатними до динамічного розширення для обробки зростаючих обсягів даних та кількості запитів. У таких умовах традиційні методи шифрування та аутентифікації можуть бути неефективними через затримки в обробці. Це вимагає застосування легковагових протоколів, таких як криптографія еліптичних кривих, а також інтеграції методів виявлення загроз на основі штучного інтелекту. Крім того, розподілені обчислення створюють ризики, пов'язані з гетерогенністю інфраструктури. У більшості випадків розподілені системи складаються з вузлів, що працюють на різних платформах та використовують різні технології. Це підвищує ймовірність експлуатації вразливостей у сумісності компонентів системи. Для вирішення цієї проблеми необхідно впроваджувати стандартизовані протоколи безпеки та механізми взаємної аутентифікації між вузлами. Також важливим є управління доступом до ресурсів у розподіленій архітектурі. Через велику кількість користувачів і запитів виникає ризик перевантаження

системи та зростання ймовірності атак, таких як DDoS. Для запобігання цьому застосовуються політики обмеження доступу, розподілення навантаження між вузлами та динамічне масштабування ресурсів. Ще одним викликом є забезпечення прозорості та контролю за безпекою в умовах розподілених обчислень. Для цього використовуються інтегровані системи моніторингу, які аналізують поведінку вузлів та запитів у режимі реального часу. Такі системи здатні автоматично виявляти аномалії, що можуть свідчити про можливу атаку, та вживати заходів для її нейтралізації.

Ефективний захист великих даних у розподілених мережах вимагає впровадження сучасних систем моніторингу та виявлення загроз, які здатні працювати в реальному часі, враховуючи динамічну природу таких мереж. Ці системи базуються на аналізі великих обсягів трафіку, виявленні аномалій у поведінці користувачів і мережевих вузлів, а також на здатності оперативно реагувати на інциденти безпеки.

Одним із основних компонентів сучасних систем моніторингу є механізми збору та аналізу телеметричних даних з усіх вузлів мережі. У розподілених мережах ці механізми мають бути децентралізованими та масштабованими, щоб обробляти дані з багатьох джерел одночасно. Важливу роль у цьому відіграють технології обробки потокових даних, наприклад, Apache Kafka або Apache Flink, які дозволяють ефективно обробляти події в реальному часі. Системи виявлення загроз, у свою чергу, використовують алгоритми машинного навчання та штучного інтелекту для аналізу аномалій. Наприклад, методи кластеризації, такі як *k* *k*-середні, або нейронні мережі для виявлення складних патернів у поведінці. Ці алгоритми дозволяють не лише ідентифікувати загрози, які відповідають відомим шаблонам атак, але й прогнозувати потенційно нові типи атак, що не мають історичних аналогів.

Інтеграція моніторингу із засобами реагування на інциденти також є важливим аспектом. Такі системи, як SIEM, дозволяють автоматизувати процеси виявлення та реагування, надаючи аналітичну інформацію для операторів безпеки. Наприклад, при виявленні підозрілої активності в мережевому трафіку система може автоматично ізолювати вузол або заблокувати доступ, запобігаючи поширенню загрози. Ще одним критично важливим компонентом є забезпечення прозорості дій системи моніторингу для адміністраторів. Це досягається через побудову багаторівневої системи візуалізації даних, яка дозволяє отримувати уявлення про стан мережі та можливі загрози. Застосування графічних інтерфейсів, наприклад, на базі систем, таких як Kibana, полегшує аналіз даних навіть у великих масштабах. Окремо слід зазначити важливість захисту самих систем моніторингу, які також стають мішенню для зловмисників. Для цього впроваджуються методи захисту від атак типу «людина посередині», багаторівневе шифрування телеметричних даних і використання технологій контейнеризації для ізоляції компонентів.

Інтеграція засобів захисту в архітектуру розподілених мереж є ключовим аспектом забезпечення безпеки великих даних. Цей процес вимагає врахування специфіки розподілених середовищ, таких як децентралізація, гетерогенність інфраструктури та динамічна природа потоків даних. Комплексний підхід до інтеграції повинен забезпечувати захист на всіх рівнях архітектури, включаючи інфраструктурний, мережевий і прикладний рівні.

На інфраструктурному рівні впроваджуються технології ізоляції та сегментації, які дозволяють мінімізувати ризики поширення загроз у разі компрометації окремих вузлів. Наприклад, використання віртуалізації та контейнеризації (Docker, Kubernetes) забезпечує логічну ізоляцію процесів і середовищ виконання. Це унеможливлює доступ зловмисника до інших компонентів системи навіть при успішній атаці на окремий вузол.

На мережевому рівні інтеграція засобів захисту передбачає впровадження механізмів шифрування каналів передачі даних (наприклад, протоколи TLS/SSL) та захисту від атак на рівні маршрутизації (наприклад, BGP-Sec). Важливим є також використання сегментації мережі за допомогою віртуальних локальних мереж (VLAN) для обмеження взаємодії між різними групами вузлів. У розподілених системах все частіше застосовуються програмно визначені мережі (SDN), які забезпечують централізований контроль і адаптацію політик безпеки до динамічних змін у мережі.

На прикладному рівні інтеграція включає впровадження багатофакторної автентифікації, ролі та атрибутивних моделей управління доступом, а також інструментів контролю та аудиту. Унікальність розподілених середовищ вимагає від цих механізмів високої адаптивності: вони мають враховувати розподіл ресурсів і забезпечувати узгодженість політик між різними вузлами. Наприклад, при зміні ролі користувача політика доступу має автоматично оновлюватися у всіх компонентах системи.

Особливе значення для інтеграції засобів захисту має автоматизація. У розподілених мережах великі обсяги даних і висока динамічність процесів ускладнюють ручне управління безпекою. Тому інтеграція автоматизованих систем, таких як SOAR (Security Orchestration, Automation, and Response), дозволяє автоматично виконувати реагування на інциденти безпеки. Наприклад, при виявленні аномальної активності у певному вузлі система може ініціювати ізоляцію цього вузла, змінити політики доступу чи перенаправити трафік на інші вузли. Ще одним важливим компонентом є інтеграція механізмів самодіагностики та самовідновлення. Це дозволяє системі проактивно виявляти вразливості, перевіряти відповідність налаштувань сучасним стандартам безпеки та автоматично відновлювати функціональність у разі збоїв. Наприклад, використання реплікації даних і механізмів консенсусу (RAFT, Paxos) у розподілених сховищах даних гарантує їхню доступність навіть при втраті частини вузлів.

Висновки. У статті досліджено критичні аспекти обробки великих даних у розподілених мережах із врахуванням викликів, пов'язаних із забезпеченням їхнього захисту. Аналізуючи сучасні методи обробки, моніторингу та захисту даних, встановлено, що розподілені системи є важливою складовою інфраструктури великих даних, але їхня децентралізація, масштабованість і динамічна природа значно ускладнюють побудову ефективних систем безпеки. Зокрема, структура розподілених мереж, що характеризується географічною віддаленістю вузлів, гетерогенністю середовищ виконання та високою інтенсивністю трафіку, накладає додаткові вимоги до засобів захисту. У межах дослідження було визначено, що традиційні підходи до забезпечення конфіденційності, цілісності та доступності інформації стають недостатніми, особливо в умовах масштабування систем. У зв'язку з цим запропоновано використання легковагових криптографічних алгоритмів, зокрема криптографія еліптичних кривих, та технологій блокчейн для децентралізованого зберігання даних і верифікації транзакцій. Моніторинг загроз і управління інцидентами є критично важливими для захисту великих даних у розподілених мережах. Запропоновано інтеграцію механізмів аналізу даних у реальному часі, використання штучного інтелекту для виявлення аномалій та автоматизацію процесів реагування. Системи типу SIEM і SOAR, орієнтовані на автоматичне виконання операцій із запобігання загрозам, виявилися ефективними для зменшення часу реагування та мінімізації впливу інцидентів на функціональність системи. Інтеграція засобів захисту в архітектуру розподілених мереж повинна враховувати багаторівневий підхід. На інфраструктурному рівні рекомендується використовувати ізоляцію вузлів через віртуалізацію та контейнеризацію. На мережевому рівні застосування програмно визначених мереж (SDN) дозволяє динамічно адаптувати політики безпеки до змін у топології. На прикладному рівні важливим є впровадження багатофакторної автентифікації, адаптивного управління доступом та механізмів аудиту. Високий рівень автоматизації процесів захисту є обов'язковим компонентом сучасних систем. Завдяки впровадженню технологій самовідновлення та механізмів самодіагностики досягається проактивний підхід до усунення загроз. Інструменти реплікації даних і консенсусу, такі як RAFT і Paxos, забезпечують доступність і цілісність даних навіть у випадку втрати окремих вузлів. Отримані результати дослідження підтверджують, що ефективний захист великих даних у розподілених мережах можливий лише за умови використання інноваційних технологій, адаптивних методів і комплексного підходу до забезпечення безпеки. Інтеграція передових технологій, таких як блокчейн, штучний інтелект, програмно визначені мережі та автоматизовані системи управління інцидентами, є ключовими для створення стійких і динамічних систем захисту, які відповідають вимогам сучасних інформаційних середовищ.

ЛІТЕРАТУРА:

1. M. S. Rahman and H. Reza, "A Systematic Review Towards Big Data Analytics in Social Media", in *Big Data Mining and Analytics*, vol. 5, no. 3, pp. 228–244, September 2022, doi: https://doi.org/10.26599/BDMA.2022.9020009

2. D. Syed, A. Zainab, A. Ghrayeb, S. S. Refaat, H. Abu-Rub and O. Bouhali, "Smart Grid Big Data Analytics: Survey of Technologies, Techniques, and Applications", in *IEEE Access*, vol. 9, pp. 59564–59585, 2021, doi: https://doi.org/10.1109/ACCESS.2020.3041178

3. X. Sun, Y. He, D. Wu and J. Z. Huang, "Survey of Distributed Computing Frameworks for Supporting Big Data Analysis", in *Big Data Mining and Analytics*, vol. 6, no. 2, pp. 154–169, June 2023, doi: https://doi.org/10.26599/BDMA.2022.9020014.

4. J. Liao and J. Lin, “A Distributed Deep Reinforcement Learning Approach for Reactive Power Optimization of Distribution Networks”, в *IEEE Access*, том. 12, стор. 113898–113909, 2024, doi: https://doi.org/10.1109/ACCESS.2024.3445143

5. D. Park, S. Kang and C. Joo, “A learning-based distributed algorithm for scheduling in multi-hop wireless networks”, in *Journal of Communications and Networks*, vol. 24, no. 1, pp. 99–110, Feb. 2022, doi: https://doi.org/10.23919/JCN.2021.000030

6. M. W. S. Atman and A. Gusrialdi, “Finite-Time Distributed Algorithms for Verifying and Ensuring Strong Connectivity of Directed Networks”, in *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 6, pp. 4379–4392, 1 Nov.-Dec. 2022, doi: https://doi.org/10.1109/TNSE.2022.3200466

7. T. Lenard, A. Collen, M. Benyahya, N. A. Nijdam and B. Genge, “Exploring Trust Modeling and Management Techniques in the Context of Distributed Wireless Networks: A Literature Review”, in *IEEE Access*, vol. 11, pp. 106803–106832, 2023, doi: https://doi.org/10.1109/ACCESS.2023.3320945

REFERENCES:

1. Rahman, M. S., & Reza, H. (2022). A systematic review towards big data analytics in social media. *Big Data Mining and Analytics*, 5(3), 228–244. https://doi.org/10.26599/BDMA.2022.9020009

2. Syed, D., Zainab, A., Ghrayeb, A., Refaat, S. S., Abu-Rub, H., & Bouhali, O. (2021). Smart grid big data analytics: Survey of technologies, techniques, and applications. *IEEE Access*, 9, 59564–59585. https://doi.org/10.1109/ACCESS.2020.3041178

3. Sun, X., He, Y., Wu, D., & Huang, J. Z. (2023). Survey of distributed computing frameworks for supporting big data analysis. *Big Data Mining and Analytics*, 6(2), 154–169. https://doi.org/10.26599/BDMA.2022.9020014

4. Liao, J., & Lin, J. (2024). A distributed deep reinforcement learning approach for reactive power optimization of distribution networks. *IEEE Access*, 12, 113898–113909. https://doi.org/10.1109/ACCESS.2024.3445143

5. Park, D., Kang, S., & Joo, C. (2022). A learning-based distributed algorithm for scheduling in multi-hop wireless networks. *Journal of Communications and Networks*, 24(1), 99–110. https://doi.org/10.23919/JCN.2021.000030

6. Atman, M. W. S., & Gusrialdi, A. (2022). Finite-time distributed algorithms for verifying and ensuring strong connectivity of directed networks. *IEEE Transactions on Network Science and Engineering*, 9(6), 4379–4392. https://doi.org/10.1109/TNSE.2022.3200466

7. Lenard, T., Collen, A., Benyahya, M., Nijdam, N. A., & Genge, B. (2023). Exploring trust modeling and management techniques in the context of distributed wireless networks: A literature review. *IEEE Access*, 11, 106803–106832. https://doi.org/10.1109/ACCESS.2023.3320945