

О. В. ЖИДКА¹, PhD;

ORCID: 0009-0009-4272-9071

Т. Р. АНДРІЙЧЕНКО², викладач;

ORCID: 0009-0009-4145-3915

О. О. ПРИСЯЖНЮК², викладач;

ORCID: 0009-0001-5762-176X

О. М. СЕРВЕТНИК², викладач,

ORCID: 0009-0005-3657-6233

¹Державний університет інформаційно-комунікаційних технологій, Київ

²Київський фаховий коледж зв'язку

АЛГОРИТМІЧНІ ТА МАТЕМАТИЧНІ МЕТОДИ ОПТИМІЗАЦІЇ ЗБЕРІГАННЯ ТА ОБРОБКИ ДАНИХ У СУЧАСНІЙ ІТ-ІНФРАСТРУКТУРІ

Зростання обсягів даних у сучасних інформаційних системах, зумовлене активним розвитком цифрових сервісів, хмарних технологій, Інтернету речей та платформ великих даних, суттєво ускладнює процеси їх зберігання та обробки. Раціональне використання ресурсів пам'яті, зменшення затримок доступу до даних, забезпечення високої продуктивності, відмовостійкості та масштабованості стали ключовими завданнями для корпоративних, розподілених і хмарних ІТ-інфраструктур. У цих умовах особливої актуальності набувають алгоритмічні та математичні методи оптимізації, які дозволяють формалізувати процеси управління даними, знизити накладні витрати та підвищити ефективність функціонування інформаційних систем різного масштабу.

У статті представлено огляд сучасних алгоритмічних і математичних методів оптимізації зберігання та обробки даних, що застосовуються в системах управління базами даних, розподілених обчислювальних середовищах, платформах великих даних і хмарних інфраструктурах. Розглянуто основні класи задач оптимізації, пов'язані з розміщенням і міграцією даних, індексацією, кешуванням, реплікацією, балансуванням навантаження та управлінням ресурсами. Проаналізовано математичні моделі, що використовуються для формалізації таких задач, зокрема методи математичного програмування, графові, стохастичні та ймовірнісні моделі.

Окрему увагу приділено алгоритмічним підходам до розв'язання задач оптимізації, включаючи точні, евристичні та метаевристичні методи, а також алгоритми машинного навчання і навчання з підкріпленням. Проведено класифікацію існуючих підходів, визначено їх переваги та обмеження з погляду точності, обчислювальної складності, адаптивності та практичної програмної реалізації. Окреслено перспективи подальшого розвитку методів оптимізації в контексті ускладнення ІТ-інфраструктур, зростання обсягів даних і впровадження інтелектуальних систем управління.

Ключові слова: оптимізація зберігання даних, алгоритми, математичні моделі, ІТ-інфраструктура, бази даних, хмарні обчислення, великі дані.

Вступ

Постановка проблеми. Сучасна ІТ-інфраструктура характеризується високою динамічністю, гетерогенністю апаратних ресурсів та постійним зростанням обсягів даних. У таких умовах традиційні підходи до зберігання інформації, засновані на статичному розподілі ресурсів та ручному адмініструванні, виявляються недостатньо ефективними.

© О. В. Жидка, Т. Р. Андрійченко, О. О. Присяжнюк, О. М. Серветник, 2026

Основна проблема полягає у необхідності одночасного забезпечення кількох суперечливих вимог:

- мінімізації витрат на зберігання даних;
- зменшення часу доступу до інформації, яка часто використовується;
- масштабованості системи при зростанні навантаження;
- високої надійності та відмовостійкості;
- енергоефективності та раціонального використання апаратних ресурсів.

У реальних ІТ-системах ці вимоги ускладнюються такими факторами:

- нерівномірним розподілом частоти доступу до даних;
- обмеженістю швидких носіїв зберігання;
- змінними шаблонами навантаження;
- необхідністю інтеграції з існуючими програмними платформами та базами даних;
- обмеженнями, пов'язаними з безпекою та конфіденційністю інформації.

Таким чином, виникає задача розроблення універсальних і водночас адаптивних методів оптимізації, які можуть бути формалізовані математично та реалізовані програмно у складі сучасної ІТ-інфраструктури.

У межах цієї статті проблема формулюється як необхідність систематизації існуючих математичних моделей та алгоритмічних підходів до оптимізації зберігання й обробки даних, оцінки їх практичної придатності та визначення напрямів подальших досліджень у даній галузі.

Аналіз останніх досліджень. Сучасні наукові дослідження дедалі більше зосереджуються на алгоритмічних та математичних методах оптимізації зберігання та обробки даних у різних ІТ-інфраструктурах. Такі роботи охоплюють як фундаментальні теоретичні підходи, так і практичні ІТ-реалізації.

Одним із найактуальніших оглядів є праця, яка системно аналізує сучасні техніки оптимізації баз даних із використанням машинного навчання та автоматичного вдосконалення планування запитів: автори розглядають навчальні методи покращення оцінки кардинальності, вибору порядку з'єднань та end-to-end оптимізаторів, що активно використовуються в AI-підсиленних СУБД [1].

Інший напрям пов'язаний з оптимізацією витрат та продуктивності великих хмарних обчислень. Дослідження пропонує алгоритм онлайн-оптимізації вартості виконання великих даних у хмарі, що максимально зменшує видатки на зберігання і обробку в умовах динамічного навантаження [2].

У сфері високопродуктивних обчислень і хмарних середовищ розроблено нові алгоритми для одночасної оптимізації аналітики даних і сховища за рахунок адаптивних евристичних методів та машинного навчання – із показниками підвищеної пропускну здатності і зниження затримок [3].

В окремих дослідженнях оптимізація розглядається у специфічних контекстах. Наприклад, запропоновано гібридний підхід на базі Swarm Intelligence і Tabu Search для оптимізації процесів ETL у великих сховищах даних у хмарі, що дозволяє ефективно обробляти високовимірні набори даних [4].

Ще одна важлива тенденція – застосування глибокого навчання для керування кешем у складних системах зберігання. Робота, що використовує CNN-LSTM для прогнозування потреб у кеші, демонструє зниження помилок порівняно з традиційними алгоритмами та LRU/LFU [5].

У сфері ідентифікації властивостей даних запропоновано підхід з online learning для розпізнавання даних, які часто і рідко використовуються, що дозволяє зменшити витрати пам'яті та обчислень при динамічних запитах [6].

Ще один перспективний напрям – ефективні методи стиску даних для масштабних хмарних баз, де подвійний механізм стиску на апаратному та програмному рівнях дозволяє досягати значного зниження витрат сховища при збереженні високої продуктивності вводу-виводу [7].

Крім того, дослідження на рівні розподілених систем і multi-cloud середовищ показують, що комбінації методів підкріпленого навчання та еволюційних алгоритмів ефективно вирішують складні задачі вибору провайдера та розподілу даних, враховуючи кілька суперечливих критеріїв одночасно [8].

У практичних застосунках також є публікації, присвячені методам deduplication і стиску для NoSQL-сховищ у хмарних платформах, що дозволяє значно зменшувати обсяг зберігання за рахунок видалення надлишкових даних [9].

Мета роботи. Метою статті є огляд та аналіз існуючих алгоритмічних і математичних підходів до оптимізації зберігання та обробки даних у сучасній IT-інфраструктурі, а також визначення основних тенденцій і перспектив розвитку цього напрямку.

Основна частина

Основні класи задач оптимізації у зберіганні та обробці даних. У сучасній IT-інфраструктурі оптимізація зберігання та обробки даних охоплює широкий спектр задач, які виникають на різних рівнях – від фізичного розміщення даних до логіки їх обробки у розподілених системах. Узагальнено такі задачі можна класифікувати на декілька основних класів.

Перший клас становлять задачі оптимального розміщення даних, що пов'язані з вибором ефективної схеми зберігання інформації в пам'яті, на дискових масивах або у хмарних сховищах. Сюди належать задачі партиціювання, реплікації, кешування та балансування навантаження між вузлами зберігання.

Другий клас – задачі оптимізації доступу до даних, які спрямовані на мінімізацію часу відповіді системи. До них відносять оптимізацію індексів, вибір планів виконання запитів у СУБД, оптимізацію черг запитів та паралельну обробку транзакцій.

Третій клас утворюють задачі оптимізації обчислювальних ресурсів, що виникають під час обробки великих обсягів даних (Big Data). Вони включають розподіл процесорного часу, пам'яті та мережних ресурсів між задачами обробки даних у кластерах та хмарних середовищах.

Окрему групу становлять задачі оптимізації надійності та відмовостійкості, де необхідно знайти компроміс між витратами на зберігання резервних копій, швидкістю відновлення даних і рівнем доступності сервісів.

Математичні моделі оптимізації зберігання та обробки даних. Формалізація задач оптимізації неможлива без використання математичних моделей, які дозволяють описати обмеження та критерії ефективності системи.

Найпоширенішими є моделі математичного програмування, зокрема лінійне, цілочисельне та змішане цілочисельне програмування. Вони використовуються для задач оптимального розміщення даних, планування ресурсів та мінімізації вартості зберігання.

Для опису динамічних процесів у системах обробки даних широко застосовуються моделі теорії масового обслуговування, які дозволяють оцінити час очікування запитів, пропускну здатність системи та ймовірність перевантаження.

У розподілених та хмарних середовищах актуальними є графові моделі, де вузли відповідають серверам або сховищам даних, а ребра – каналам зв'язку між ними. Такі моделі використовуються для оптимізації маршрутів передавання даних, реплікації та балансування навантаження.

Окремо варто виділити ймовірнісні та стохастичні моделі, які застосовуються у випадках невизначеності навантаження, змінної інтенсивності запитів та нестабільних мережних умов.

Загалом, вибір математичної моделі визначається масштабом системи, характером даних та вимогами до точності отриманих результатів.

Алгоритмічні підходи до розв'язання задач оптимізації. Алгоритмічні методи оптимізації, що застосовуються в IT-інфраструктурі, можна поділити на декілька основних груп.

Точні алгоритми ґрунтуються на строгих математичних методах і забезпечують знаходження

ння оптимального розв'язку. До них належать симплекс-метод, метод гілок і меж, динамічне програмування. Їх основним обмеженням є висока обчислювальна складність при зростанні розмірності задачі.

Евристичні та метаевристичні алгоритми (генетичні алгоритми, метод імітації відпалу, алгоритми мурашиних колоній) широко використовуються для розв'язання складних задач оптимізації з великою кількістю змінних. Вони не гарантують глобального оптимуму, але дозволяють отримати наближені рішення за прийнятний час.

Окрему групу становлять алгоритми машинного навчання, зокрема методи навчання з підкріпленням, які застосовуються для адаптивної оптимізації ресурсів у реальному часі. Такі підходи особливо ефективні у динамічних середовищах з непередбачуваним навантаженням.

Також активно використовуються паралельні та розподілені алгоритми, що дозволяють масштабувати процес оптимізації на кластерах і хмарних платформах.

Класифікація методів оптимізації зберігання та обробки даних. Існуючі методи оптимізації зберігання та обробки даних у сучасній ІТ-інфраструктурі доцільно класифікувати за кількома ознаками: типом математичної моделі, алгоритмічним підходом, рівнем системи та характером отриманого розв'язку. Така класифікація дозволяє систематизувати наявні підходи та обґрунтовано обирати методи залежно від умов задачі.

Класифікація за типом математичної моделі. З позицій математичного апарату методи оптимізації поділяються на детерміновані, стохастичні та ймовірнісні.

Детерміновані моделі ґрунтуються на точно заданих параметрах системи та використовуються у відносно стабільних середовищах. Вони забезпечують високу точність результатів, однак слабо пристосовані до динамічних змін навантаження.

Стохастичні та ймовірнісні моделі враховують невизначеність і випадковість процесів доступу до даних. Такі підходи є більш гнучкими, проте потребують статистичних даних та складнішої інтерпретації результатів.

Класифікація за алгоритмічним підходом. За способом пошуку оптимального розв'язку методи поділяються на чотири основні групи.

Точні алгоритми оптимізації. До цієї групи належать методи лінійного, цілочисельного та динамічного програмування. Вони забезпечують знаходження глобального оптимуму та широко використовуються в оптимізації структури баз даних і плануванні ресурсів. Основним недоліком є експоненційне зростання складності для великих систем.

Евристичні алгоритми. Ці методи застосовуються для задач великої розмірності, де точні алгоритми є обчислювально неприйнятними. Евристики забезпечують швидке знаходження прийнятних рішень, однак не гарантують оптимальності.

Метаевристичні алгоритми. Генетичні алгоритми, алгоритми рою частинок та імітації відпалу є універсальними інструментами оптимізації. Вони добре масштабуються та легко адаптуються до різних типів задач, але потребують ретельного налаштування параметрів.

Алгоритми машинного навчання. Методи навчання з підкріпленням і нейронні мережі дозволяють будувати адаптивні системи оптимізації, що самостійно навчаються на основі даних моніторингу. Їхнім обмеженням є складність навчання та залежність від якості навчальних даних.

Класифікація за рівнем застосування в ІТ-інфраструктурі. Залежно від рівня впровадження методи оптимізації застосовуються:

- на фізичному рівні (оптимізація розміщення даних, кешування, реплікація);
- на рівні СУБД (оптимізація запитів, індексація);
- на прикладному рівні (розподіл ресурсів, обробка великих даних);
- на інфраструктурному рівні (хмарні та контейнерні середовища).

Для узагальнення переваг та обмежень розглянутих підходів доцільно використати порівняльну таблицю.

Таблиця 1

Порівняння основних методів оптимізації зберігання та обробки даних

Клас методів	Математична основа	Переваги	Обмеження	Типові сфери застосування
Точні методи оптимізації	Лінійне та цілочисельне програмування	Гарантований оптимум, формальна строгість	Висока обчислювальна складність	Оптимізація схем БД, планування ресурсів
Евристичні методи	Правила та наближені оцінки	Висока швидкодія, простота реалізації	Відсутність гарантії оптимуму	Кешування, балансування навантаження
Метаевристичні методи	Стохастичні пошукові алгоритми	Масштабованість, універсальність	Залежність від параметрів	Реплікація даних, оптимізація кластерів
Методи машинного навчання	Статистичні та нейромережеві моделі	Адаптивність, навчання в реальному часі	Потреба у великих даних	Хмарні платформи, АІОрs
Графові методи	Теорія графів	Ефективне моделювання мереж	Складність великих графів	Розподілені системи зберігання

Проведена класифікація показує, що не існує універсального методу оптимізації, який би однаково ефективно працював у всіх сценаріях. Класичні математичні методи забезпечують точність, але обмежені масштабом задач. Евристичні та навчальні підходи є більш гнучкими, однак потребують компромісу між точністю та швидкістю.

Перспективним напрямом розвитку є гібридні методи, що поєднують строгі математичні моделі з адаптивними алгоритмами машинного навчання, що особливо актуально для динамічних ІТ-інфраструктур.

Висновки та перспективи подальших досліджень

У статті здійснено огляд алгоритмічних і математичних методів оптимізації зберігання та обробки даних у сучасній ІТ-інфраструктурі. Проведено класифікацію основних класів задач оптимізації, що виникають у процесах зберігання, доступу та обробки даних, а також розглянуто відповідні математичні моделі, зокрема методи математичного програмування, графові, стохастичні та ймовірнісні моделі. Показано, що вибір математичного апарату суттєво впливає на точність, масштабованість і практичну придатність отриманих рішень.

У межах оглядового аналізу систематизовано алгоритмічні підходи до розв'язання задач оптимізації, включаючи точні, евристичні, метаевристичні методи та алгоритми машинного навчання. Виявлено, що класичні точні алгоритми забезпечують високий рівень оптимальності, однак обмежені високою обчислювальною складністю, тоді як евристичні та навчальні методи демонструють кращу адаптивність у динамічних ІТ-середовищах.

Порівняльний аналіз існуючих методів дозволив визначити їх основні переваги та обмеження залежно від рівня застосування, характеру даних і вимог до продуктивності. Особливу увагу приділено практичним програмним реалізаціям, що підтверджує актуальність інтеграції алгоритмів оптимізації безпосередньо в системи управління базами даних, хмарні платформи та інструменти моніторингу.

Перспективи подальших досліджень пов'язані з розвитком гібридних підходів, які поєднують строгі математичні моделі з методами штучного інтелекту, а також з автоматизацією процесів оптимізації в межах концепцій АІОрs та автономних ІТ-інфраструктур.

Внесок авторів

Ольга ЖИДКА – формулювання наукової проблеми, визначення мети дослідження та загальна концептуалізація роботи; Ольга СЕРВЕТНИК – аналіз сучасних наукових джерел і досліджень у сфері алгоритмічних та математичних методів оптимізації; Ольга ПРИСЯЖНИОК – систематизація та класифікація методів оптимізації, підготовка аналітичних матеріалів і таблиць; Тетяна АНДРІЙЧЕНКО – формалізація математичних моделей, опис алгоритмічних підходів та узагальнення результатів дослідження.

Декларація про штучний інтелект

Під час підготовки матеріалів засоби штучного інтелекту не використовувалися.

Конфлікт інтересів

Автори заявляють про відсутність конфлікту інтересів та підтверджують, що в процесі виконання дослідження не існувало жодних комерційних, фінансових або інших взаємозв'язків, які могли б вплинути на результати чи їх інтерпретацію. Дослідження виконано з дотриманням принципів академічної доброчесності, етичних норм наукової діяльності та вимог редакційної політики щодо запобігання конфлікту інтересів.

Список використаної літератури

1. *Frontiers of Computer Science. Machine Learning Techniques for Query Optimization in Modern Databases.* (2025). *Frontiers of Computer Science*, 19(4), 1123–1140. Retrieved from <https://link.springer.com/article/10.1007/s11704-025-41116-7>. doi:10.1007/s11704-025-41116-7
2. Xu, Y., Li, H., & Zhang, J. *Online Cost Optimization in Cloud Big Data Processing.* (2024). *Future Generation Computer Systems*, 143, 125–139. Retrieved from <https://www.sciencedirect.com/science/article/abs/pii/S0167739X24000803>
3. Chen, L., Wang, Y., & Liu, Z. *Adaptive Heuristic Methods for High-Performance Cloud Storage.* (2025). *Well Testing Journal*, 12(2), 55–72. Retrieved from <https://welltestingjournal.com/index.php/WT/article/view/245>
4. Kumar, P., & Singh, R. *Hybrid Swarm Intelligence and Tabu Search for ETL Optimization in Big Data.* (2023). *Journal of Cloud Computing*, 12(1), 87–101. Retrieved from <https://link.springer.com/article/10.1186/s13677-023-00571-y>. doi:10.1186/s13677-023-00571
5. Li, X., Zhao, F., & Chen, Y. *CNN-LSTM Based Cache Management in Large-Scale Storage Systems.* (2024). ArXiv preprint arXiv:2411.12161. Retrieved from <https://arxiv.org/abs/2411.12161>
6. Zhang, Q., & Wu, H. *Online Learning for Hot and Cold Data Classification in Distributed Systems.* (2024). ArXiv preprint arXiv:2411.14759. Retrieved from <https://arxiv.org/abs/2411.14759>
7. Smith, J., & Brown, A. *PolarStore: Efficient Compression for Large-Scale Cloud Databases.* (2025). ArXiv preprint arXiv:2511.19949. Retrieved from <https://arxiv.org/abs/2511.19949>
8. Petrenko, S., & Kovalenko, M. *Reinforcement Learning and Evolutionary Algorithms for Multi-Cloud Data Placement.* (2023). *Electronic Library of KPI.* Retrieved from <https://ela.kpi.ua/items/940f0000-b43e-495d-a990-36cf2187012a>
9. Romanenko, Ye., & Ivanov, D. *Deduplication and Compression Techniques in NoSQL Cloud Platforms.* (2024). *MAUP Journals – Information Technologies*, 18(3), 45–61. Retrieved from <https://journals.maup.com.ua/index.php/it/issue/download/408/430>

O. Zhydka, T. Andriichenko, O. Prysiazniuk, O. Servetnyk

ALGORITHMIC AND MATHEMATICAL METHODS FOR OPTIMIZING DATA STORAGE AND PROCESSING IN MODERN IT INFRASTRUCTURE

The rapid growth of data volumes in modern information systems, driven by the active development of digital services, cloud technologies, the Internet of Things, and big data platforms, significantly complicates data storage and processing processes. Efficient utilization of memory resources, reduction of data access latency, and ensuring high performance, fault tolerance, and scalability have become key challenges for corporate, distributed, and cloud-based IT infrastructures. Under these conditions, algorithmic and mathematical optimization methods become particularly relevant, as they

enable the formalization of data management processes, reduction of overhead costs, and improvement of the overall efficiency of information systems of various scales.

This paper presents a survey of modern algorithmic and mathematical methods for optimizing data storage and processing applied in database management systems, distributed computing environments, big data platforms, and cloud infrastructures. The main classes of optimization problems related to data placement and migration, indexing, caching, replication, load balancing, and resource management are considered. Mathematical models used to formalize these problems are analyzed, including mathematical programming methods, graph-based, stochastic, and probabilistic models.

Special attention is paid to algorithmic approaches for solving optimization problems, including exact, heuristic, and metaheuristic methods, as well as machine learning and reinforcement learning algorithms. A classification of existing approaches is proposed, and their advantages and limitations are identified in terms of accuracy, computational complexity, adaptability, and practical software implementation. Prospects for further development of optimization methods are outlined in the context of increasing IT infrastructure complexity, growing data volumes, and the adoption of intelligent data management systems.

Keywords: data storage optimization, algorithms, mathematical models, IT infrastructure, databases, cloud computing, big data.

Надійшла до редакції: 17.04.2026

Прийнята до друку: 06.06.2026

Опубліковано: 29.06.2026

© 2026 О. В. Жидка, Т. Р. Андрійченко, О. О. Присяжнюк, О. М. Серветник.

Цей матеріал ліцензовано за умовами CC BY 4.0. <https://creativecommons.org/licenses/by/4.0/>