

БАГАТОВИМІРНЕ АНАЛІТИЧНЕ ОПРАЦЮВАННЯ ВЕЛИКИХ ДАНИХ

BIG DATA MULTIDIMENSIONAL ANALYTICAL PROCESSING

Актуальність використання аналітичного опрацювання великих за обсягом наборів та колекцій даних для новітніх фізико-технічних технологій полягає у необхідності швидкої та точної обробки експоненційно зростаючих обсягів даних, що генеруються високочутливими сенсорами, прискорювачами частинок та складними фізичними моделями, для оптимізації процесів та інновацій. Це дає змогу в режимі реального часу виявляти закономірності, прогнозувати поведінку систем, керувати міськими системами тощо. У контексті аналітичного опрацювання великих за обсягом наборів та колекцій даних багатовимірна аналітика великих даних відіграє провідну роль, здебільшого завдяки широкому переліку реальних застосувань, де її можна успішно застосувати. На даний момент часу, дослідники запропонували обширний перелік прикладних застосувань, зокрема [1]:

- великі за обсягом набори та колекції соціальних даних;
- великі за обсягом набори та колекції сенсорних даних;
- великі за обсягом набори та колекції графових даних;
- великі за обсягом набори та колекції енергетичних даних

Внутрішня природа багатовимірних абстракцій надає явну перевагу над класичними, наприклад, SQL-орієнтованими, підходами завдяки використанню розлогого класу аналітичних методологій. В результаті, декілька доменів даних можуть бути легко змодельовані в термінах вимірів, ієрархій, рівнів та мір, і на їхній основі можуть бути побудовані потужні аналітичні інструменти.

Традиційні системи оперативної аналітичної обробки (OLAP) та бізнес-аналітики (BI) помітно розширюють можливості SQL-орієнтованих інструментів аналітичного опрацювання даних. Оскільки добре відомо, що, як правило, набори даних у реальному житті проявляють багатовимірну та багаторівневу природу. Тому багатовимірні аналітичні інструменти більш придатні для видобування корисної інформації з великих за обсягом наборів та колекцій даних, ніж традиційні, безвимірні підходи.

Багатовимірна аналітика великих за обсягом наборів та колекцій даних, що зберігає приватність, є актуальним напрямком досліджень в галузі аналітики великих даних, що зберігає приватність, де фокус зосереджується навколо комбінації методологій багатовимірної аналітики великих даних та парадигм збереження приватності. Цей дослідницький контекст передбачає збереження приватності чутливого діапазону даних при одночасній підтримці завдань багатовимірної аналітики великих за обсягом наборів та колекцій даних. Цей напрямок досліджень активно стимулюється нещодавнім вибуховим розвитком сучасних хмарних інформаційно-технологічних платформ.

Література

1. Ohana, Ruben, et al. "The well: a large-scale collection of diverse physics simulations for machine learning." *Advances in Neural Information Processing Systems* 37 (2024): 44989-45037.