УДК 004.8:004.62

ІНТЕЛЕКТУАЛЬНІ АГЕНТИ ЯК ЗАСІБ АДАПТИВНОГО КЕРУВАННЯ КОНВЕЄРАМИ ДАНИХ У СЕРЕДОВИЩАХ BIG DATA

Лобода Ю. Г., к.пед.н, доцент [1], Кричун О. С.[2]
[1,2] Національний університет «Одеська юридична академія»
[1,2] Україна, Одеса
[1] loboda@onua.edu.ua; [2] aleksa.krychun@gmail.com

Анотація. *У статті розглянуто підхід до інтелектуалізації процесів обробки даних у середовищах Big Data на основі мультиагентних систем. Запропоновано використання автономних агентів для підвищення адаптивності, надійності та ефективності керування конвеєрами даних. Проаналізовано архітектурні рівні системи, визначено функції агентів і перспективи їх застосування в оркестраційних середовищах типу Apache Airflow. Підкреслено роль інтелектуальних агентів у забезпеченні самонавчання, реконфігурації та автономного прийняття рішень у динамічних середовищах обробки даних.*

Ключові слова: *інтелектуальні агенти; конвеєри даних; адаптивне керування; мультиагентні системи; великі дані.*

Вступна частина. Сучасні системи обробки великих даних характеризуються високою динамічністю потоків, різнорідністю джерел та необхідністю підтримки стабільної продуктивності при змінних навантаженнях. Традиційні підходи до керування конвеєрами даних ґрунтуються на статичних правилах і наперед визначених сценаріях, що знижує їхню здатність адаптуватися до непередбачуваних змін середовища. Класичні ETL-системи та оркестратори, заокрема Apache Airflow, використовують ациклічні графи з фіксованою топологією, які ускладнюють масштабування й інтеграцію нових джерел без ручного налаштування [1].

Інтелектуальні агенти, як автономні програмні сутності, здатні сприймати стан середовища, аналізувати його та приймати рішення на основі знань та досвіду, пропонують принципово новий підхід до організації адаптивного керування конвеєрами даних [2]. Завдяки використанню агентних технологій можливим стає децентралізоване прийняття рішень, прогнозування змін навантаження та оптимізація розподілу ресурсів у реальному часі.

Мета дослідження полягає у розробленні архітектури інтелектуальної агентної системи для адаптивного керування конвеєрами даних і дослідження механізмів автономної оптимізації на основі методів машинного навчання.

Основна частина. Конвеєри даних (Data Pipeline) розглядаються як багаторівнева система, що охоплює процеси збору, очищення, трансформації, збереження та аналізу інформації [3]. З позицій системного аналізу можна виокремити чотири основні рівні архітектури:

– Data Layer – джерела даних, сховища та транспортні протоколи (Kafka, S3);

– Processing Layer – обчислювальні фреймворки (Apache Spark, Apache Flink, Apache Beam);

– Orchestration Layer – інструменти керування виконанням і моніторингом процесів (Apache Airflow, Prefect, Dagster);

– Intelligent Control Layer – рівень, на якому реалізується агентна система адаптивного керування.

Інтелектуальні агенти на рівні Intelligent Control Layer забезпечують реалізацію принципів самоорганізації (self-organization) та самоадаптації (self-adaptation). Агент виступає як автономна програмна сутність, що сприймає стан середовища, аналізує його та приймає рішення на основі внутрішніх моделей або знань [4].

Типова мультиагентна система керування конвеєрами даних включає кілька агентів:

– Monitor Agent – відстежує стан вузлів і метрики продуктивності;

– Scheduler Agent – оптимізує чергу виконання завдань а допомогою евристичних методів або алгоритмів планування;

– Reconfiguration Agent – перебудовує структуру DAG при зміні джерел або відмовах вузлів;

– Analytics Agent – прогнозує перевантаження, відмови чи порушення SLA на основі аналізу історичних даних;

– Coordinator Agent – узгоджує взаємодію між іншими агентами через протокол Contract Net [5].

Такі агенти обмінюються повідомленнями та використовують спільну базу знань, що дає змогу зменшити навантаження на центральний оркестратор і забезпечити децентралізоване прийняття рішень. У разі збою одного з вузлів система здатна перебудувати граф виконання завдань і відновити процес без участі користувача.

Використання агентного підходу до керування конвеєрами даних сприяє зменшенню кількості ручних втручань, скороченню часу реакції на зміни в джерелах даних, підвищенню ефективності використання обчислювальних ресурсів і якості даних завдяки автоматизованому моніторингу та самокорекції.

Висновки. Інтелектуальні агенти формують основу для створення адаптивних систем керування конвеєрами даних, здатних до самонавчання, реконфігурації та автономного прийняття рішень у середовищах Big Data. Запропонований підхід сприяє підвищенню стійкості систем, зменшенню кількості ручних втручань і підвищенню ефективності обробки інформації. Подальші дослідження доцільно спрямувати на розроблення прототипу мультиагентної системи на базі Python-агентів та інтеграцію з хмарними оркестраційними платформами для практичного підтвердження ефективності підходу.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Chanda, D. (2024). Automated ETL Pipelines for Modern Data Warehousing: Architectures, Challenges, and Emerging Solutions. *The Eastasouth Journal of Information System and Computer Science*, 1(03), (pp. 209–212). https://doi.org/10.58812/esiscs.v1i03.523

[2] Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., ... & Gui, T. (2025). The rise and potential of large language model based agents: A survey. *Science China Information Sciences*, *68*(2), https://doi.org/10.48550/arXiv.2309.07864

[3] Munappy, A. R., Bosch, J., & Olsson, H. H. (2020). Data pipeline management in practice: Challenges and opportunities. In *International Conference on Product-Focused Software Process Improvement* (pp. 168-184). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-64148-1_11

[4] Cheng, Y., Zhang, C., Zhang, Z., Meng, X., Hong, S., Li, W., ... & He, X. (2024). Exploring large language model based intelligent agents: Definitions, methods, and prospects. *arXiv preprint.* https://doi.org/10.48550/arXiv.2401.03428

[5] Luzolo, P. H., Elrawashdeh Z., Tchappi I. (2024). Combining Multi-Agent Systems and Artificial Intelligence of Things: Technical challenges and gains. *Internet of Things*, Vol. 28. https://doi.org/10.1016/j.iot.2024.101364.

Loboda Yu.G., Krychun O.S.

Intelligent agents as a means of adaptive control of data pipelines in Big Data environments

Abstract. *The paper examines an approach to the intelligent management of data processing in Big Data environments based on multi-agent systems. The use of autonomous agents is proposed to enhance adaptability, reliability, and efficiency in data pipeline control. The article analyzes the main architectural layers, defines agent functions, and outlines prospects for applying such systems in orchestration tools like Apache Airflow. The study highlights the potential of intelligent agents for self-learning, reconfiguration, and autonomous decision-making in dynamic data environments.,*

Keywords: *intelligent agents; data pipelines; adaptive control; machine learning; stream processing; Big Data*