

DOI: 10.34185/1991-7848.itmm.2023.01.098

## **ПРОБЛЕМИ ТА РІШЕННЯ МАСШТАБУВАННЯ ПОТОКІВ ДАНИХ У ЕКОСИСТЕМІ «ІНДУСТРІЇ 4.0» В ПРОМИСЛОВІСТІ ТА БІЗНЕСІ**

Ситник Р.С., Гнатушенко Вік.В.

*Український державний університет науки і технологій, Україна*

«Індустрія 4.0» або «Промисловість 4.0» — це термін, який використовується для опису так званої «четвертої промислової революції», яка характеризується інтеграцією передових технологій, таких як штучний інтелект, Інтернет речей (IoT), хмарні обчислення та аналітика великих даних, у виробництво та інші галузі. «Індустрія 4.0» стимулює цифровізацію, автоматизацію та взаємозв'язок у промислових секторах, що призводить до появи розумних фабрик і об'єднаних ланцюжків поставок [1].

Таким чином, «Індустрія 4.0» характеризується імплементацією цифрових технологій у виробничі процеси, що призводить до значного збільшення обсягу даних, які генеруються та обмінюються між різними компонентами промислової екосистеми. Але також цей збільшений потік даних створює ряд проблем та викликів, які можуть не дозволяти промисловості та бізнесу повністю реалізувати переваги концепту «Індустрії 4.0». До таких викликів можна віднести:

1. Виклики зі зберігання та обробки великих обсягів даних від різних джерел.
2. Аналітика даних, отриманих від різних ланок промислової системи.
3. Безпека та надійність даних, які записуються та зчитуються з інформаційної системи виробництва.
4. Масштабування потоків даних та їх взаємодія з іншими компонентами та ланцюгами виробничих та логістичних процесів, які обслуговує інформаційна система.

Класичні інформаційні системи, які використовуються при реалізації принципів «Індустрії 4.0», для зберігання даних використовують реляційні та нереляційні бази даних, такі як MySQL та MongoDB. У MySQL дані зберігаються у таблицях, які складаються з рядків і стовпців. Кожна таблиця містить унікальний набір стовпців, які визначають тип даних, який може зберігатися в кожному стовпці. Для створення таблиць MySQL використовується мова структурованих запитів (SQL). У свою чергу, MongoDB – це нереляційна база даних, яка використовує документоорієнтований підхід для зберігання даних. На відміну від реляційних баз даних, де дані зберігаються у вигляді таблиць,

MongoDB дані зберігаються у вигляді документів, схожих на JSON. Кожен документ може мати свій набір полів та значень, і MongoDB може масштабуватись для зберігання величезних обсягів даних [2].

Доступ до цих даних отримується через головний сервер, на якому база даних хоститься. Сервер обробляє запити клієнтів, комп'ютерів, які хочуть отримати або змінити дані у базі даних, на отримання та запису даних. Такий підхід носить назву “клієнт-серверна архітектура” і є загальноживаним у індустрії. Але у великих інформаційних системах ланцюгів постачань та системах промисловості, де використовується постійний обмін даними між різними системами та їхніми клієнтами, для взаємодії у виробництві та логістиці, звичайна клієнт-серверна архітектура може виконувати свою роботу неефективно [3].

Для проєктування інформаційних систем для поліпшення управління потоками даних, які працюють за концепціями «Індустрії 4.0», в цій роботі пропонується розглянути використання технологію блокчейну, яка на даний час знайшла своє застосування у сфері електронних фінансів.

Блокчейн – це розподілена структура даних, яка реплікується та поширюється між членами мережі [4]. Технологічною основою, якою будується блокчейн, є Distributed Ledger Technology (DLT). Ця технологія є розподіленою базою даних, яка зберігає записи транзакцій у кількох вузлах мережі, а чи не в єдиній централізованій базі даних. DLT не має центрального апарату, який контролює базу даних, і всі вузли мають однаковий статус. Кожна транзакція записується в блок ланцюжка блоків (блокчейн) і містить хеш попереднього блоку, що забезпечує безперервність ланцюжка. Крім того, DLT забезпечує можливість створення смарт-контрактів, які автоматично виконуються під час виконання певних умов.

Одним із головних елементів DLT є хеш-функція, яка використовується для генерації унікального ідентифікатора для блоку даних. Хеш-функція може бути представлена такою формулою:

$$\overline{H(x)} = y, \quad (1)$$

де  $H(x)$  – хеш-функція, застосована до даних  $x$ , а  $y$  – результуюче хеш-значення.

Іншим основним елементом DLT є протокол консенсусу. Протоколи консенсусу використовуються в технології блокчейн, щоб гарантувати, що всі вузли в мережі погоджуються щодо стану блокчейну. Так, наприклад, блокчейн Ethereum з 2022 року використовує алгоритм Proof-of-Stake, в якому ймовірність того, що вузол буде обрано для створення нового блоку,

пропорційна сумі частки, яку він займає в мережі.

Формула для розрахунку ймовірності того, що вузол буде обраний творцем блоку:

$$P(\text{node}) = (S(\text{node}) / \sum S(i))^k, \quad (2)$$

де  $P(\text{node})$  – це ймовірність того, що вузол  $i$  буде обрано творцем блоку,  $S(\text{node})$  – це розмір частки, яку має вузол  $i$ , а  $\sum S(i)$  – це загальна сума частки, яку має мати мережа, змінна  $k$  є константою, яка визначає рівень випадковості в процесі вибору та зазвичай встановлюється протоколом мережі.

Для перевірки цілісності транзакцій та зберігання даних у блокчейні використовується структура даних Дерево Меркла. В ньому кожен блок містить в собі набір транзакцій, і дерево Меркла дозволяє швидко перевіряти, чи не були змінені транзакції в блоці.

Для побудови дерева Меркла кожен блок ділиться на кілька частин, наприклад, на транзакції, які входять в блок. Потім кожна частина хешується з уживанням криптографічної хеш-функції, і результати об'єднуються в пари. Кожна пара потім хешується знову, і цей процес повторюється до тих пір, поки не залишиться тільки один хеш, який називається кореневим хешом дерева Меркла.

Ми починаємо з хешування кожного окремого елемента даних у блоці (наприклад, транзакцій), щоб отримати їхнє хеш-значення. Потім ми об'єднуємо хеш-значення попарно і хешуємо їх разом, щоб отримати нові хеш-значення, які є корінням піддерев дерева Меркла. Цей процес повторюється до тих пір, поки всі хеш-значення не будуть об'єднані в одне хеш-значення, яке є коренем всього дерева Меркла.

Математично це може бути представлено такою формулою:

$$H = H_n(H_{n-1}(H_{n-2}(\dots(H_2(H_1(T_{x_1}, T_{x_2}))\dots)T_{x_{n-1}}), T_{x_n})), \quad (3)$$

де  $H$  – підсумкове хеш-значення кореня дерева Меркла,  $H_n$  – функція хешування,  $T_{x_n}$  – окремий елемент даних, які потрібно захистити.

Дерево Меркла має багато переваг у застосуванні у блокчейні та інформаційних системах, наприклад швидку перевірку цілісності даних та відсутність необхідності зберігати всі дані у кожному блоці. А для авторизації та додаткового захисту транзакцій в блокчейні, кожна транзакція може додатково захищена цифровим підписом.

В «Індустрії 4.0» дерево Меркла може бути використане для забезпечення цілісності даних у виробничих системах, таких як системи управління якістю або системи моніторингу виробничих процесів.

**Висновки.** Таким чином, дерево Меркла дає можливість ефективно захищати великі обсяги даних, забезпечуючи цілісність їх зберігання та передачі, що можна використати у інформаційних системах «Індустрії 4.0». Так, використовуючи алгоритми дерева Меркла, цифрового підпису, розподіленої програмованої бази даних та валідації транзакцій за допомогою алгоритмів консенсусу, тобто технологій та алгоритмів, на яких базується блокчейн та Distributed Ledger Technology, дає можливість створити інформаційні системи для промисловості та ланцюгів постачання з потенціалом для створення нових форм довіри та співпраці в контексті «Індустрії 4.0», надаючи безпечний і прозорий спосіб зберігання та обміну даними.

### Література

1. H. Lasi, P. Fettke, H.G. Kemper, T. Feld, M. Hoffmann. Industry 4.0. Business & information systems engineering, 6, pp.239-242, 2014
2. C. Győrödi, R. Győrödi, G. Pecherle, A. Olah. A comparative study: MongoDB vs. MySQL, 2015. 13th International Conference on Engineering of Modern Electric Systems (EMES) (pp. 1-6). IEEE.
3. H.S. Oluwatosin, Client-server model. IOSR Journal of Computer Engineering 16.1 (2014): 67-71.
4. Satoshi Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, 2008, [Електронний ресурс]. – Режим доступу: <https://bitcoin.org/bitcoin.pdf>

## CHALLENGES AND SOLUTIONS FOR SCALING DATA FLOWS IN THE INDUSTRY 4.0 ECOSYSTEM IN INDUSTRY AND BUSINESS

Sytnyk Roman, Hnatushenko Viktoriia

**Abstract.** The information technology integration processes into manufacturing and logistics are often described by the term "Industry 4.0". This term includes various approaches in digitalization, automation and interconnection in various industrial sectors. The integration of advanced technologies into manufacturing creates large flows of data that conventional industry information systems are often unable to process and respond to in a timely manner. That can create delays and errors that can lead to inefficiencies in supply chains, delays in production, and communication between industrial sectors in addition in problems of transparency and accountability of information systems in the "Industry 4.0" ecosystem, problems with security, data integrity, real-time process tracking, etc. Therefore, it is an important task to research new methods of managing data flows in information systems in the "Industry 4.0" ecosystem, which can improve interaction

and accelerate the adoption of information technologies in industry and logistics.

**Keywords:** industry, industry 4.0, data flows, information systems, blockchain, logistics, smart contracts, supply chains.

### Reference

1. H. Lasi, P. Fettke, H.G. Kemper, T. Feld, M. Hoffmann. Industry 4.0. Business & information systems engineering, 6, pp.239-242, 2014
2. C. Győrödi, R. Győrödi, G. Pecherle, A. Olah. A comparative study: MongoDB vs. MySQL, 2015. 13th International Conference on Engineering of Modern Electric Systems (EMES) (pp. 1-6). IEEE.
3. H.S. Oluwatosin, Client-server model. IOSR Journal of Computer Engineering 16.1 (2014): 67-71.
4. Satoshi Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, 2008, URL: <https://bitcoin.org/bitcoin.pdf>