Decentralized Industrial IoT Data Management Based on Blockchain and IPFS

Xiaochen Zheng, Jinzhi Lu, Shengjing Sun, Dimitris Kiritsis

Decentralized Industrial IoT Data Management Based on Blockchain and IPFS*

Xiaochen Zheng[1[0000−0003−1506−3314]], Jinzhi Lu[1[0000−0001−5044−2921]], Shengjing Sun[2[0000−0002−8790−7371]], and Dimitris Kiritsis[1[0000−0003−3660−9187]]

[1] Institute of Mechanical Engineering, EPFL, 1015 Lausanne, Switzerland
{xiaochen.zheng, jinzhi.lu, dimitris.kiritsis}@epfl.ch
[2] ETSII, Universidad Politécnica de Madrid, 28006 Madrid, Spain
shengjing.sun@alumnos.upm.es

Abstract. The wide application of Internet of Things (IoT) has fostered the development of Industry 4.0. In manufacturing domain, Industrial IoT (IIoT) are key components of the Factories of the Future (FoF). The big IIoT data are the foundation of implementing data-driven strategies. In current industrial practice, most of these IIoT data are wasted or fragmented in data silos due to security and privacy concerns. Novel data management approaches are required to replace traditional centralized data management systems. The rapid development of blockchain technologies provides a novel solution for this challenge leveraging its unique characteristics such as decentralization, immutability and traceability. However, blockchain is inefficient for exchanging big data due to transaction throughput limits. The peer-to-peer InterPlanetary File System (IPFS) provides a suitable complement for blockchain. Therefore, this paper aims to propose a decentralized IIoT data management approach based on blockchain and IPFS technology. The architecture and enabling technologies of the proposed system are introduced. A proof-of-concept implementation is realized and relevant experiments are conducted. The results demonstrated the feasibility of the proposed approach.

Keywords: Blockchain · IPFS · Data management · Industrial IoT.

1 Introduction

The advancements of digital technologies, such as Internet of Things (IoT), Artificial Intelligence (AI) and Cyber-Physical Systems (CPS) etc., are reshaping different sectors of the modern society. In industrial domain, the Industry 4.0 paradigm has been proposed empowered by Cyber-Physical Production Systems (CPPS), Industrial Internet of Things (IIoT) and big manufacturing data analytic etc. [1–3]. To realize the vision of Industry 4.0, all these enabling technologies and systems need to be integrated seamlessly. A smart manufacturing system

* Supported by EU Commission within the research projects QU4LITY (EU H2020 825030) Digital Reality in Zero Defect Manufacturing and FACTLOG (EU H2020 869951) Energy-aware Factory Analytics for Process Industries.

requires connectivity among various manufacturing units, facilities, machinery, suppliers and retailers as well as other manufacturing supporting industries, to form a valuable smart manufacturing network through the entire manufacturing value chain [4, 5]. This remains a challenging task for manufacturing enterprises due to the concerns about security, trust, traceability, reliability, and agreement automation within the manufacturing value chain [6, 7].

According to the Data-Information-Knowledge-Wisdom (DIKW) model [8], data are the basis of higher levels of intelligence. For the Factories of the Future (FoF), the heterogeneous data produced by the wide deployed IIoT devices are the foundation of all data-driven applications for smart manufacturing. However, in reality the storing and sharing of big IIoT data are still challenging tasks for most manufacturing companies. Overwhelming amount of these data are either discarded due to high storage and processing cost, or remain fragmented in isolated data bases [9].

One of the main obstacles hindering IIoT data sharing is the concerns about data privacy and security issues. Traditional centralized databases and data exchange protocols might be susceptible to various attacks and tampering risks [10]. The cost of data transferring is another limitation preventing IIoT data flow freely. Although, from a technological point of view, data sharing cost is relatively lower than before, in reality it is still expensive for manufacturing companies to transfer large volume of fine and granular IIoT data in real-time due to intermediary fees [11]. To cope with the above mentioned issues, advanced data management systems are required. Distributed Ledger Technologies (DLT), which have been developing rapidly in recent years, provide suitable solutions for this purpose. A distributed ledger is a distributed database, maintained by a consensus protocol run by nodes in a peer-to-peer network without any central administrator [12]. Popular DLT structures include blockchain and Directed Acyclic Graph (DAG) among others [13]. Blockchain was first successfully applied to cryptocurrency field, such as Bitcoin [14] and Ethereum [15], and has gained attention from both academia and industry owing to its unique features, such as decentralized control, high anonymity and distributed consensus mechanisms [16–18].

In recent years, blockchain technology has been widely adopted in manufacturing domain to enable smart manufacturing under the Industry 4.0 context [19]. Many Blockchain-based applications have been developed such as cloud manufacturing [20], manufacturing supply chain management [21], manufacturing processes tracing[22] etc. Some recent studies have investigated the application of blockchain in industrial IoT data exchange [23], knowledge management [24] and trading and energy trading in smart grid [25–27] among others. Despite of the such advantages, blockchain protocols are facing some challenges in terms of throughput limitations, high transaction fees and long approval time.

DAG-based protocols, such as IOTA [28], have been proposed as the new generation of DLT to solve the scalability and transaction cost limitations of blockchain. Theoretically they can reach unlimited throughput without transaction fees, making them suitable to transfer high frequency and large amount of data [29]. However, currently they are still in their early phases with limited throughput. The InterPlanetary File System (IPFS) [30] provides a suitable complement to DLT protocols for big data handling. IPFS is a content-addressing, peer-to-peer network for storing and sharing arbitrary data in a distributed file system. It uses a hash to access a stored file and the hash value changes when any changes are made in the file content. This paper aims to propose a decentralized data management system by combining IPFS and DAG-based DLT to address big IIoT data storing and sharing challenges.

The rest of the paper is organized as follows. The architecture of the proposed data management system and the key components are introduced in Section 2. A proof-of-concept prototype implementation and experiments are demonstrated in Section 3. The paper is concluded in Section 4.

2 System architecture

The architecture of the proposed IIoT data management system is as shown in Fig. 1. It includes four main components: IIoT data providers, IPFS data storage, blockchain-based meta data sharing platform, and IIoT data consumer.

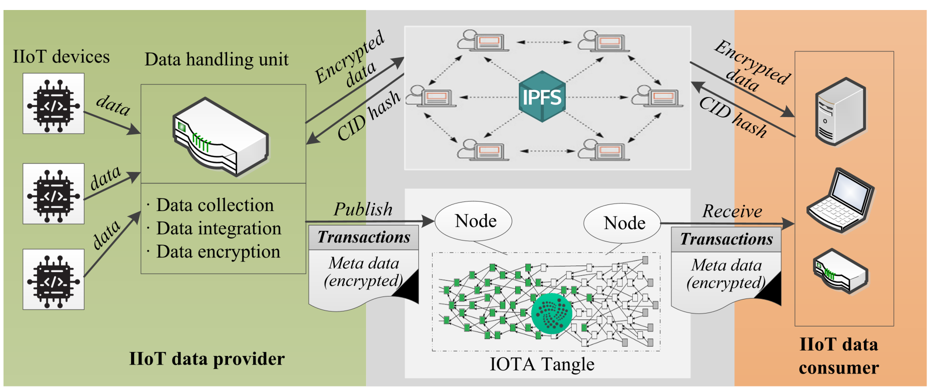

Fig. 1. Architecture of the proposed IIoT data management system based on IOTA Tangle and IPFS.

- IIoT data providers refer to the IIoT devices and relevant data handling units in a manufacturing system. The raw data generated by IIoT devices are transferred to a connected handling unit which can be a normal computer or any edge devices with certain computing capabilities like local servers and single-board computers. The handling units collect and integrate the raw data before uploading to IPFS in batches. Depending on the privacy by design, the data can be encrypted before uploading.

- IPFS data storage: Once the data are stored in the IPFS network, a cryptographic hash is generated as the Content Identifier (CID), which can be used to retrieve the uploaded data. After received the CID, the handling unit encode and encrypt the CID together with other meta data to create a transaction, to be transmitted to data sharing platform.
- Blockchain-based meta data sharing platform: The adopted protocol in this study is DAG structured IOTA Tangle, which is considered as the third generation blockchain. Although it is not exactly 'blockchain', we follow the common naming rules [13] to avoid confusing. The main reason of choosing this solution is that it solves the main limitations of previous blockchains, i.e. scalability and transaction fees. The transactions published on the Tangle can be searched or recommended to interested stakeholders. The content of a transaction might have different privacy levels depending on if it is encrypted or not. If encrypted, only receivers with authenticated decryption keys can consume the content.
- IIoT data consumers decode, and decrypt if required, the content of the transaction. Following the CID in the content, they can retrieve the data stored in IPFS. Another decryption might be required depending on the privacy level.

3 Proof-of-concept and experiments

A prototype of has been developed as proof-of-concept based on the IOTA Tangle API *(iota.lib.js)* and IPFS API *(ipfs Javascript implementation)*. The Javascript source codes with implementation instructions are open access on GitHub (https://github.com/zhengxiaochen/ipfs_iota_data_management.git).

3.1 Experiment and result

Experiments are conducted simulating the process of storing and sharing sensor data in JSON format collected from environment monitoring devices located in a factory. Details about data collection are introduced in a previous study [31]. As shown in Fig. 2, the raw data generated by different sensors were streamed to the data handling unit, where they were integrated into JSON format. Then the JSON data were encrypted according to a certain interval and uploaded to the IPFS network. Once stored successfully, a CID *(e.g. QmbFPdXRP8EuEHNuZuoqJ3vZeSzyE8TnKy7djLsrYf9rHE)* was created and return to the data handling unit. The uploaded content can be viewed on the IPFS network *(https://ipfs.io/ipfs/)* using the CID as shown in Fig. 2.

After received the CID, the handling unit created a transaction containing the CID and other meta data describing the shared data, such as data owner, data type, brief description and decryption keys to the IPFS content etc. The transactions can be published in different privacy modes [29]. For example, in Fig. 2 the transaction was published in public mode so that anyone who knows the transaction address can view the content of the transaction. In restricted

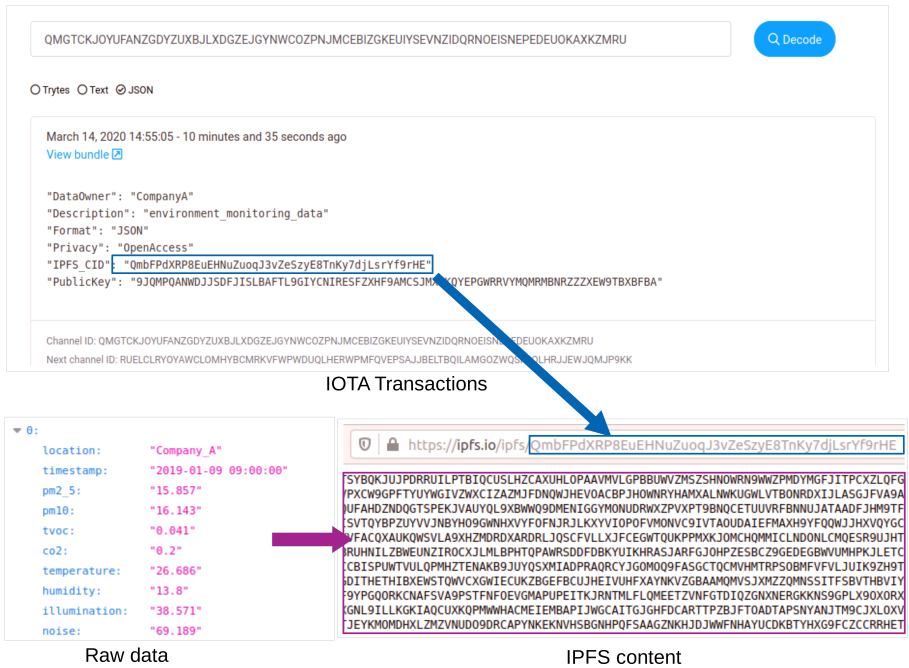

Fig. 2. Screenshots of the experiment for storing and sharing JSON sensor data using proposed approach.

mode, a extra key will be required to view the content, enabling the data publisher to flexibly control the access to the data.

Once the transaction is published on IOTA Tangle, interested data consumers can find it by proactive searching or passive subscription and recommendation. As shown in Fig. 2, the consumer followed the IFPS CID to retrieve the raw data from IPFS network. After downloaded the content, they might need to decrypt it with a specific key (the *"PublicKey"* in Fig. 2), which can be included in the transaction or distributed in separately.

3.2 Privacy and security analysis

The proposed data management system supports different privacy modes as demonstrated in the experiment. It allows nonsensitive data to be shared in public mode making them accessible to as many users as possible. For sensitive data, it provides two layers of encryption to protect the privacy and security. First, the data uploaded to IPFS are encrypted before uploading. It prevents data leakage even the IPFS network is compromised. Second, the transactions published to the Tangle can also be encrypted with the signature of the publisher and an extra key. Moreover, both IOTA protocol and IPFS are designed to be decentralized and distributed. This could eliminate the single-point-failure risks, and increase the tamper-resistance capability of the system. After a transaction is confirmed, it cannot be modified unless the majority of the computing power of entire network is controlled the the attacker. Any modification to the original content stored in IPFS will also produce a totally different CID because of the hash function.

4 Discussion

With the evolution of the Industry 4.0 and smart manufacturing, the needs of data integration and sharing from different sources is increasing. The proposed approach can be applied as an enabling tool for many aspects of advanced manufacturing. As an example, the proposed approach is been applied in a factory producing construction materials. This factory is facing difficulties in handling the data related to customer orders data from the Enterprise Resource Planning (ERP), Manufacturing Execution System (MES) data, the loading and delivery monitoring data etc. Due to privacy and security concerns, the manufacturer, logistic providers and customers are reluctant to share their data, which is essential for data integration and higher level data analysis. The proposed approach has been adopted in a more complex data management system to support data sharing among different stakeholders. By combining the proposed approach with semantic modelling and machine learning, advanced data handling solutions can be developed thus to empower the development of Industry 4.0.

5 Conclusions

This paper proposed a IIoT data storing and sharing approach by integrating advanced blockchain and IPFS technologies. A DAG structured IOTA Tangle was used to empower high scalability, low cost and tamper-resistance data sharing; and IPFS was adopted to handle the large volume data storage challenge. The proof-of-concept implementation and experiment results demonstrate the feasibility of the proposed approach. Such data management systems could accelerate IIoT data exchange among FoF manufacturing systems enabling higher level data-driven strategies for smart manufacturing.

References

1. L. Wang, M. Törngren, and M. Onori, “Current status and advancement of cyber-physical systems in manufacturing,” *Journal of Manufacturing Systems*, vol. 37, pp. 517–527, 2015.
2. J. Wang, Y. Ma, L. Zhang, R. X. Gao, and D. Wu, “Deep learning for smart manufacturing: Methods and applications,” *Journal of Manufacturing Systems*, vol. 48, pp. 144–156, 2018.
3. J. Ordieres-Meré, J. Villalba-Díez, and X. Zheng, “Challenges and opportunities for publishing iiot data in manufacturing as a service business,” *Procedia Manufacturing*, vol. 39, pp. 185–193, 2019.

4. J. Lee, B. Bagheri, and H.-A. Kao, "A cyber-physical systems architecture for industry 4.0-based manufacturing systems," *Manufacturing letters*, vol. 3, pp. 18–23, 2015.
5. Y. Lu, "Industry 4.0: A survey on technologies, applications and open research issues," *Journal of Industrial Information Integration*, vol. 6, pp. 1–10, 2017.
6. N. Mohamed and J. Al-Jaroodi, "Applying blockchain in industry 4.0 applications," in *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 0852–0858, IEEE, 2019.
7. K. Zhou, T. Liu, and L. Zhou, "Industry 4.0: Towards future industrial opportunities and challenges," in *2015 12th International conference on fuzzy systems and knowledge discovery (FSKD)*, pp. 2147–2152, IEEE, 2015.
8. M. Frické, "The knowledge pyramid: a critique of the dikw hierarchy," *Journal of information science*, vol. 35, no. 2, pp. 131–142, 2009.
9. Industrial-Internet-Consortium-(IIC), "Smart factory applications in discrete manufacturing," *Industrial Internet Consortium White Paper*, 2017.
10. F. Callegati, W. Cerroni, and M. Ramilli, "Man-in-the-middle attack to the https protocol," *IEEE Security & Privacy*, vol. 7, no. 1, pp. 78–81, 2009.
11. S. David, "Iota data marketplace," 2017.
12. J. Brogan, I. Baskaran, and N. Ramachandran, "Authenticating health activity data using distributed ledger technologies," *Computational and structural biotechnology journal*, vol. 16, pp. 257–266, 2018.
13. Y. Xiao, N. Zhang, W. Lou, and Y. T. Hou, "A survey of distributed consensus protocols for blockchain networks," *IEEE Communications Surveys & Tutorials*, 2020.
14. S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," tech. rep., Manubot, 2008.
15. G. Wood *et al.*, "Ethereum: A secure decentralised generalised transaction ledger," *Ethereum project yellow paper*, vol. 151, no. 2014, pp. 1–32, 2014.
16. R. Böhme, N. Christin, B. Edelman, and T. Moore, "Bitcoin: Economics, technology, and governance," *Journal of economic Perspectives*, vol. 29, no. 2, pp. 213–38, 2015.
17. M. A. Harlev, H. Sun Yin, K. C. Langenheldt, R. Mukkamala, and R. Vatrapu, "Breaking bad: De-anonymising entity types on the bitcoin blockchain using supervised machine learning," in *Proceedings of the 51st Hawaii International Conference on System Sciences*, 2018.
18. X. Zheng, R. R. Mukkamala, R. Vatrapu, and J. Ordieres-Mere, "Blockchain-based personal health data sharing system using cloud storage," in *2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pp. 1–6, IEEE, 2018.
19. J. Lee, M. Azamfar, and J. Singh, "A blockchain enabled cyber-physical system architecture for industry 4.0 manufacturing systems," *Manufacturing Letters*, vol. 20, pp. 34–39, 2019.
20. C. Yu, L. Zhang, W. Zhao, and S. Zhang, "A blockchain-based service composition architecture in cloud manufacturing," *International Journal of Computer Integrated Manufacturing*, pp. 1–15, 2019.
21. A. E. C. Mondragon, C. E. C. Mondragon, and E. S. Coronado, "Exploring the applicability of blockchain technology to enhance manufacturing supply chains in the composite materials industry," in *2018 IEEE International conference on applied system invention (ICASI)*, pp. 1300–1303, IEEE, 2018.

22. M. Westerkamp, F. Victor, and A. Küpper, "Tracing manufacturing processes using blockchain-based token compositions," *Digital Communications and Networks*, 2019.
23. C. H. Liu, Q. Lin, and S. Wen, "Blockchain-enabled data collection and sharing for industrial iot with deep reinforcement learning," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3516–3526, 2018.
24. X. Lin, J. Li, J. Wu, H. Liang, and W. Yang, "Making knowledge tradable in edge-ai enabled iot: A consortium blockchain-based efficient and incentive approach," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6367–6378, 2019.
25. Z. Li, J. Kang, R. Yu, D. Ye, Q. Deng, and Y. Zhang, "Consortium blockchain for secure energy trading in industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3690–3700, 2018.
26. S. Wang, A. F. Taha, J. Wang, K. Kvaternik, and A. Hahn, "Energy crowdsourcing and peer-to-peer energy trading in blockchain-enabled smart grids," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 8, pp. 1612–1623, 2019.
27. Z. Zhou, B. Wang, M. Dong, and K. Ota, "Secure and efficient vehicle-to-grid energy trading in cyber physical systems: Integration of blockchain and edge computing," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 1, pp. 43–57, 2019.
28. S. Popov, "The tangle," *White paper*, p. 131, 2016.
29. X. Zheng, S. Sun, R. R. Mukkamala, R. Vatrapu, and J. Ordieres-Meré, "Accelerating health data sharing: A solution based on the internet of things and distributed ledger technologies," *Journal of medical Internet research*, vol. 21, no. 6, p. e13583, 2019.
30. J. Benet, "Ipfs-content addressed, versioned," *P2P file system*, 2014.
31. S. Sun, X. Zheng, J. Villalba-Díez, and J. Ordieres-Meré, "Indoor air-quality data-monitoring system: Long-term monitoring benefits," *Sensors*, vol. 19, no. 19, p. 4157, 2019.