EAI/Springer Innovations in Communication and Computing

Luis Carlos Méndez-González
Luis Alberto Rodríguez-Picón
Iván Juan Carlos Pérez Olguín *Editors*

Innovation and Competitiveness in Industry 4.0 Based on Intelligent Systems

EAI
RESEARCH MEETS INNOVATION

Springer

Architecture for Initial States Algorithm for Blockchain Scalability in Local OnPrem IIoT Environments

Alfonso José Barroso-Barajas, Jesús Andrés Hernández-Gómez, Roberto Antonio Contreras-Masse, and Salvador A. Noriega-Morales

1 Introduction

Currently, many activities are supported by networked electronic devices and information systems, in the so-called Internet of Things, IoT, or Industrial Internet of Things IIoT. IoT is a network made up of physical electronic devices, integrated software, and connectivity, which allows them to connect to exchange data [4]. The main idea of IoT is to connect millions of devices capable of reacting to stimuli in their environment [16]. The goal of IoT is to enable devices to monitor their environment, communicate, and create a better environment for humans [14]. IoT is present in transportation, smart homes/buildings, public safety, environmental monitoring, and medicine, as well as in industrial processing, agriculture, and livestock, among others [10]. In this way, IoT contributes to working more intelligently, establishing a man-machine and machine-machine relationship.

In this sense, the accelerated growth of IoT devices and the benefits they offer mean that the usual data storage architectures present great technological, performance, and economic challenges. In an IoT network, thousands of transactions from thousands of devices must be verified for authenticity, causing network bottlenecks [1, 12].

On the other hand, blockchain technology is being implemented in various environments more frequently, due to the traceability and immutability of stored

A. J. Barroso-Barajas (✉) · J. A. Hernández-Gómez · S. A. Noriega-Morales
Department of Industrial Engineering and Manufacturing, Universidad Autónoma de Ciudad Juárez, Chihuahua, Mexico
e-mail: al206593@alumnos.uacj.mx; rhernand@uacj.mx; snoriega@uacj.mx

R. A. Contreras-Masse
Instituto Tecnológico de Ciudad Juárez, Chihuahua, Mexico
e-mail: rcontreras@itcj.edu.mx

L. C. Méndez-González et al. (eds.), *Innovation and Competitiveness in Industry 4.0 Based on Intelligent Systems*, EAI/Springer Innovations in Communication and Computing, https://doi.org/10.1007/978-3-031-29775-5_12

data. It can be said that blockchain is a decentralized database, organized in blocks of information, where each connected node stores a copy of the database [6]. The blocks are connected to form a chain, where the transactions are stored, recording the moment they occurred, within a network governed by rules agreed by the participants [8] highlighting its main characteristics such as decentralization, persistence, anonymity, and auditability [18].

Blockchain technology has often been questioned, due to the massive amount of information that is recorded and replicated on the network, thus causing the storage to eventually run out, threatening the stability of the system.

In the words of [7, 13], the storage of data in each blockchain block generates significant difficulties. If blockchain is used within IIoT devices, these difficulties can become a bottleneck for scalability, since a network of IIoT devices can generate a large amount of data in a few minutes or hours; therefore, the requirements of node storage will grow along with reported negative latency and throughput side effects [1, 13].

As a result of these problems, an algorithm is proposed to generate and manage initial states for blockchain. In this way, the blockchain theory and its concepts and variables are also enriched. On the other hand, companies that implement this technique will be able to take better advantage of the storage spaces of their servers, as far as blockchain storage is concerned. Finally, it is important to note that, as more information is stored, more electrical energy is needed to keep the equipment running and, in addition, electrical energy to maintain the cooling of the sites; therefore, by reducing the size of blockchain, these energy costs will decrease considerably, contributing positively against the ecological footprint.

This article is organized into sections. The following section explains the main concepts for the reader, in the form of background. Next, we discuss the methodology with the components required in our architecture. Then, in Sect. 4, the local architecture for IIoT is presented as a materialization of the proposed architecture. Finally, Sect. 5 presents our conclusions and future work.

2 Background

Blockchain, IIoT, and initial states are concepts that are worthy of review to grasp the same context in this paper.

2.1 Blockchain

Blockchain is similar to an accounting ledger, where movements or transactions are recorded in a list of data. Similarly, all transactions are digitally validated and signed, by all blockchain nodes. Data is added to each block, including the date and time the transaction was recorded. In turn, each block is linked to its predecessor,

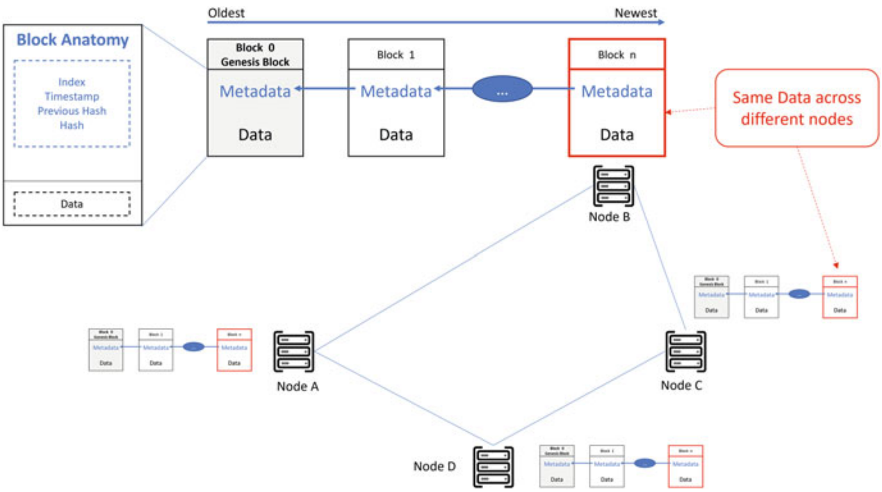


Fig. 1 Blockchain conceptual diagram and block anatomy

thanks to a set of reference data or key. The blockchain uses elliptic curve cryptography (ECC) and a SHA-2 hash scheme to provide strong cryptographic proof for authentication and data integrity [2]. The main idea of blockchain is to add blocks of information regardless of a central authority, which controls the system, with the added value of being a decentralized and encrypted data storage structure [11].

All system nodes share information and record transactions, applying a consensus algorithm to validate that the data to be stored is correct. Any attempt to destabilize the system must include simultaneous attacks on at least 50% plus one of the nodes in the blockchain to affect a single block.

Figure 1 shows the anatomy of a block in the blockchain. At a minimum, each block on the blockchain is made up of the following components: (1) current block index, (2) timestamp, (3) previous block hash in the blockchain, (4) current block hash, and (5) a record of transactions contained in this [9, 15] block.

The first block generated for blockchain is called "Genesis Block" or zero block. Therefore it is not possible to have a block $n + 1$ older than block zero. This in turn contributes to data consistency, since the timestamps of an n block will be greater than the $n - 1$ block and less than the $n + 1$ block. In this way, blockchain technology guarantees the integrity, non-repudiation, and security of the data due to the list of digitally signed transactions.

2.2 Industrial/Internet of Things (IIoT/IoT)

IIoT or IoT is one of the nine pillars of Industry 4.0 [5]. IIoT is a network made up of various devices that can be sensors and actuators managed by microcontrollers or household appliances such as a domestic coffee machine or even vehicles and complex machines. It is called the Internet of Things because these devices form an inter-network where they can send and receive information. However, very often these devices send information to the cloud, and, in that case, they do depend on the Internet.

IoT is a general concept. It can be made up of commonly used household devices. The "things" used in different industries and business solutions have different requirements. In this sense, IIoT is a more particular concept, limited to industrial activity, where specific protocols and proprietary networks known as Operational Technology or simply OT are used.

As the use of things grows and becomes more and more popular, the term industrial IoT is created to refer specifically to "things" that are primarily used in industry; however, today, other sectors such as healthcare, logistics, and commerce implement IIoT as well.

2.3 Initial State

As discussed above, each blockchain must have a genesis block. This block must be initialized considering that it has no predecessor and must be replicated in all the nodes in an identical way, that is, the genesis block must be exactly the same in all the nodes of the blockchain networks. As more blocks are added to the chain, the amount of data varies, but it is mandatory that all nodes have the same data stored, in order to carry out the consensus algorithm at the time of a [3] transaction.

On the other hand, when it comes to deploying in the cloud, allocating storage space is easier; in general, all hyperscalers have the ability to easily add blocks. However, what if the blockchain runs locally, on premises, where access to storage space is limited and restricted by physical media or availability time? This is where an algorithm to generate and manage initial states becomes relevant, since it can solve space and availability problems. In this sense, to address this problem in local environments, it is necessary to have an architecture to execute tests and later implement in production. The next section discusses the methodology and approach used to propose a local architecture that fits test and production environments, along with the main concept of a proposed algorithm to handle initial states.

3 Methodology and Materials

For a better understanding of these ideas, two issues must be addressed. The first is the construction of an architecture to implement and test the operation of the blockchain. The second is the algorithm for managing initial states and optimizing storage for the different blocks generated by IIoT devices.

3.1 Network Architecture

The proposed architecture basically comprises three elements: (1) IIoT devices, (2) servers to store the control logic and blockchain, and (3) NoSQL database to store metadata over time. These components must communicate as shown in Fig. 2.

The materials for this architecture are the following:

- IIoT components: Five virtual machines managed by Virtual Box version 6.1 running Raspberry Pi OS version 4.19, each with 2GB memory. These three virtual machines act as clients. Each client has previously been loaded with Python 3.10 programs to transmit the telemetry of CPU temperature and two random numbers in the range [1..100], respectively.
- Three virtual machines with operating system GNU/Linux Ubuntu 20.04, one 1 vCPU each and memory in each server is 4GB (these are referred also as servers). The file system will host the blockchain files. Here is also hosted the initial states algorithm. Each of those three virtual machines is acting as independent blockchain nodes. Each node has the following software:
  - Blockchain Manager. Developed in Python 3.10.2 with the responsibility of constantly reviewing the status of the blockchain in relation to the available capacity on the server.
  - Database MongoDB NoSQL version 5.0.3. MongoDB is a database that stores data in flexible structures called collections. In the database for the present experiment, only one collection is generated to store the data, such as the HASH IDs of the files, date, time, and the IIoT device ID data.

For this experiment, the IIoT device obtains data, either by the processor temperature or by generating a random number ranging from 1 to 100. This data is sent to the server. One of the servers is running a service developed in Python

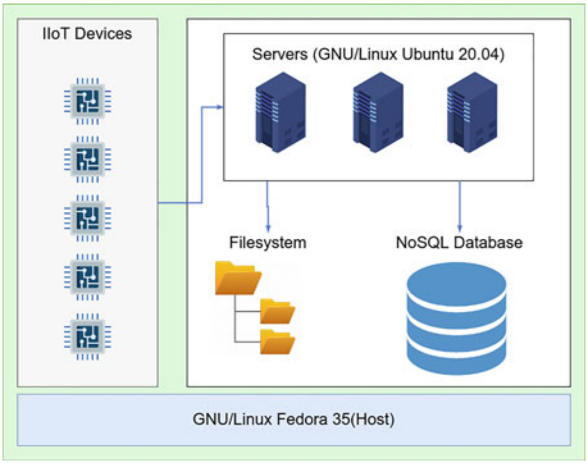


Fig. 2 Materials for this architecture

that, through the POST method, receives the data to process it. The received data is added to the transaction collection. Subsequently, the transactions are stored in the file system, adding an identification HASH and, in turn, as a reference to a previous file, forming the block chain. On the other hand, there is a MongoDB database, in which a backup record of the HASH identifiers of the files is kept, including the date, time, and identification data of the IIoT device. The data registry service has an authentication system in the MongoDB database. The application authenticates its access with the parameters of the connection string: host, port, database, username, and password. On the other hand, the database server has a user account configured with access privileges only to the database and the application's own collections. To carry out the experiment, a temperature sensor was simulated with an application developed in Python. The application runs continuously on the client, sending data every 0.5 s. The devices are too small to store a blockchain. It is important to note that one of the main characteristics of blockchain is decentralization. In this sense, it is decided to have, initially, two servers dedicated to store blockchain information in a decentralized way.

The literature suggests that the amount of data generated will grow considerably over time, and therefore, in a short time there will be numerous blocks in the blockchain and the storage in the file system will quickly become saturated. This is where it is important to develop an algorithm that allows managing initial states over time, in order to optimize storage space.

3.2 Initial States Algorithm

In our algorithm, we adopt the Initial Entries accounting concept. According to [17], an opening entry or initial states is a journal entry to preserve the balances of various assets, liabilities, and equity that appear on the balance sheet from the previous accounting period. Figure 3 depicts the operation of the initial states algorithm proposed.

The algorithm's logic works in three steps. First, transactions arrive and are recorded into a block. Second, it is registered in a new chain. Third, calculate available storage space, and validate thresholds to make a decision to truncate the chain, calculate hashes, and link new "genesis block" pointing to the previous historic blockchain to be available for queries, and do not lose information. The detailed algorithm is explained as follows:

- Start. Transactions are recorded in the blocks of the chain. IIoT devices sense their environment and send the information to the data server. The server runs an application developed in Python. The app receives data from the IIoT device. The data is recorded in a text file on the server to be evaluated.
- Registration of the new chain block.
- Calculation of available storage space on the server(SCL). The server has a secondary storage unit, of limited capacity. Blockchain appends fixed-size blocks

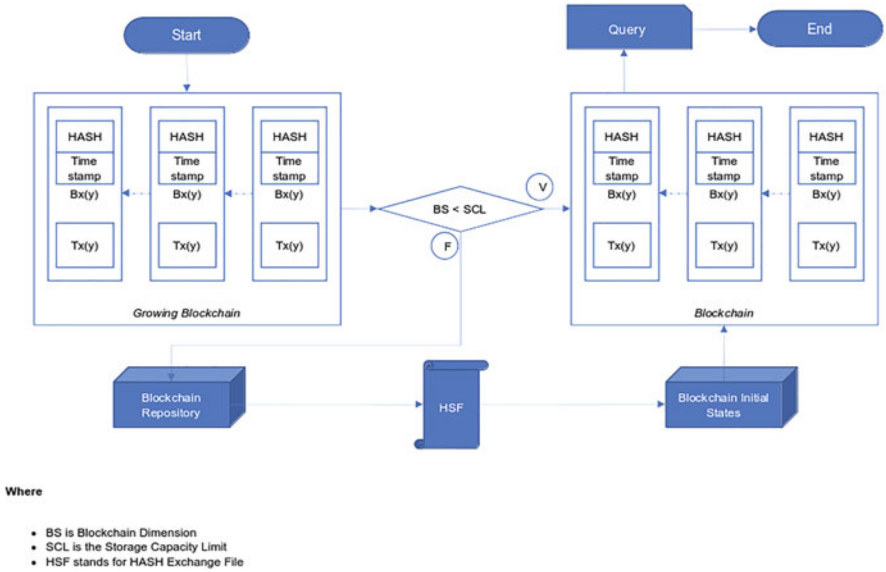


Fig. 3 Blockchain initial states algorithm

with the IIoT data(BS). The algorithm takes both parameters for comparison and decision-making.

- Initial or opening statements are a concept coined in accounting. In the case of the present algorithm, it has the function of generating an "exercise" closure and generating a new one, called the initial or opening state. The closure is executed based on the dimension of the blockchain with respect to the available storage capacity of the server. The algorithm will be constantly reviewing the growth of the blockchain and its relationship with the available space. Once the indicated value is reached, the algorithm will obtain the HASH of the last block. Subsequently, a "swap" file is generated where the HASH of the last block is stored with the HASH of the next block generated, to maintain the continuity of the chain. Finally, the "closing blockchain exercise" is packaged and compressed to be sent to a repository, cleaning the file system of the servers.
- If the remaining storage space is enough to store an additional block, it is added.
- If the storage space is not enough,
  - ∗ Calculate the hash of the last block. A hash key is calculated to allow for the continuity of the blockchain once it is stored in the repository.
  - ∗ Blockchain is compressed, packaged, and stored in the repository. The blockchain compressed package is pushed to the repository. In this way,

space is freed in the storage unit, so that the unit does not become saturated and avoids the risk of system collapse.

∗ A file is generated with the hash key of the last block (HSF). A new blockchain is started with its first block pointing to the hash previously stored in HSF.

Once this algorithm finishes, the new chain is available for new queries or ready to receive new transactions.

4 Results

For the performance evaluation and validation of the proposed algorithm, a virtualized small blockchain network has been created. This experiment was carried out on a GNU/Linux Fedora 35 machine, with an 8 × 11th generation Intel® Core™ i7-1165G7 processor at 2.80 GHz and 7.6 GiB of RAM. Three Raspberry Pi virtual terminals are executed on this physical equipment, as IIoT devices and two Ubuntu 20.04 virtual servers, for blockchain storage, connected to a network, using the TCP/IP protocol. IIoT devices run a client application that generates data based on processor temperature and random number generation. One of the servers receives the information and stores it in a MongoDB database.

Figure 4 shows the amount of data generation with a frequency of 0.5 s, with one device, then with two and finally with three IIoT devices. Throughout the experimentation period, it is observed that the data is generated, received, and stored without loss or latency.

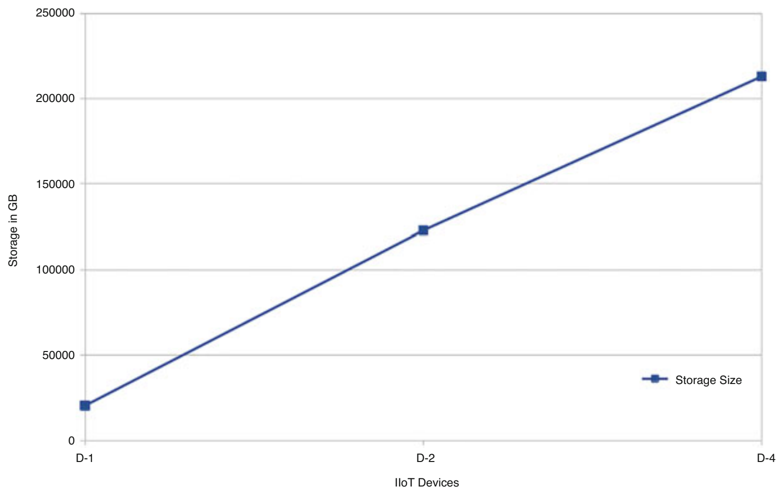


Fig. 4 Storage occupation

At this stage of development, blockchain compression was done manually using the .tar.gz format. Figure 5 shows that the directory occupies 3.1M of disk storage, while the generated compressed file occupies 124K, that is, 4.51% of the total space.

Data generation is done with a frequency of 0.1 s, from five virtualized IIoT terminals. Throughout the experimentation period, it is observed that the data is generated, received, and stored without loss or latency. Figure 6 shows initial state of the database, before system boot.

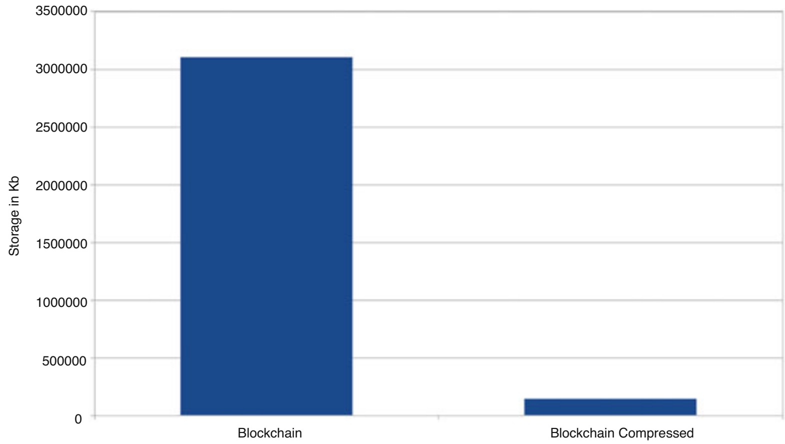


Fig. 5 Compressed blockchain partition

Fig. 6 Database initial state

Figure 7 shows the initial state of the directory where the blockchain structure is stored.

In a period of 5 min, 770 files were generated for the blockchain on the server, each containing 10 transactions. Figure 8 shows the used command to get the number of generated files.

Figure 9 shows the total occupancy.

The database in MongoDB, reached the following values, as seen in Fig. 10

At this stage of development, compression of the blockchain directory was done manually using the .tar.gz format. The directory occupies 3.1M of storage in disk, while the generated compressed file occupies 124K, that is, 4.51 of the total space, as shown in Fig. 11.

The experiment was repeated in 10 min, giving the following results

Figure 12 shows the difference between the regular storage for blockchain and, on the other hand, the storage required for the compressed blockchain structure.

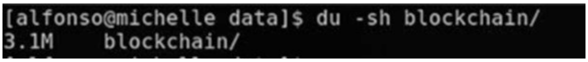
Fig. 7 Directory initial state

```
[alfonso@michelle data]$ du -sh blockchain/
0       blockchain/
[alfonso@michelle data]$ 
```

Fig. 8 Generated files

```
[alfonso@michelle data]$ ls blockchain/ | wc -l
770
```

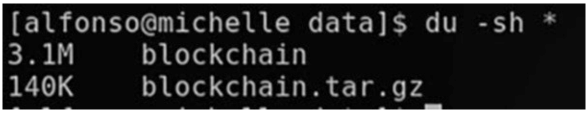
Fig. 9 Used space for blockchain into file system



```
[alfonso@michelle data]$ du -sh blockchain/
3.1M    blockchain/
```

Fig. 10 Database size at the start of the process

```
> db.stats()
{
        "db" : "iiot_db",
        "collections" : 1,
        "views" : 0,
        "objects" : 7229,
        "avgObjSize" : 158,
        "dataSize" : 1142182,
        "storageSize" : 401408,
        "indexes" : 1,
        "indexSize" : 258048,
        "totalSize" : 659456,
        "scaleFactor" : 1,
        "fsUsedSize" : 171877691392,
        "fsTotalSize" : 510405902336,
        "ok" : 1
}
>
```



```
[alfonso@michelle data]$ du -sh *
3.1M    blockchain
140K    blockchain.tar.gz
```

Fig. 11 Comparison of space used by blockchain after compression

```
> db.stats()
{
        "db" : "iiot_db",
        "collections" : 1,
        "views" : 0,
        "objects" : 14233,
        "avgObjSize" : 158,
        "dataSize" : 2248814,
        "storageSize" : 790528,
        "indexes" : 1,
        "indexSize" : 495616,
        "totalSize" : 1286144,
        "scaleFactor" : 1,
        "fsUsedSize" : 170770812928,
        "fsTotalSize" : 510405902336,
        "ok" : 1
}
>
```

Fig. 12 Database size at the end of the process

5 Conclusions and Future Work

The experiment, with the indicated materials, characteristics, and configurations, showed that, once compressed, the blockchain package occupies only 4.51% of the original space; this means that the application of the algorithm could effectively solve the storage problems. On the other hand, data generation and storage did not present any problem. It will proceed to increase both the amount of execution time and the frequency of the data, to look for a point where the service collapses. Future work is to determine what the optimal storage value is and what the optimal size of the blockchain is, so that the algorithm can decide when to generate the initial inputs, pack the original chain, and generate a new chain.

References

1. A. Abdelmaboud, A.I.A. Ahmed, M. Abaker, T.A.E. Eisa, H. Albasheer, S.A. Ghorashi, F.K. Karim, Blockchain for IoT applications: taxonomy, platforms, recent advances, challenges and future research directions. Electronics **11**(4), 630 (2022). https://doi.org/10.3390/electronics11040630
2. R. Almadhoun, M. Kadadha, M. Alhemeiri, M. Alshehhi, K. Salah, A user authentication scheme of IoT devices using blockchain-enabled fog nodes, in *Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA*, vol. 2018-Novem (2019), pp. 1–8. https://doi.org/10.1109/AICCSA.2018.8612856

3. F.M. Benčić, I.P. Žarko, Distributed ledger technology: blockchain compared to directed acyclic graph, in *Proceedings—International Conference on Distributed Computing Systems*, vol. 2018-July (2018), pp. 1569–1570. https://doi.org/10.1109/ICDCS.2018.00171
4. R. Contreras, A. Ochoa, E. Cossío, V. García, D. Oliva, R. Torres, Design and implementation of an IoT-based háptical interface implemented by Memetic Algorithms to improve competitiveness in an industry 4.0 model for the manufacturing sector. Adv. Intell. Syst. Comput. **939**, 103–117 (2019). https://doi.org/10.1007/978-3-030-16681-6_11
5. R. Contreras-Masse, A. Ochoa-Zezzatti, V. García, L. Pérez-Dominguez, M. Elizondo-Cortés, Implementing a novel use of multicriteria decision analysis to select IIoT platforms for smart manufacturing. Symmetry **12**(3) (2020). https://doi.org/10.3390/sym12030368
6. P. Danzi, M. Angjelichinoski, C. Stefanovic, P. Popovski, Distributed proportional-fairness control in microgrids via blockchain smart contracts, in *2017 IEEE International Conference on Smart Grid Communications, SmartGridComm 2017*, vol. 2018-Janua (2018), pp. 45–51. https://doi.org/10.1109/SmartGridComm.2017.8340713
7. A. Gorkhali, L. Li, A. Shrestha, Blockchain: a literature review. J. Manag. Anal. **7**(3), 321–343 (2020). https://doi.org/10.1080/23270012.2020.1801529
8. F. Gürrmann, B. Banusch, Fundamentals of Blockchain (October) (2018)
9. D. Khan, L.T. Jung, M.A. Hashmani, Systematic literature review of challenges in blockchain scalability. Appl. Sci. **11**, 9372 (2021); **11**, 9372 (2021). https://doi.org/10.3390/APP11209372
10. A. Khanna, S. Kaur, *Internet of Things (IoT), Applications and Challenges: A Comprehensive Review*, vol. 114. (Springer US, 2020). https://doi.org/10.1007/s11277-020-07446-4
11. B. Lashkari, P. Musilek, A comprehensive review of blockchain consensus mechanisms. IEEE Access **9**, 43620–43652 (2021). https://doi.org/10.1109/ACCESS.2021.3065880
12. K. Peng, M. Li, H. Huang, C. Wang, S. Wan, K.K.R. Choo, Security challenges and opportunities for smart contracts in Internet of Things: a survey. IEEE Internet of Things J. **8**(15), 12004–12020 (2021). https://doi.org/10.1109/JIOT.2021.3074544
13. A. Reyna, C. Martín, J. Chen, E. Soler, M. Díaz, On blockchain and its integration with IoT. Challenges and opportunities. Fut. Gener. Comput. Syst. **88**(2018), 173–190 (2018). https://doi.org/10.1016/j.future.2018.05.046
14. A.K. Singh, N. Firoz, A. Tripathi, K. Singh, P. Choudhary, P.C. Vashist, Internet of Things: from hype to reality (2020). https://doi.org/10.1016/b978-0-12-821326-1.00007-3
15. S. Singh, A.S. Sanwar Hosen, B. Yoon, Blockchain security attacks, challenges, and solutions for the future distributed IoT network. IEEE Access **9**, 13938–13959 (2021). https://doi.org/10.1109/ACCESS.2021.3051602
16. E. Sisinni, A. Saifullah, S. Han, U. Jennehag, M. Gidlund, Industrial internet of things: challenges, opportunities, and directions. IEEE Trans. Ind. Inform. **14**(11), 4724–4734 (2018). https://doi.org/10.1109/TII.2018.2852491
17. P. Tulsian, Financial Accounting, 2nd edn. (Tata McGrawHill Publishing, New Delhi, 2007)
18. Z. Zheng, S. Xie, H.N. Dai, X. Chen, H. Wang, Blockchain challenges and opportunities: a survey. EC Tax Review **28**(2), 83–89 (2019)