Blockchain and artificial intelligence for 5G-enabled Internet of Things: Challenges, opportunities, and solutions

Ashutosh Dhar Dwivedi¹ | Rajani Singh² | Keshav Kaushik³ | Raghava Rao Mukkamala² | Waleed S. Alnumay⁴

¹Section for Cyber Security, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark
²Centre for Business Data Analytics, Department of Digitalization, Copenhagen Business School, Frederiksberg, Denmark
³School of Computer Sciences, University of Petroleum and Energy Studies, Dehradun, India
⁴Computer Science Department, King Saud University, Riyadh, Saudi Arabia

Correspondence to: Ashutosh Dhar Dwivedi, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark. Email: adhdw@dtu.dk, ashudhar7@gmail.com

Abstract

Internet of Things (IoT) has revolutionized the digital world by connecting billions of electronic devices over the internet. IoT devices play an essential role in the modern era when conventional devices become more autonomous and smart. On the one hand, high-speed data transfer is a major issue where the 5G-enabled environment plays an important role. On the other hand, these IoT devices transfer the data by using protocols based on centralized architecture and may cause several security issues for the data. Merging artificial intelligence to 5G wireless systems solves several issues such as autonomous robots, self-driving vehicles, virtual reality, and engender security problems. Building trust among the network users without trusting third party authorities is the system’s primary concern. Blockchain emerged as a key technology based on a distributed ledger to maintain the network’s event logs. Blockchain provides a secure, decentralized, and trustless environment for IoT devices. However, integrating IoT and blockchain also has several challenges; for example, major challenge is low throughput. Currently, the ethereum blockchain network can process approximately 12 to 15 transactions per second, while IoT devices require relatively higher throughput. Therefore, blockchains are incapable of providing functionality for a 5G-enabled IoT based network. The limiting factor of throughput in the blockchain is their network. The slow propagation of transactions and blocks in the P2P network does not allow miners and verifiers to fastly mine and verify new blocks, respectively. Therefore, network scalability is the major issue of IoT based blockchains. In this work, we solved the network scalability issue using blockchain distributed network while to increase the throughput of blockchain, this article uses the Raft consensus algorithm. Another most important issue with IoT networks is privacy. Unfortunately, the blockchain distributed ledgers are public and sensitive information is available on the network for everyone are private, but in such cases, third party editing is not possible without revealing the original contents. To solve privacy issues, we used zkLedger as a solution that is based on zero knowledge-based cryptography.
1 INTRODUCTION

Communication network plays an important role and becomes the nervous system of today's digital era. The network needs to transfer a large amount of data at a much higher speed. Internet of Things (IoT) devices and its usage in the industry are growing exponentially. Millions of IoT devices are embedded in various applications such as smart homes, smart cities, airspace devices, and so on. Fifth-generation (5G) will play an important role in fully realizing IoT that connects people and computing resources, for example, sensors, vehicles, wearable devices, and so on. Sixth generation networking plays an important role in developing a network with a low latency network. Major IoT system nowadays uses centralized servers and storage databases, and the biggest issue with the centralized system is the lack of trust between entities involved and single point failure. To overcome such issues, decentralized architecture can be useful for peer to peer communication among network nodes. Nowadays, the most popular decentralized system is a blockchain that plays an important role in improving the trust between nodes in the network. To operate a distributed ledger called blockchain, network peers must provide the following functionalities: wallet service, storage, routing, and mining. The keys used to order transactions are provided by the wallet service. Storage is used to keep the copies of the chain in the node. Routing functionality is used for block and transaction propagation, while mining functionality is responsible for creating new blocks by solving the cryptographic puzzle for proof of work (PoW) mining schemes. Once a miner solves this complex cryptographic problem, it publishes the new block in the network. Peers of the network verifies this new block before adding it to the blockchain. However, several blockchain models exist with different implementation designs, and each has some pros and cons. Blockchain has many potential in the several field such as drone system, artificial intelligence (AI), fog computing, voting scheme, healthcare, fake news identification, forestalling pandemics, digital rights management system, and so on. However, the major concern with the integration of blockchain and IoT is scalability and throughput issues. The initial blockchain used for the Bitcoin network uses a PoW-based mining system with very low throughput and very high energy consumption and cannot be used for other applications. However, several other upgrades of blockchain allow high throughput but are mostly suitable for a small network. It is hard to scale them for a large network, and therefore a network consist of a large number of IoT devices is hard to implement. Another issue with blockchain is storage capacity that is deeply questioned. The chain is continuously growing, and in every 10 minutes, the chain grows its storage with 1 MB per block in Bitcoin. The copies of this chain stored in different nodes across the network. As the chain grows, these networks require more and more resources. Blockchain has mainly four important components that are as follows:

(a) Distributed ledger: Blockchain uses a distributed database and nodes in the network use the ledger’s replicated copy. Due to this distributed property, blockchain is immutable, and therefore information is much secured. The new blocks containing information are only added in the network when verified by the network’s major portion.

(b) Smart contract: Blockchain also uses the smart contract concept, which refers to a program or protocol that allows the contract to be automatically executed using some predefined conditions. Major blockchain systems now use a smart contract, and the pioneer blockchain system that used this concept was Ethereum. Hyperledger, a blockchain project that allows companies to implement their system, also uses smart contracts.

(c) Security: Data blocks are connected by using some hash function. Changing data into a single block will disturb the hash sequence of the chain. Therefore, it is practically infeasible to change the stored data.

(d) Consensus: Every blockchain has some consensus algorithm that allows nodes to agree with a certain decision. These consensus plays an important role to add new blocks in the network under certain rules and agreements.

IoT represents a network consist of various electrical and electronic devices that interact with each other by using some channel such as the internet. Several technologies, such as radio frequency identification, sensor networks, and near field communication, are used to connect the network. However, certain issues or limitations are available with these IoT devices and need to be addressed, such as:

(a) Latency: The IoT devices used nowadays are latency issues that refer to data transfer delay. In some cases, latency does not make much difference like command given to washing machine or thermostat, but in some cases, like an automatic car or satellite, this higher value of latency may cause serious problems.
## Comparison of existing literature with the proposed framework

<table>
<thead>
<tr>
<th>Author</th>
<th>Description</th>
<th>Year</th>
<th>Merits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skouby et al</td>
<td>Authors proposed a four-layer framework that combines smart cities, smart homes and Internet of Things devices.</td>
<td>2014</td>
<td>The framework uses smart technologies such as 5G and artificial intelligence (AI).</td>
</tr>
<tr>
<td>Marco et al</td>
<td>In this digital era, mobile networks are connected with everything such as cloud resources, sensors, vehicles and even robotics. The authors presented use cases and technologies of 5G networks.</td>
<td>2019</td>
<td>Authors discussed technologies that evolve wireless networks towards 5G networks. Paper summarizes the use cases, potentials and main challenges of enabling technologies.</td>
</tr>
<tr>
<td>Nguyen et al</td>
<td>Authors presented a state-of-art survey about blockchain for 5G and beyond networks. Authors explored opportunities of blockchain in 5G networks and presented extensive discussions based on literature papers.</td>
<td>2019</td>
<td>Paper also discussed several challenges of integrating IoT and blockchain and highlighted the motivation to integrate the 5G network with blockchain.</td>
</tr>
<tr>
<td>Liu et al</td>
<td>Authors proposed Ethereum blockchain-based secure sharing and data collection framework.</td>
<td>2019</td>
<td>The whole system can provide strong resistance against database attacks. The system has high security due to the use of blockchain functionality.</td>
</tr>
<tr>
<td>Shen et al</td>
<td>Authors proposed a privacy-preserving secure support vector machine (SVM) that is based on blockchain.</td>
<td>2019</td>
<td>The proposed system tackled the issues of data integrity and data privacy using secure SVM.</td>
</tr>
<tr>
<td>Wu et al</td>
<td>Authors developed an application-aware consensus management framework to enhance the flexibility between Internet of Things and blockchain.</td>
<td>2020</td>
<td>Current static consensus management cannot provide intelligent configuration capabilities, and the proposed system resolve this issue.</td>
</tr>
<tr>
<td>Hewa et al</td>
<td>Authors presented a paper about the role of blockchain in 6G where authors briefly describe challenges, opportunities, and research directions of integrating 6G enabled devices with blockchain.</td>
<td>2020</td>
<td>6G wireless networks driven by heterogeneous and enormous demands of hyperconnected existence of everything. This article canvassed and highlighted the blockchain role to mitigate some of the issues.</td>
</tr>
<tr>
<td>Alsharif et al</td>
<td>Authors presented an article that describes the vision, challenges and potential solutions of 6G wireless networks. More specifically, the paper explored the critical issues and key potential features, including research activities, key features and vision, challenges and specific solutions.</td>
<td>2020</td>
<td>The study shows key features and visions at forecasting 6G in the following dimensions: intelligence, energy efficiency, privacy, security, spectral efficiency, secrecy, affordability, and customization.</td>
</tr>
<tr>
<td>Mistry et al</td>
<td>In this article, authors presented a systematic review of 5G-enabled IoT for industrial automation. Authors presented an application of blockchain in the industry, smart city, healthcare and supply chain, and so on</td>
<td>2020</td>
<td>Paper presents several issues, challenges and solutions of 5G-enabled IoT. A comparison of existing proposals are also presented along with various parameters.</td>
</tr>
<tr>
<td>Qu et al</td>
<td>In this article, authors developed a D2C platform through the combined use of blockchain and federated learning for Industry 4.0.</td>
<td>2021</td>
<td>Federated learning addresses issues of efficiency and privacy. Blockchain facilitates poisoning-attack-proof functionality.</td>
</tr>
<tr>
<td>Gupta et al</td>
<td>In this article, authors presented a blockchain-based secure drone communication.</td>
<td>2021</td>
<td>The framework provide efficient drone communication system using smart technologies such as 5G and AI.</td>
</tr>
</tbody>
</table>
(b) Privacy: IoT devices produce a large amount of data transmitted through a channel and stored somewhere. To store such data, the user has to believe in third party providers, and therefore, data leaks at some point are possible.
(c) Security: IoT devices are resource-constrained, and they have very low computational power and memory. Some well standard encryption algorithms like AES do not fit with them, and therefore these devices required some lightweight encryption schemes.
(d) Storage: Millions of IoT devices produces a massive amount of data in real-time that is not easy to store somewhere. A single block in a blockchain only stores transactions, and the size of it is generally 1 to 2 MB, and therefore this issue is the main bottleneck in the integration of IoT devices with blockchain.

1.1 Contribution

The first part of the paper presents the main challenges of integrating blockchain and AI with 5G-enabled IoT devices. IoT devices produce billions of data in a second, while major blockchain protocols have slow throughput and not scalable to deal with these small IoT devices. This article solves the network scalability issue with the help of blockchain distributed network (BDN) and slow throughput issue with the Raft consensus algorithm’s help. The other issues, such as privacy and anonymity, can be solved using zero-knowledge proof (ZKP). We also presented several applications of such combined technologies in healthcare, supply chain, smart home, and so on.

1.2 Outline

Section 2 describes the preliminary of blockchain where we discussed the architecture of blockchain, types of blockchain and several consensus algorithms. Sections 3 and 4 describes the application of blockchain in 5G enabled industrial automation such as in healthcare, smart home, supply chain, industry, and E-voting. Section 5 gives detailed information about the challenges of integrating blockchain with IoT. These challenges are storage and throughput scalability, privacy and anonymity, network scalability, and so on. Section 6 gives the solution to these challenges mentioned before. These solutions are based on using BDN that improves the network scalability issues and using Raft consensus that provides high throughput. To remove privacy and anonymity issues, ZKP is used. Finally, Section 7 gives the conclusion of this whole work. Table 1 shows the review of related literature.

2 PRELIMINARY OF BLOCKCHAIN

Blockchain is distributed ledger technology (DLT) initially used for crypto-currency Bitcoin. It has attracted the whole world in the past few years due to its transparent and distributed behavior. Various nodes of the peer maintain a DLT to peer network. The essential features of blockchain are security, scalability, and decentralization. Though, to date, only two of them can be truly achieved by blockchain technology. The concept of achieving two out of three is called “The Blockchain Trilemma.” On the one hand, “Decentralization” is the most common word that is used in blockchain technology research, but on the other hand, this is also perhaps the most poorly defined word. Centralization and decentralization are mostly related to the levels of control of the system. In a centralized system, the system control is given to one entity, while in the decentralized system control is shared among various entities. When we talk about distribution, it refers to the location differences, that is, all parts of the system do not exist in the same place but have a different physical location (For all three types, see Figure 1). The bitcoin blockchain is decentralized as well as DLT.

The blockchain is organized into a chain of blocks that contains information. Several blocks can be added to blockchain but follow an append-only structure. Other data structures in a blockchain do not follow chain rules and represented in the form of a directed acyclic graph (DAG); however, they are not very popular. Nowadays blockchain is very popular in various other research areas such as: electronic voting, supply chain management, healthcare, digital right management system, and so on. Another important component of a blockchain is smart contract. Smart contracts are programmable applications used to perform a task based on specific terms and conditions automatically. These smart contracts are similar to traditional contracts, but instead of running by central authorities, they are programmed to perform the task itself in a decentralized framework. A general architecture of the whole blockchain system is defined in
Table 2: Public, private, and consortium blockchain

<table>
<thead>
<tr>
<th>Properties</th>
<th>Public blockchain</th>
<th>Private blockchain</th>
<th>Consortium blockchain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access</td>
<td>Public</td>
<td>Restricted</td>
<td>Restricted</td>
</tr>
<tr>
<td>Permissioned</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>System throughput</td>
<td>Slow</td>
<td>Fast</td>
<td>Fast</td>
</tr>
<tr>
<td>Consensus participants</td>
<td>All nodes</td>
<td>An organization</td>
<td>Multiple organization</td>
</tr>
</tbody>
</table>

Figure 2. Each box represents several options to choose based on the requirement. The blockchain system can be divided into three categories (also see Table 2):

(a) **Public blockchain**: Public blockchain are treated as truly decentralized blockchains where any node can mine or validate new blocks. Public blockchains termed as *permissionless* as any node can join the network and do not require any permission to validate new blocks. Bitcoin and Ethereum are few examples of public blockchains. These types of blockchains are designed to allow a large number of participants in the network. In such a type of blockchain, each transaction has a certain processing fee used as an incentive to the nodes who publish the block.

(b) **Private blockchain**: Private blockchains termed as *permissioned* and are not like public blockchain where anyone can join the network and are allowed to mine new blocks. These blockchains are suitable single organizations or enterprise solutions where blocks are published by only a few delegated nodes from the same network. Such blockchain does not require any type of processing fee or token to publish new blocks. Organizations may roll back the blockchain to any point in the past, and therefore these blockchains are not completely treated as decentralized. Ripple\(^{27}\) is a crypto-currency and treated as a private blockchain.

(c) **Consortium blockchain**: The consortium blockchain is a kind of hybrid blockchain with some private and some public blockchain property. These blockchains are also called the federated blockchain. Federated blockchains use *permissioned* network but involve multiple organizations. These blockchains do not pose any processing fee. some examples of consortium blockchains are: Hyperledger,\(^{28}\) Quorum,\(^{29}\) and so on.

The other important component of a blockchain is a consensus mechanism. This mechanism aims to achieve consensus in any blockchain network where the network participants do not trust each other, and the network does not have any central authorities. A brief introduction (also see Table 3) of a few important consensuses are as follows:

(a) **Proof of work** The first consensus protocol was PoW, a Permissionless blockchain, seen in Bitcoin. In the bitcoin network, any node can participate the network to publish the block. The nodes which create a new block are called
FIGURE 2  Blockchain architecture

TABLE 3  Comparison of few known blockchain systems

<table>
<thead>
<tr>
<th>Blockchain system</th>
<th>Consensus</th>
<th>Permissionless</th>
<th>Smart contract language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bitcoin(^{22})</td>
<td>PoW</td>
<td>Yes</td>
<td>C++, Golang</td>
</tr>
<tr>
<td>Ethereum(^9)</td>
<td>PoS</td>
<td>Yes</td>
<td>Serpent, Solidity</td>
</tr>
<tr>
<td>Hyperledger(^{10})</td>
<td>pBFT</td>
<td>No</td>
<td>Java, Golang</td>
</tr>
<tr>
<td>Ripple(^{34})</td>
<td>Ripple</td>
<td>No</td>
<td>C++, Golang</td>
</tr>
<tr>
<td>IOTA(^{35})</td>
<td>DAG</td>
<td>Yes</td>
<td>Java</td>
</tr>
<tr>
<td>ZCash(^{36})</td>
<td>PoW</td>
<td>Yes</td>
<td>C++</td>
</tr>
<tr>
<td>Litecoin(^{37})</td>
<td>PoW</td>
<td>Yes</td>
<td>C++, Golang</td>
</tr>
<tr>
<td>Quorum(^{29})</td>
<td>QuorumChain</td>
<td>No</td>
<td>Golang</td>
</tr>
</tbody>
</table>

Abbreviations: DAG, directed acyclic graph; pBFT, practical Byzantine fault tolerance; PoW, proof of work.

miners. These miners solve a cryptographic puzzle that has certain difficulties. To solve such puzzles, miners require a lot of computational power.\(^{30,31}\) Once a node solves the puzzle, it broadcast their block to the whole network. Other nodes in the network verify the block, and once the block is added to the blockchain, the miner gets the reward in terms of bitcoin and also get the transaction fee. However, the major issue with PoW based consensus is a high computational requirement by the nodes, and therefore energy consumption in such case is also very high.

(b) **Proof of stake** The proof of stake (PoS) protocol\(^{32}\) is also a Permissionless blockchain where a validator replaces the miners. Instead of solving puzzles, a validator directly adds a block to the network. Every validator should have a stake in the network and deposit an amount into the system. The validator is chosen in a pseudorandom fashion. However, a node that has more share in the network has more probability of becoming a validator. PoS protocol does not require high computational power to mine a new block, and therefore energy consumption is also very low. There is no reward in PoS based consensus, but only a transaction fee is given to the validator as an incentive.

(c) **Practical Byzantine fault tolerance (pBFT)** pBFT\(^{33}\) based consensus belongs to the permissioned blockchain system. There are always possibilities of malicious nodes in the network. The network’s ability to establish consensus when
malicious nodes send the wrong information to the network is called Byzantine fault tolerance (BFT). pBFT consensus provides 33% BFT, which means if less than 33% of the network nodes are malicious, in such case, the network performs well.

3 | BLOCKCHAIN USAGE IN 5G-ENABLED SMART INDUSTRIAL AUTOMATION

Blockchain has several industrial applications in the 5G-enabled network, which includes healthcare 5.0, autonomous vehicles, industry 5.0, supply chain management, e-voting, smart home, and so on. Blockchain and 5G together can be used to improve the overall performance and security parameters discussed in the above sections. The application of blockchain in these industrial networks is discussed as follows.

3.1 | Healthcare 5.0

The growth of the population worldwide demands more enhancement in healthcare technology. Remote health monitoring becomes more popular nowadays, which uses wireless healthcare applications. The use of the AI concept and high-speed data transmission and smart, intelligent devices set a benchmark in the Healthcare industry. Nowadays, the healthcare sector is adopting several advanced IoT devices to support remote patient monitoring. These devices can be divided into four types:

(a) Wearable health monitoring devices, for example, Fuelband, Fitbit, and so on.
(b) Medical wearable devices prescribed by doctors, for example, insulin pump.
(c) Medical embedded devices that can be implanted inside the body, for example, pacemakers
(d) Stationary medical devices can be used anywhere at a certain physical location, for example, chemotherapy dispensing stations for home-based healthcare.

Electronic health records (EHR) is a collection of patients’ information, while personal health record is a record of an individual patient. EHR allows real-time monitoring of patients by using shared patients medical data. Dwivedi et al.38 presented a paper about decentralized privacy-preserving healthcare system for the IoT devices that used blockchain as a distributed ledger for storing healthcare events. The authors proposed an overlay network based distributed network that uses cloud servers to store patient healthcare records, and the hash of the cloud data was stored on the blockchain. Any changes on cloud data can easily be detected as the hash of the data was stored on the blockchain, and therefore changing a single bit in the cloud will cause a different hash value of that data. The authors also proposed several lightweight cryptographic algorithms for the fast implementation of the network.

3.2 | Smart home

IoT privacy and security is a major challenge due to the massive amount of devices used nowadays. Dorri et al.39 presented a case study of a blockchain for IoT security and privacy used in smart homes. The authors proposed an overlay based network and eliminated the PoW and concept of coins. The network consists of three main components: smart home, cloud storage and overlay. An overlay network is also known as peer to peer network and used for distributed network architecture. To decrease data transmission and overhead delay, nodes are grouped into clusters, and each cluster elects a cluster head. The ideas in the paper were discussed using the smart home as a representative case study. The authors also presented the privacy and security analysis of the system. The simulation result shows that the method used has less overhead to the network and manageable for low resource IoT devices.

3.3 | Supply chain management

Singh et al.40 presented a supply chain management paper based on the combination of blockchain and the IoT to prevent counterfeit pharmaceuticals and to monitor the temperature of medicines throughout the chain. A pharmaceutical
supply chain system has mainly three components: manufacturer, wholesalers and pharmacies. The proposed IoT and blockchain-based supply chain also take care of the cold chain process of drugs. In general, a cold chain system also consists of three components: cold storage, cold transport and cold processing & distribution. During the distribution and transport of drugs, the cold chain has to ensure sanitary conditions. Therefore, it requires a certain kind of box that maintains the required temperature, and therefore a continuous monitoring process is required throughout the chain. The authors made the following contribution to this article:

(a) Authors proposed a scalable distributed network that uses high-speed servers to forward blocks quickly in the network. The blockchain system uses Raft consensus algorithm.
(b) The proposed framework uses Hyperledger. Hyperledger uses separate channels where the transaction ledger of one channel is hidden from other channels, and therefore many organizations can use the same system by keeping the privacy of their ledger.
(c) The proposed framework guarantee that drugs packet or sensor data have not been modified throughout the chain. Sensor and QR code used to enable temperature monitoring of packets.

3.4 Industry 5.0

In the present era, the business process and industry become completely automated. Massive improvement and development in technologies have resulted in the emergence of new approaches to production known as Industry 5.0. The industry aims to combine several technological domains such as cyber-physical system, blockchain, AI, and the IoT. Due to high competition in the industry, all business bodies aim to reduce costs and get more business advantages. To increase the profit in business, companies applying automated process in a business that helps in getting more advantage but also impose security risks. Business process management (BPM) system in industry 5.0 to automate and digitize the system plays an important role. Using blockchain with BPM also reduces the security risk and provide the following benefits:

(a) Accelerate transactions: Automated process settles the transaction instantaneously and do not require to wait for a long time.
(b) Cost reduction: After making the whole process automated and removing middlemen removes the overhead cost associated with intermediaries.
(c) Trust building: Distributed network to increase the trust among the parties involved in the business process.

3.5 E-Voting

A fraudulent election is one of the major issues in most of the countries and the largest democracies like the United States and India still suffers from a flawed electoral system. EVM machine hacking and vote-rigging and booth capturing are major challenges. Srivastava et al. proposed an e-voting model using blockchain. In this article, the authors use a DAG of blocks, also known as blockDAG as blockchain protocol that provides a secure and fast implementation. Authors break the system into two major parts:

(a) Authors use the PHANTOM protocol with large transaction throughput compared with another classical blockchain consensus, and it utilizes a DAG of blocks, aka blockDAG.
(b) To reduce the problem of booth capturing, authors considered Borda count method of vote counting that is based on ranked based voting.

4 BLOCKCHAIN FOR AI ENABLED 5G

On a very high level of abstraction, the blockchain is a systematic connection of blocks through cryptographic hashes that provides transparency, consistency, and reliability. Blockchain technology is one of the game-changer for 5G networks. 5G is a breakthrough mobile telecommunications innovation that promises to be 20 times faster than the 4G
technology of today. 5G’s latest features can be used to enable new business models and programs that include seamless connections between different stakeholders, including network carriers, companies, broadband suppliers, government providers, and so on regulators, and suppliers of facilities. Blockchain technology, however, has emerged as an empowering, destructive, disruptive technology that has begun to be implemented in several vertical industries of the industry. In a trusted, decentralized and protected way, Blockchain has been widely used to register, authenticate and verify properties and transactions, control communications, record data and manage identity amongst multiple parties. Figure 3 shows the scope of blockchain in AI-enabled 5G technology. Moreover, the brief descriptions about each subcategory are as below:

4.1 5G framework for commercializing

Crowdsourcing helps smaller players in networks to carry out telecommunications towers that can be part of the structure of the overall provider. By using their towers, such individual investors need to be licensed, approved, handled, and often immediately compensated. In providing a telecommunications service to a designated location, the scattered cellular sites can be independently or collectively operated by investors/operators. In a given geographical region, several coalitions may also exist, and a central authority will be needed to handle the dissemination of signatures for each coalition. A realistic approach for registering towers, handling used services, and automated fees, billing, and payment in crypto tokens can be provided through blockchain and smart contracts in a decentralized, trusting way while securing transparency and traceability at the same time.

4.2 5G framework trading

Framework trading in 5G is substantially important, which leads to sharing of telecom services by the mobile network operator (MNO). In this method, the mobile cellular towers or some part of these cellular towers are shared. It is classified into two categories: Active and Passive trading of 5G framework. Due to the introduction of network virtualization, active sharing is known to be the most effectively used strategy. On the other hand, passive sharing happens when an MNO, room, cooling share the cellular tower mast, and telecommunication rooms in separate buildings are reserved.

4.3 Distribution of 5G networks

In wireless networks, the bandwidth has become a scanty and very costly estate. Actually, providers pay spectrum regulators high fees. Usually, an operator purchases a subband or multiple subbands from a regulator. The provider either utilizes these subbands for its own purposes or rents them to other operators. The telecommunications operator will use its advantages and allow new small operators to offer 5G coverage without paying high license fees. It is also understood that main users, known as entrenched users, now fill most of the spectrum bands available. These bands include cable, digital radio, digital government systems, and satellites.
4.4 | Roaming of networks

Roaming is one of the most difficult problems in the telecommunications industry, as it requires brokers and third parties to negotiate payment and fee laws among them. It is clear that in 5G, multiple stakeholders would be interested in the use of 5G networks. Such stakeholders may include several carriers, international broker exchanges, and broker networks. Smart contracts incorporate the terms of the arrangement and the rationale of all parties and register, verify and control all their communications so that they can all be traced, recorded and audited by all parties in a cost-effective manner.

4.5 | Regulatory and compliance

The implementation of such a vast number of IoT products brings up the possibilities of providing innovative business models and facilities to revolutionary smartphone consumers. It is expected that 5G will handle these IoT devices by trusted centralized intermediary providers. Blockchain smart, autonomous storage contracts are more efficient and useful substitutes to those centralized providers under which they run. Control tasks may be carried out in a decentralized way, with outstanding trust, transparency, regulatory compliance, and automatic payment.

5 | CHALLENGES IN IOT-BLOCKCHAIN INTEGRATION

This section is dedicated to major challenges to be addressed when blockchain is integrated with IoT. Blockchain technology was initially made for the digital currency with a pioneering platform called Bitcoin. The initial version of blockchain was designed for the scenario where nodes in the network were powerful computers. However, the IoT reality is far away with this blockchain set-up, and the whole IoT network is full of resource-constrained devices. Some of the challenges are as follows:

(a) Storage and throughput scalability: The primary issue in blockchain-IoT integration is throughput and storage scalability. System throughput in Bitcoin is measured by the number of transactions per second (TPS). The average throughput of Bitcoin is approximately 2.5 to 3.5 TPS. Similarly, the transaction throughput of Ethereum is roughly 12 to 15 TPS. However, the average number of IoT devices installed is 20.4 billion by 2020. Therefore, the blockchain-based IoT systems require far better throughput to store only the events produced by IoT devices. Another challenge of such system is storage scalability. The blockchain was not designed to store large data, and the average block size of any blockchain is few megabytes, while IoT based systems can produce gigabytes (GBs) of data per second.

(b) Privacy and anonymity: Many IoT devices require privacy when transferring data on the blockchain network, for example, health-related information. In a remote patient monitoring scenario, the patient uses wearable IoT devices and transfer data to the health provider. The problem of data privacy for public blockchain is already known and discussed. Monero, a digital cryptocurrency uses a ring signature to protect user’s privacy. To preserve privacy in case of resource-constrained IoT devices, it is more difficult to use standard cryptographic algorithms.

(c) Consensus: Due to IoT’s limited resource nature, several popular consensus algorithms are not suitable for these devices, such as the PoW based algorithm requires miners with high computational power. However, there are several consensus proposals, but they are not standard one and still need to be tested for IoT based models. The major key challenge with resource-constrained IoT devices is mining. Off-chain could be one solution, but in such case, information has to move outside the blockchain to reduce latency. A high amount of energy consumption in PoW based blockchain is another issue that makes it unsuitable for IoT devices.

(d) Network scalability: To improve the throughput of blockchain, several consensus algorithms were proposed. However, the limiting factor of these consensus was their network. The blockchain’s added block should be verified by the network’s maximum number of nodes to get a truly distributed ledger. However, due to the slow propagation speed of the network, the verification of blocks takes time. If transactions and blocks are instantly propagated to the network, mining and verification of blocks can be performed quickly. To improve the network scalability, all nodes in the network should perform fast in propagating the blocks that are practically impossible for a public network.
6 | PROPOSED SOLUTIONS

An outline of the proposed framework can be given as follows:

(a) In the proposed framework, Raft consensus is used to improve throughput for IoT devices. Raft consensus is well suited for small organizations or networks. To apply the same consensus for a very large network, network scalability is an issue. However, to improve network scalability “bloXroute: A Scalable Trustless Blockchain Distribution Network” can be used.45

(b) A Consortium blockchain is used with Hyperledger Fabric operating system as a platform. A consortium blockchain is used by several organizations where one organization prefer to keep information private from other organization. Hyperledger uses separate channels for the individual organization.

(c) The framework uses a decentralized and distributed cloud storage network. Instead of storing data on the blockchain, the technology breaks the data into chunks called sharding and distributed to different cloud network storage peers. The blockchain only stores the event and hash of these data produced by IoT devices.

(d) User privacy is another issue in blockchain and IoT based systems. For the privacy issue, ZKP can be used. Zero-knowledge is a method by which one party can prove to another party that they know the fact without conveying any information apart from the fact.

The blockchain framework is divided into three important parts: distributed database network, overlay network, consensus, blockchain operating system or software, and ZKP for privacy.

6.1 | A scalable BDN

The Blockchain Trilemma is the major problem of any blockchain system. The trilemma involves three components: scalability, decentralization and security. These three components cannot be truly achieved at the same time. The term scalability determines the upper limit of how large a network can grow. Bitcoin22 provides strong security with approximately six million users, but the throughput and latency performance of bitcoin are poor. A blockchain’s latency is the time between submitting a transaction to the network and when the transaction becomes confirmed. Once the miner includes transactions in a block, other nodes in the network validate the block. To add a new block in a blockchain, any consensus requires communication between the nodes, and therefore, the network’s communication capacity is an important attribute. A peer to peer network generally consists of several computers with various computational speed. A network can have fast nodes as well as slow nodes. These slow nodes reduce the propagation speed of data in the network. bloXroute,45 a BDN is a global network to boost scalability (see Figure 4).

The bloXroute consist of two networks:

![Blockchain distribution network](image-url)
(a) bloXroute servers are low latency and high-speed servers designed to quickly propagate the blocks for multiple networks. Note that these servers do not control the P2P network but only used for the propagation of blocks to improve the network's latency.

(b) Peer to Peer networks uses bloXroute servers to propagate the blocks throughout the network and audit the server's behavior. Instead of propagating the blocks using peer to peer transmission, the blocks are propagated through servers, and therefore it reduces the network latency.

### 6.2 Distributed ledger software

Hyperledger Fabric established under Linux Foundation is an open-source permissioned DLT. The most important characteristic of Fabric is that it supports pluggable consensus protocols and any consensus can be used according to need. However, the current version of Fabric uses Raft consensus that is mainly suitable for single enterprises where only trusted nodes are added to the network. However, Raft is not byzantine fault-tolerant consensus. In situations, for example, public blockchain where nodes in the network are not trusted, byzantine fault-tolerant consensus like pBFT are much suitable. Another issue with any blockchain is its privacy issues. Hyperledger Fabric also provides a certain level of privacy through its channel architecture. Consider if a subnetwork does not want to share its information with other subnetworks, in such case, fabric uses channels for different subnetworks. The member of one channel cannot see the transaction details of other channels.

### 6.3 Consensus mechanism (Raft)

Consensus involves different peers or servers to agree on a particular decision on values. Once these peers or servers of the network reach a decision on these values, the decision is treated as a final decision. To make such a decision, most of the peers should be available; for example, if a network has five nodes and three of them agree on a certain value while the other two are not working, the network can make progress using these nodes. In the proposed framework, we suggest using Raft consensus in Hyperledger. For the smaller and private network, Raft is very suitable due to high throughput. However, to scale the network for a large number of nodes, bloXroute server can be used along with Raft or any other consensus. Raft consensus works by keeping replicated log in all nodes. This log follows the blockchain data structure where data structure can only be appended. The consensus is divided into three types of nodes: leader, follower, and candidate. However, these nodes change their role from time to time, and any node can become a leader based on voting. The write request is sent to the leader, and the leader distributes it to the follower nodes. Let us understand how a node becomes a leader.

#### 6.3.1 Leader election

To become a leader, a node's first vote sends a request to all other nodes and expects a response in a certain time (see Figure 5 from Raft). A certain time amount is given to nodes, say 150 ms, in which they have to respond, this time is called term. All node terms do not start simultaneously because if it happens, they all would also timeout at the same time and it will be harder for one node to collect the majority of votes as everyone might send the request to become the leader at the same time. If the majority of nodes give votes for it, then the node becomes the leader. The followers in the network expect a heartbeat from the leader, and if they do not receive this heartbeat for a certain time (called term), they assume that the leader is dead and start a new process for the selection of a leader. For a certain term, only one node can be a leader.

#### 6.3.2 Log replication

The main part of a consensus is keeping the replicated log. Once a node becomes leader, all the request is sent to it. If a follower receives a request, it can redirect it to the leader or return a message to the client by indicating which node is a leader. Once a leader receives a request, it initially updates the ledger and sends the request to all followers
to update. Though the message was appended to the log at this point, the leader cannot respond to the client until it does not get confirmation from the majority of the nodes. After getting a response from the majority, the message is treated as committed. The followers also expect the next heartbeat (empty message) from the leader to know that message is committed.

Figure 6 from Raft\textsuperscript{41} shows the logs and these logs composed of entries. Each entry consists of a term in which it was created and the state machine's command. Committed entries are only those for which the majority agrees.

### 6.4 Distributed database network

It is impossible to store GBs of data produced by IoT devices after every second. The only possibility to store the data in the cloud. However, to get the property of a distributed database, the data sharding concept has been used in the proposed framework. The working distributed database network is explained below:

(a) Data storage: The produced data is broken into multiple data chunks, and the process is called data sharding. Let's say the data is broken into \( n \) pieces and only \( k \) (less than \( n \)) parts are required to generate the original file. An
attacker would need the secret key for all the \( k \) files while these files are stored on \( k \) different and random storage nodes. A metadata is created for all the chunks that are sent to the different nodes. The concept is very similar to Shamir’s Secret Sharing Scheme developed by Adi Shamir in 1952, where a secret is divided into several parts and distributed each individual secret to a different participant. To construct the secret, a minimum number of parts is required.

(b) Data retrieval: By referencing the metadata, the client or user can easily identify the location of previously stored chunks. It is unnecessary to identify all the \( n \) chunks stored at different places, but a threshold is already defined, say \( k \), and original data can be retrieved by using \( k \) chunks out of \( n \) in the client computer system.

(c) Data maintenance: This technology helps to retrieve the data even when few storage nodes stop working or a few chunks become damage. The original data can be reconstructed with a minimum number of chunks required to construct the original message.

6.5 Solving privacy issues using ZKPs

The dark side of the IoT is its users’ privacy breach. Because of unawareness of data privacy, most IoT users even do not realize that they are gradually losing their privacy and so their data ownership. For example, if you click on some website, they generally have a long privacy policy that the user doesn’t care about and accept the terms and conditions. You are obliged to share your location, contact information, and other personal details by agreeing. On the other hand, these data can have monetary values and, therefore, benefit the companies or organizations who collect these data for their self-interest and later on sell them in the data market. For example, healthcare data generated using sensors or other wearable devices has great monetary value and the company that collects these data can sell it to other clients without notifying you.

Moreover, IoT devices and IoT networks are highly vulnerable to cyber-attacks as they are becoming the main target for cybercriminals. This result in the breach of IoT user data security as well as privacy. A new threat report from security firm Symantec reported the 600% rise in IoT attacks only in 1 year. In 2016, IoT attacks reported were 6000, however in 2017, IoT attacks reported were 50 000. Main IOT attacks (about 21%) are originated from China than the USA (about 11%), followed by Brazil (about 7%), and Russia (about 6%). Therefore, it is essential to mitigate the privacy issue in IoT devices by using the simple sharing concept with caring.

6.5.1 ZKP as a privacy solution

An effective way for privacy-preserving data sharing in 5G enabled networks is to share these network data using ZKP. ZKP is an old but very powerful method developed in the 1980s as a result of ground-breaking research by the researcher’s Goldwasser et al from MIT. A ZKP is a cryptographic mechanism allowing one party (prover) to prove to another party (verifier) that they possess knowledge of certain information without saying anything about the information. Moreover, such proofs are publicly verifiable, meaning that anyone can verify these proofs. For example, using ZKP, a customer can prove to the bank that he/she is old enough to borrow the loan without telling the actual age. Another example is Wher’s Waldo game, developed by Bethesda Softworks. In this puzzling game, one has to find the Waldo in the picture who is hidden somewhere in a massive crowd. Using ZKP, one can prove that he or she knows Waldo’s location without revealing his location’s exact coordinate. In a network, we can consider the sender as a prover while the receiver as a verifier.

Using the ZKP, a data owner can answer all the queries related to that data without providing the original data. This will help the data owner to monetize its own data. In addition, this data is always only with the data owner, he/she can sell the values generated by these data, such as Machine learning results several times to different customers. Any ZKP should have the following three properties:

(a) Completeness: It guarantees that if the statement is true, the prover will always be able to prove to a verifier that the statement is true and thus, the verifier has always to accept such proof. In the bank and customer example, if the customer’s age is above 18 and he is providing proof that he is above 18, the bank has to accept his proof.
(b) Soundness: It guarantees that if the statement is false, the prover will never be able to prove to a verifier that the statement is true and thus the verifier will always reject such proof. For example, nobody can identify himself as somebody else in a simple user authentication protocol.

(c) Zero-knowledge: It guarantees that the verifier will learn nothing from these proofs except the statement’s validity, meaning that he has zero knowledge about the statement or information. Banks do not have and can never guess the customer's original age in the bank and customer example.

Proving any statement using a ZKP is a three-stage process known as zero-knowledge protocol: namely, commit, challenge, and verify. If the prover knows the solution (which he/she wants to keep secret) of a given problem, he/she commits to this solution by using some mathematical function called commitment, and this process is called a commitment scheme or technique. It helps in preventing the malicious prover from cheating to the verifier as the prover cannot change the number later on without changing the commitment value. For example, in a sealed bid auction, all bidders put their bid (secret value) into a sealed envelope and give it to the auctioneer. This envelope has a secret bid inside, is similar to the commitment. By doing so, bidders cannot cheat by saying that they did something else as their bid is already there in the envelope at the time of the envelope opening.

After the commitment phase, the verifier sends a random challenge to the prover to solve the given challenge’s original problem. If the prover is honest and telling the truth, he/she will also be able to provide this solution.

At the final stage, the verifier verifies all the answers by using the commitment, and if all answers are correct, then he/she accepts the proof; otherwise, he rejects it.

Let's take a simple example of Schnorr’s digital signature protocol based on the discrete log assumption meaning that it is hard to find \( x \) such that \( A = g^x \) as long as the discrete problem is hard enough to be nonsolvable. Here, \( x \) is the private key and \( A \) is the public key.

The protocol allows the signer to prove that he knows the secret key or knowledge of \( x \) without revealing the key itself. The process is as follows:

1. Commitment: The prover or signer chooses a random scalar \( r \) and submits \( B = g^r \) to the verifier or receiver.
2. Challenge: The verifier reply back to the prover with a random challenge scalar say \( c \). The prover sends the value of \( s = r + xc \) to the verifier. Note that \( s \) is a point on a line passing through the random number chosen by prover and called secret key \( x \).
3. Verify: Since the protocol assume that the discrete problem is hard, therefore, the verifier will not be able to calculate the \( r \) but can check whether \( s \) was computed correctly because \( s = r + xc \) is equivalent to \( g^s = g^r \cdot g^{xc} \) which is equivalent to \( g^s = B \cdot A^c \). Now the verifier knows \( B, A \) and \( c \), hence he can easily verify if the signature is correct or not (Figure 7).

ZKP protocol can either be interactive or noninteractive. An interactive protocol is similar to, as explained above, with several message exchanges between the prover and verifier. However, the interactive protocol has two major drawbacks: first, if the verifier sends multiple challenges, it is quite probable that he may learn something about the problem; second, it requires a lot of communication and message exchange between the prover and verifier, which is time-consuming and costly. However, to mitigate these issues, a noninteractive protocol is being used mostly in real-life applications. In a non-interactive version of the protocol, the prover generates the challenge using the hash function instead of sending it by the
The idea of making the protocol noninteractive from interactive is developed by using the Fiat Shamir heuristic. When ZKPs are combined with blockchain, it opens up the gate to several opportunities for businesses/policymakers. Blockchain with ZKP has various applications in real-life problems. Some of them that are already being developed are listed below:

(a) Privacy-preserving smart contract, for example, Zether by Bünz et al.\textsuperscript{48}
(b) Privacy-preserving bank audit system, for example, ZkLedger by Narula et al.\textsuperscript{49}
(c) Confidential transactions and Cryptocurrency, for example, Confidential assets by Poelstra et al.\textsuperscript{50} Mimblewimble by Jedusor\textsuperscript{51} and Zerocash by Sasson et al.\textsuperscript{52}
(d) Privacy-preserving location tracking, for example, COVID-19 Contact Tracing App by Liu et al.\textsuperscript{53}

The ZKP mechanism has already been adopted and implemented in practice by some organizations. For example, ING Bank Introduced a Notary service based on ZKP, which enables the users to evaluate a transaction’s validity without revealing anything about it except that the transaction is valid. This is a blockchain breakthrough that substantially improves the privacy and security of transactions on Corda, an open-source blockchain platform. Applications of blockchain with ZKPs that are not being investigated are (see also Figure 8):

(a) Privacy-preserving supply chain: Zero knowledge-based location tracking and tracking systems will be very useful in supply chain systems. Moreover, It can also help the temperature based supply chain monitoring system by validating if the food or drug temperature lies in a predefined range throughout the process.
(b) Privacy-preserving data exchange or healthcare medical record: All medical records contain additional personal information; thus, a patient doesn’t feel comfortable to share the whole medical record. Using ZKPs, a patient can share only the required information from his/her medical records without revealing any extra or personal information. For example, let us consider that a healthy patient wants to buy insurance from an insurance company. To choose an insurance premium, insurance companies need to know about your health status; for example, they want to know whether a patient is diabetic or not. Using ZKPs, a patient can prove to the insurance company that he/she is healthy by providing a zero knowledge-based range proof that his/her sugar level lies in a normal range.
(c) A privacy-preserving financial system for organizations: ZKPs are mainly developed for cryptocurrencies and confidential transactions. Almost every organization has its financial system, which contains highly sensitive information such as beneficiary name, account details, salary amount, and so on. Using a zero knowledge-based financial system will improve the privacy and efficiency in the existing financial systems. These financial systems can also be auditable, which can help in fraud detection.
Moreover, ZKP is not limited to the applications mentioned above. A recent advancement of ZKP and blockchain is their applicability in the machine learning area. For example, researchers Froelicher et al.\textsuperscript{54} from the Laboratory for Data Security, EPFL proposed a Drynx system where no party or user trusts each other. Using Drynx, a querier can train machine-learning models and compute statistics on distributed datasets.

7 CONCLUSION

In the era of 5G, the integration of blockchain with the IoT plays an important role. This article provides various opportunities and industrial applications of 5G enabled IoT devices such as supply chain, e-voting, industry 5.0, smart home, and so on. The paper also provides the major challenges of integrating blockchain with IoT devices such as storage and throughput scalability, network scalability and privacy. The proposed framework solved network scalability issues using BDN and the slow throughput issue by using Raft consensus. We discussed all the available detailed solutions to overcome these challenges of integrating both technologies together. Another major issue of blockchain is privacy, and proposed framework uses a ledger (ZKLedger) based on ZKP to avoid privacy issues in blockchain and IoT. Zero-knowledge is one of the emerging areas in computer science, and this technology is still under investigation to answer several unsolved research problems.

ACKNOWLEDGMENTS

The work of Rajani Singh is funded by the Danish Ministry of Education and Science, Digital Pilot Hub and Skylab Digital. The work of Waleed S Alnumay is funded by the Researchers Supporting Project number (RSP-2020/250), King Saud University, Riyadh, Saudi Arabia.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Ashutosh Dhar Dwivedi \(\text{https://orcid.org/0000-0001-8010-6275}\)

REFERENCES


27. Ripple.


29. Quorum.


35. Popov Serguei. The tangle.


44. Cryptocurrency MONERO is a reasonably private digital currency. https://www.getmonero.org/.

How to cite this article: Dhar Dwivedi A, Singh R, Kaushik K, Rao Mukkamala R, Alnumay WS. Blockchain and artificial intelligence for 5G-enabled Internet of Things: Challenges, opportunities, and solutions. Trans Emerging Tel Tech. 2021;e4329. https://doi.org/10.1002/ett.4329