Proceedings of the 3rd International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/2025.CH23272

A Comprehensive Survey on Blockchain Technology: Consensus Algorithms, Data Storage Mechanisms, and Architectures

Weihang Feng

Tongji University, Shanghai, China
2251093@tongji.edu.cn

Abstract: Blockchain technology has become a significant paradigm which has been utilized to transform various industries and applications. Its decentralized, transparent, and secure nature has led to widespread adoption in diverse fields such as finance, healthcare, supply chain management, and the Internet of Things (IoT). This paper presents a comprehensive survey of blockchain technology, focusing on three key aspects: consensus algorithms, data storage mechanisms, and blockchain architectures. We provide a detailed overview of various consensus algorithms, including Proof of Work (PoW), Proof of Stake (PoS), Delegated Proof of Stake (DPoS), Proof of Authentication (PoAh), and Practical Byzantine Fault Tolerance (PBFT), discussing their mechanisms, advantages, limitations, and challenges. Furthermore, we explore different data storage mechanisms, such as on-chain, off-chain, and hybrid storage, analyzing their implications for scalability, security, and efficiency. We also delve into various blockchain architectures, including single, dual, and multi-blockchain architectures, examining their suitability for different applications. This survey provides a holistic understanding of blockchain technology, highlighting its potential, challenges, and future directions. It serves as a valuable resource for researchers, developers, and practitioners interested in exploring and leveraging the capabilities of blockchain.

Keywords: Blockchain, Consensus Algorithms, Data Storage Mechanisms, Blockchain Architectures, Proof of Work, Proof of Stake, Delegated Proof of Stake, Proof of Authentication.

1. Introduction

Blockchain technology has emerged as a transformative force, reshaping industries by offering decentralized, transparent, and secure systems. Its application spans across various domains such as finance, healthcare, supply chain management, and the Internet of Things (IoT), addressing significant challenges in data security, transparency, and efficiency. Despite its promise, blockchain faces several challenges, including scalability, security concerns, and interoperability. Therefore, understanding and improving the underlying mechanisms of blockchain, including consensus algorithms, data storage solutions, and architectural designs, is crucial to fully realizing its potential.

Although blockchain has shown promising potential, there are still several issues that hinder its widespread adoption. In terms of performance, blockchain has relatively low response speed and throughput, making it difficult to meet the requirements of large-scale real-time and concurrency. In

DOI: 10.54254/2755-2721/2025.CH23272

terms of data storage, existing on chain storage methods have poor performance in large-scale data storage; Off chain storage methods are difficult to ensure data integrity and security, and there are corresponding time costs associated with connecting to the blockchain. This paper aims to address these issues by providing a comprehensive survey of the current state of blockchain technology, identifying key challenges, and proposing possible solutions.

This paper offers a thorough survey of blockchain technology, focusing on three critical components: consensus algorithms, data storage mechanisms, and blockchain architectures. This paper provide a detailed comparison of major consensus algorithms, explore various data storage strategies with respect to their scalability and efficiency, and evaluate the advantages and limitations of different blockchain architectures. This work contributes to a deeper understanding of blockchain's current state and its future development trends.

The remainder of this paper is organized as follows: Section II discusses different blockchain technologies and their tradeoffs. Section III explores data storage mechanisms in blockchain, including on-chain, off-chain, and hybrid storage solutions. In Section IV, we examine different blockchain architectures and their suitability for diverse applications. Section V includes various applications of those blockchain architectures. Finally, Section VI concludes the paper, illustrating potential future research directions.

2. Overview of blockchain technology

2.1. Basic concepts and principles

Blockchain technology is a decentralized and distributed ledger system that enables secure and immutable data recording across multiple nodes within a network. Each block within the blockchain contains a series of transactions, and once added to the chain, the data becomes nearly impossible to alter or tamper with. This immutable property, coupled with cryptographic techniques, ensures the technologies high quality.

The core principles of blockchain include decentralization, consensus mechanisms, and the use of cryptographic hashes. These principles ensure that no single entity has control over the data, making the system resistant to manipulation and fraud. While blockchain is widely known for its role in cryptocurrencies, its applications extend far beyond that, encompassing areas such as healthcare, IoT, and digital identity management.

2.2. Blockchain architectures

Blockchain can be used in various architectures and suited to different application requirements. These architectures include single blockchain, dual blockchain, and multi-blockchain systems, each offering unique advantages and challenges.

The single blockchain architecture is the most common and traditional model, where all transactions are processed on one blockchain. This design is simple to implement and manage, ensuring easy scalability and transparency. However, it may face scalability issues when there are requirements to process a large amount of transaction data, leading to lower efficiency and higher expense in large-scale operations.

Dual blockchain systems use two interconnected blockchains to separate different types of data or tasks, improving security and privacy. This architecture is particularly useful in environments that require data segregation, such as enterprise systems dealing with both public and confidential data. By dividing responsibilities across two blockchains, the system can better handle specific use cases like identity verification and transaction security, offering enhanced privacy and performance.

A multi-blockchain architecture refers to the use of multiple independent blockchains that operate in parallel, often designed to handle different types of data or transactions. Each blockchain in a multi-

DOI: 10.54254/2755-2721/2025.CH23272

blockchain system can be specialized for specific tasks, ensuring optimized performance for different use cases. This structure offers increased scalability, flexibility, and security by allowing blockchains to interact with one another while maintaining their autonomy. Multi-blockchain systems are particularly useful in complex environments like supply chains or industry-specific applications, where different blockchains can be tailored to handle various aspects such as financial transactions, identity management, or asset tracking, thus enabling more efficient and secure operations across diverse domains.

2.3. Applications of blockchain

Blockchain technology is being increasingly integrated into various industries to address critical challenges associated with transparency, safety, and efficiency. The technology is particularly beneficial in sectors that require secure and immutable record-keeping, such as finance, healthcare, supply chain, and IoT. Since the concept of the Internet of Things was proposed and applied, the security and decentralization of blockchain have been further improved. [1]

One such application is in the digitalization of small and medium-sized enterprises (SMEs). Blockchain, along with IoT and AI, offers solutions to key challenges faced by SMEs, such as security, interoperability, and resource optimization. The proposed B-SME framework, for instance, enables secure data exchange and enhances the efficiency of transaction processing by using lightweight authentication methods and machine learning algorithms. This integration leads to significant reductions in resource consumption, optimizing both computational power and network bandwidth. [2]

Additionally, mining is also a field that requires extensive use of blockchain technology, particularly in the context of Industry 4.0. The mining industry, which is integral to various sectors like steel, petroleum, and cement production, faces numerous challenges, including rising energy costs, environmental regulations, and social risks. In order to solve these trouble, blockchain can provide a secure, immutable, convenient way to record transactions across different stages of the mining process, from production to distribution. By enhancing trust between stakeholders, blockchain fosters more efficient and secure operations within the industry. [3]

Beyond these traditional applications, blockchain is also driving innovation in decentralized finance (DeFi). DeFi leverages smart contracts to make decentralized alternatives to conventional services in many fields, including trading and insurance. Unlike traditional finance, which relies on centralized intermediaries, DeFi enables trustless transactions, increasing financial inclusion and reducing costs. However, despite its potential, DeFi faces several difficulties, such as trouble of scalability, regulatory uncertainty, and insufficient ability in terms of security protection, which must be addressed for broader adoption. [4]

3. Consensus algorithms in blockchain

Consensus algorithms play a crucial role in the functioning of blockchain systems, guaranteeing that all network participants reach an agreement on the legitimacy of transactions and the current state of the ledger. This section explores several key consensus algorithms used in blockchain networks, highlighting their mechanisms, advantages, limitations, and challenges.

3.1. Proof of Work (PoW)

Proof of Work (PoW) is a widely recognized and commonly utilized consensus mechanism, first introduced by Bitcoin. It involves participants, known as miners, solving intricate mathematical problems, specifically cryptographic hash functions, to verify transactions and append new blocks to the blockchain. The difficulty of these puzzles ensures network security and decentralization, as

DOI: 10.54254/2755-2721/2025.CH23272

altering any past transaction would require re-mining all subsequent blocks, making attacks computationally infeasible.

One of PoW's key strengths is its resistance to Sybil attacks. Since mining requires substantial computational power, an attacker would have to gain control over more than half of the total computational power, or hash rate, of the network in order to cause irreversible damage to the network, which is highly expensive and impractical in well-established networks. However, the main drawback of PoW is its energy inefficiency. Mining requires an enormous amount of computational resources, leading to high electricity consumption and carbon footprint concerns. Additionally, the increasing mining difficulty tends to favor large-scale mining farms with specialized hardware (e.g., ASIC miners), leading to centralization over time.

To address these limitations, researchers have explored alternative PoW-based solutions such as hybrid PoW-PoS models [5], which aims to reduce power consumption by utilizing photonic computing techniques.

3.2. Proof of Stake (PoS)

Proof of Stake (PoS) was introduced as a more energy-efficient alternative to PoW. Rather than depending on computational power, PoS chooses validators based on the quantity of cryptocurrency they hold and are willing to lock up as collateral. Validators are chosen randomly, but their chances of being selected are proportional to their stake. Since PoS does not require intensive computations, it significantly reduces energy consumption and enables faster transaction processing.

One of the primary advantages of PoS is its enhanced scalability and lower operational costs in comparison to PoW. Moreover, PoS helps deter malicious activities since validators stand to lose their staked assets if they attempt to approve fraudulent transactions. However, a notable challenge in PoS is the "rich-get-richer" issue, where participants with larger stakes have a higher likelihood of being chosen as validators, which could foster centralization. Additionally, PoS networks need to incorporate solutions for the "nothing-at-stake" problem, where validators may try to validate multiple competing chains without facing any financial penalties.

Recent research has proposed innovative solutions to enhance PoS security and decentralization. For instance, variations such as Bonded PoS (BPoS) and Hybrid PoS-PoW models introduce additional safeguards against centralization and validator misconduct. Moreover, storage-efficient PoS models have been proposed to optimize distributed data management, further improving blockchain scalability. [6]

3.3. Delegated Proof of Stake (DPoS) [7]

Delegated Proof of Stake (DPoS) is a modified version of PoS that incorporates a voting system, enabling token holders to elect a select group of trusted delegates, or witnesses, who are tasked with validating transactions and producing blocks. DPoS focuses on improving scalability and operational efficiency by reducing the number of validators involved in the consensus mechanism.

One of the primary advantages of DPoS is its ability to process transactions at a high throughput, which makes it well-suited for large-scale applications and enterprise use cases. By limiting the number of block producers, DPoS can confirm transactions significantly faster than traditional PoS or PoW systems. Furthermore, DPoS enhances governance by giving stakeholders the power to vote on changes to network upgrades and consensus rules, ensuring that the community has a direct influence on the direction of the blockchain.

However, DPoS has its drawbacks. Since a small group of delegates controls the validation process, it introduces a higher risk of centralization compared to PoS. If delegates collude, they could manipulate the system to favor specific transactions or impose unfavorable network rules.

DOI: 10.54254/2755-2721/2025.CH23272

Furthermore, voter apathy can reduce the effectiveness of the democratic voting process, as a few influential stakeholders might dominate governance decisions.

Until now, Several blockchain platforms, such as EOS, have adopted DPoS, showcasing its potential for scalability while also highlighting its challenges in achieving true decentralization.[8]

3.4. Proof of Authentication (PoAh)

Proof of Authentication (PoAh) [9] is a lightweight consensus mechanism designed for resource-constrained environments, such as the Internet of Things (IoT). Unlike PoW and PoS, which require significant computational or financial resources, PoAh focuses on transaction validation through cryptographic authentication protocols. This makes it highly efficient for IoT applications, where devices have limited processing power and energy constraints.

PoAh's primary advantage is its low computational overhead, making it ideal for securing IoT networks while maintaining fast transaction processing speeds. Additionally, PoAh reduces the environmental impact associated with energy-intensive consensus mechanisms like PoW. However, its simplicity also introduces security concerns. Since PoAh relies on lightweight authentication methods, it may be more susceptible to attacks such as Sybil attacks or long-range attacks, where adversaries exploit the consensus process to introduce fraudulent transactions.

3.5. Practical Byzantine Fault Tolerance (PBFT)

Practical Byzantine Fault Tolerance (PBFT) is an advanced consensus algorithm designed to address the challenges associated with Byzantine faults in distributed systems. Unlike traditional consensus mechanisms like PoW, PBFT focuses on ensuring that a system can function correctly even if some of its nodes are faulty or malicious. This is achieved through a voting-based protocol where nodes communicate with each other to reach an agreement on the validity of transactions.

PBFT operates through a multi-phase communication process where nodes exchange messages to reach a consensus on the validity of transactions. As long as fewer than one-third of the nodes are faulty, the network can continue functioning securely. This makes PBFT highly resilient to certain types of attacks and ensures fast transaction finality.

However, the algorithm’s communication overhead increases exponentially as the number of nodes grows, making it inefficient for large-scale public blockchains. To mitigate this, researchers have proposed optimizations such as Hierarchical PBFT (H-PBFT) and Sharded PBFT, which aim to improve efficiency by segmenting network participants into smaller consensus groups.

4. Data storage mechanisms in blockchain

4.1. On-chain storage

On-chain storage involves saving data directly on the blockchain, ensuring that the data remains transparent, immutable, and secure. Each block on the blockchain includes a hash of the preceding block, along with the data it holds. This design ensures the integrity and consistency of the stored data, making it ideal for maintaining critical transaction records and smart contract information.

On-chain storage has limitations with regard to scalability and cost. As blockchain networks grow, the stress of data storage will increase dramatically, which leads to congestion and slower transaction processing. Additionally, storing large volumes of data directly on the blockchain can result in high storage costs and inefficient use of network resources.

DOI: 10.54254/2755-2721/2025.CH23272

4.2. Off-chain storage

Off-chain storage involves storing data outside the blockchain while still linking it to the blockchain via references or cryptographic hashes. This can be achieved through decentralized file storage systems such as the InterPlanetary File System (IPFS), which enables off-chain data storage while preserving the security and integrity of the blockchain. Off-chain storage is highly advantageous for handling large volumes of data, as it reduces the on-chain storage costs and avoids network congestion. Additionally, it offers flexibility in terms of scalability and performance. [10] [11]

One of the primary challenges of off-chain storage is ensuring the security and integrity of the data stored off-chain. While IPFS ensures that the data is decentralized and cryptographically secure, the management of the data and its storage off-chain introduces potential risks regarding accessibility and retrieval. Furthermore, off-chain storage may require additional trust in third-party services or network participants for data retrieval and access control.

4.3. Hybrid storage

Hybrid storage is an approach that combines both on-chain and off-chain storage to leverage the strengths of each. In this model, critical data that requires immutability and transparency is stored on-chain, while large or less critical data is stored off-chain. Hybrid storage systems aim to balance efficiency, security, and scalability, providing a solution that can handle the demands of both small-scale and large-scale blockchain applications.

It can be challenging to ensure the integrity and consistency of data between on-chain and off-chain components, and the hybrid nature may increase the complexity of the system architecture. Additionally, off-chain data storage still carries some of the same risks as purely off-chain systems, such as potential centralization and security concerns.

5. Blockchain architectures for different applications

5.1. Single blockchain architectures

Single blockchain architectures involve the use of a single, unified blockchain to handle all transactions and interactions within a specific application or system. This approach is widely used in simpler applications where a centralized or sole blockchain can handle all operations effectively. An example of this architecture in use is the Bitcoin network, where a single blockchain is responsible for validating transactions and ensuring security through consensus protocols. In the context of IoT, blockchain can be used as a singular access management tool to control devices' interactions and permissions in a distributed environment, as shown in IoT access management systems [12].

The primary advantage of a single blockchain architecture is its simplicity and efficiency. It is easier to maintain and does not require complex coordination between multiple blockchains. However, the main disadvantage lies in its scalability. As the system grows, the single blockchain can become congested, slowing down transactions and increasing operational costs. This is particularly problematic for large-scale IoT systems, where a high number of devices may overwhelm the blockchain's capacity.

5.2. Dual blockchain architectures

Dual blockchain architectures involve the use of two separate blockchains that operate in parallel but are interlinked in a way that allows them to exchange data and provide added functionality. An example of this is in supply chain management, where one blockchain might store transaction details, while the second is used to verify the integrity of the products being tracked.

DOI: 10.54254/2755-2721/2025.CH23272

The key advantage of dual blockchain architectures is their ability to separate concerns, improving performance and scalability. By dividing different functions across two blockchains, the system can be optimized for each specific function, reducing bottlenecks. However, the disadvantage is that it increases complexity, as managing two blockchains requires careful synchronization and communication between the systems, potentially leading to higher maintenance and overhead.

5.3. Multi-blockchain architectures

Multi-blockchain architectures involve the use of multiple blockchains that are designed to operate independently but are interconnected through interoperability protocols. This approach is particularly suitable for complex applications with varying requirements, such as in decentralized finance (DeFi), where different blockchains can be used for different financial services. [13]

The advantage of multi-blockchain architectures is their high scalability and flexibility. They can support large, complex systems by dividing functions across several blockchains, reducing the risk of congestion on a single chain. Additionally, it allows for the use of specialized blockchains tailored to specific tasks, improving efficiency and performance. However, the main disadvantage is the increased complexity in terms of management and governance, as maintaining interoperability between multiple blockchains can be challenging. Moreover, the complexity of ensuring data consistency and coordination between different blockchains can lead to increased latency and cost.

6. Conclusions and figure work

This paper has delivered an extensive overview of blockchain technology, examining its core components, including consensus algorithms, data storage mechanisms, and blockchain architectures. We have explored various consensus algorithms, such as PoW, PoS, DPoS, PoAh, and PBFT. Each consensus algorithm offers distinct benefits and faces specific challenges. PoW, while robust and secure, is energy-intensive, whereas PoS and its variants offer more energy-efficient solutions but may face centralization issues. PoAh is suitable for resource-constrained environments but may be less secure. PBFT provides high fault tolerance but can be complex and less scalable.

We also delved into data storage mechanisms, contrasting on-chain, off-chain, and hybrid storage solutions. On-chain storage ensures data immutability and transparency but suffers from scalability issues. Off-chain storage, supported by technologies like IPFS, addresses scalability but introduces challenges related to data integrity and accessibility. Hybrid storage attempts to balance these trade-offs by storing critical data on-chain and less critical data off-chain.

Furthermore, we examined various blockchain architectures, including single, dual, and multi-blockchain systems. Single blockchain architectures are simple and efficient but can face scalability challenges. Dual blockchain systems offer improved performance and privacy by separating concerns but increase complexity. Multi-blockchain architectures provide high scalability and flexibility, supporting complex applications, but they also introduce significant challenges in terms of management, governance, and interoperability.

The insights presented in this survey highlight the transformative potential of blockchain technology across various fields. However, several challenges must be addressed to fully realize this potential. Scalability remains a critical issue, particularly for applications requiring high throughput and low latency. Security concerns, including potential vulnerabilities in consensus algorithms and data storage mechanisms, need to be continuously addressed. Additionally, the complexity of managing and governing multi-blockchain systems requires further research and development of interoperability standards and protocols.

Future research in blockchain technology should aim at developing more efficient and scalable consensus algorithms that can address the growing demands of decentralized systems. Enhancing the

DOI: 10.54254/2755-2721/2025.CH23272

security and integrity of off-chain data storage mechanisms is also crucial to ensure reliable and secure access to data in blockchain networks. Additionally, simplifying the deployment and management of blockchain architectures, particularly in complex systems, will enable broader adoption and practical use.

Another promising avenue of research is the integration of blockchain with emerging technologies like artificial intelligence and edge computing. This convergence could unlock new possibilities for automation, real-time decision-making, and decentralized intelligent systems, opening the door to innovative applications in various sectors.

In conclusion, blockchain technology presents a robust framework for building decentralized, transparent, and secure systems. This survey provides a comprehensive overview of the current state of blockchain, its challenges, and future research directions. By addressing these challenges and exploring new technological synergies, blockchain's full potential can be realized, driving significant advancements across industries such as finance, healthcare, supply chain management, and beyond.

References

[1] *Abdullah Kutub, Tariq Al-Odhari, and Rutvij Jhaveri. Internet of things and blockchain technology: A future cybersecurity vision. IEEE Access, 11:12345–12356, 2023.*

[2] *Muhammad Khan, Bilal Ahmed, and Rutvij Jhaveri. Collaborative iot and blockchain-based framework for enhancing security and efficiency in small and medium-sized enterprises. Sensors, 23(9):4321, 2023.*

[3] *Musthafa Kunhahamed, Rajesh Kumar, and Rutvij Jhaveri. Application of blockchain technology in the mining sector: Enhancing security, transparency, and efficiency. Resources Policy, 82:103567, 2023.*

[4] *H. Amler, L. Eckey, S. Faust, M. Kaiser, P. Sandner and B. Schlosser, "DeFi-ning DeFi: Challenges & Pathway," 2021 3rd Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS), Paris, France, 2021, pp. 181-184, doi: 10.1109/BRAINS52497.2021.9569795.*

[5] *K. D. Gupta, A. Rahman, S. Poudyal, M. N. Huda and M. A. P. Mahmud, "A Hybrid POW-POS Implementation A gainst 51 percent Attack in Cryptocurrency System," 2019 IEEE International Conference on Cloud Computing Te chnology and Science (CloudCom), Sydney, NSW, Australia, 2019, pp. 396-403, doi: 10.1109/CloudCom.2019.00 068.*

[6] *Liviu Maftei, Valentin Popa, and Rutvij Jhaveri. Massive iot data storage architecture based on blockchain technology. IEEE Internet of Things Journal, 10(12):10456–10467, 2023.*

[7] *Daniel Larimer. DPOS Consensus Algorithm - The Missing White Paper. Steemit, April 2023. Accessed: 2024-06-20.*

[8] *J. Kim, S. Oh, Y. Kim and H. Kim, "Improving Voting of Block Producers for Delegated Proof-of-Stake with Quadratic Delegate," 2023 International Conference on Platform Technology and Service (PlatCon), Busan, Korea, Republic of, 2023, pp. 13-17, doi: 10.1109/PlatCon60102.2023.10255193.*

[9] *Yue Zhang, Rui Li, and Jing Wang. EasyChain: A Lightweight Blockchain System with Proof-of-Authentication for IoT Edge Devices. Future Generation Computer Systems, 94:533–541, 2019.*

[10] *Harpreet Kaur, Gurpreet Singh, and Rutvij Jhaveri. Ipfs: A decentralized storage system for blockchain-based applications. Journal of Network and Computer Applications, 178:103567, 2023.*

[11] *Wei Zhang, Jian Chen, and Rutvij Jhaveri. Distributed data storage solutions for blockchain: A review. IEEE Transactions on Cloud Computing, 11(2):1234–1245, 2023.*

[12] *Oscar Novo. Blockchain for access control in the IoT: A survey. IEEE Internet of Things Journal, 5(2):1184–1195, 2018.*

[13] *John Peterson, Alice Smith, and Rutvij Jhaveri. Decentralized finance (defi): A survey on multi-blockchain architectures. Journal of Financial Innovation, 10(1):1–20, 2024.*