Engineering xxx (xxxx) xxx

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Research
6G Requirements, Vision, and Enabling Technologies—Article

Blockchain for Transparent Data Management Toward 6G

Xuemin (Sherman) Shen [a], Dongxiao Liu [a,*], Cheng Huang [a], Liang Xue [a], Han Yin [a], Weihua Zhuang [a], Rob Sun [b], Bidi Ying [b]

[a] Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada
[b] Huawei Technologies Canada, Kanata, ON K2K 3J1, Canada

ARTICLE INFO

Article history:
Received 31 December 2020
Revised 6 July 2021
Accepted 18 July 2021
Available online xxxx

Keywords:
Blockchain
Data management
Decentralization
Transparency
Privacy

ABSTRACT

The wealth of user data acts as a fuel for network intelligence toward the sixth generation wireless networks (6G). Due to data heterogeneity and dynamics, decentralized data management (DM) is desirable for achieving transparent data operations across network domains, and blockchain can be a promising solution. However, the increasing data volume and stringent data privacy-preservation requirements in 6G bring significantly technical challenge to balance transparency, efficiency, and privacy requirements in decentralized blockchain-based DM. In this paper, we investigate blockchain solutions to address the challenge. First, we explore the consensus protocols and scalability mechanisms in blockchains and discuss the roles of DM stakeholders in blockchain architectures. Second, we investigate the authentication and authorization requirements for DM stakeholders. Third, we categorize DM privacy requirements and study blockchain-based mechanisms for collaborative data processing. Subsequently, we present research issues and potential solutions for blockchain-based DM toward 6G from these three perspectives. Finally, we conclude this paper and discuss future research directions.

1. Introduction

The proliferation of wireless networks has greatly impacted our ways of living and working by providing ubiquitous coverage and seamless connectivity. As the wireless networks continue to evolve, the sixth generation wireless networks (6G) will further integrate heterogeneous access technologies and network slicing [1,2] to support diversified services with dynamic quality-of-service requirements. More importantly, network intelligence plays an essential role not only in improving network resource utilization, but also in enhancing user experience with customized service provisioning [3].

1.1. Data management (DM) toward 6G

The wealth of user data and recent developments in artificial intelligence (AI) technologies lie at the heart of network intelligence toward 6G. With numerous end devices being deployed and connected, wireless big data are generated at a remarkable rate and scale [4]. Through AI-based data processing, wireless big data have great value for efficient network management toward 6G. For example, user trajectory and association history at different access points can be utilized to conduct AI-based network traffic prediction and content catching on the edge for dynamic network resource allocation [1,5]. Thus, how to effectively and efficiently manage user data—that is, DM, which includes multiple data operations in the life-cycle of user data, from data creation to deletion [6,7]—has become a key enabler of future network intelligence. However, the highly dynamic and heterogeneous nature of 6G imposes four major requirements on DM:

(1) **Decentralization:** DM requires collaborations among multiple data stakeholders, including users or machines for data generation, mobile operators for data collection and transmission, and technology vendors (e.g., edge/cloud providers) for data storage and processing. Stakeholders usually come from different network domains that cannot simply agree on a single DM authority. Thus, it is necessary to have a decentralized architecture for data stakeholders to collaboratively manage data life-cycle events [8].

(2) **Transparency:** Due to the lack of mutual trust, the DM process should be transparent and verifiable to data stakeholders. Data owners should be aware of any operation performed over their data [9]. For regulation purposes, "respective responsibilities"

* Corresponding author.
E-mail address: dongxiao.liu@uwaterloo.ca (D. Liu).

https://doi.org/10.1016/j.eng.2021.10.002

Please cite this article as: Xuemin (Sherman) Shen, D. Liu, C. Huang et al., Blockchain for Transparent Data Management Toward 6G, Engineering, https://doi.org/10.1016/j.eng.2021.10.002

should be determined transparently for stakeholders who process data collaboratively [10].

(3) **Efficiency:** The heterogeneity of DM stakeholders, increasing volume of user data, and complexity of data life-cycle events will lead to major concern on assuring efficiency in terms of distributed architecture design, authentication and authorization (AA) management for DM stakeholders, and data-processing mechanisms.

(4) **Privacy:** Privacy preservation in DM refers to both the identity privacy of data stakeholders and the content confidentiality of personal data. Although specific privacy requirements can change with different data operations, general principles are enforced in recent privacy regulations, such as European General Data Protection Regulation (GDPR) [10]. For example, users are granted full control of any operation over their data with identifiable information [11]. A data usage agreement that defines stakeholder obligation should be pre-determined and strictly followed.

It remains an ambitious task to develop a decentralized and transparent DM that satisfies the efficiency and privacy requirements.

1.2. Blockchain-based DM

A blockchain consists of a ledger of blocks of peer-to-peer (P2P) transactions [12]. The blockchain is maintained by distributed nodes in the network, where each (full) node maintains a copy of the ledger. From the perspective of functionality, the blockchain shares some features with a traditional distributed database [13], but utilizes secure consensus protocols to maintain consistency of the ledger among mutually distrusted nodes. Moreover, a blockchain can provide programmability to control the ledger updates with smart contract technology [14].

Blockchain is a promising technology for DM toward 6G, as it naturally addresses the decentralization and transparency requirements. First, DM stakeholders can use the blockchain as the trusted shared storage to record critical DM events [8,15]. Each DM stakeholder can maintain a copy of the shared ledger without relying on a centralized entity. Second, the shared ledger is transparent, and the ledger updates are verifiable to related blockchain nodes. DM stakeholders can design smart contracts in order to conduct various data operations collaboratively. These benefits have motivated many recent discussions on blockchain-based DM schemes [16,17] in future intelligent networks [3,18] and other applications, such as information-centric networks [7], supply-chain management [19], the Internet of Things (IoT) [20–22], and e-healthcare [23].

Given its decentralized and transparent nature, a blockchain-based solution may aggravate the complexity of DM in achieving the requirements of efficiency and privacy [24]. First, distributing data storage to blockchain nodes increases the overall storage overhead. At the same time, to maintain the consistent view of the shared ledger, DM stakeholders run consensus protocols to endorse transactions and verify blocks, which may limit the transaction throughput and increase the data processing burden. Second, due to the storage transparency of the blockchain, onchain data are visible to related blockchain nodes, which contradicts the privacy requirements of user data. Thus, more research efforts should be directed to new designs and practical implementations of blockchain-based DM in order to resolve the efficiency and privacy challenges.

1.3. Organization of this paper

In this paper, we discuss blockchain-based DM for 6G. To address the challenges of efficiency and privacy, we summarize state-of-the-art research progress with potential solutions. The organization of this paper is as follows:

Section 2 presents the blockchain architecture design for DM. We summarize the existing blockchain mechanisms, such as efficient consensus protocols and hybrid chain designs. Moreover, by comparing recent blockchain-based DM schemes, we discuss how DM stakeholders can serve as blockchain components. In Section 3, we explore blockchain-based AA mechanisms for the efficient and privacy-preserving identity management of DM stakeholders. In Section 4, we investigate blockchain-based data-processing mechanisms. After specifying the privacy requirements for blockchain-based data processing, we discuss an on-/off-chain computation model. We also summarize research outcomes on specific privacy-preserving data operations, including data sharing and data analytics. In Section 5, we discuss research issues and potential solutions in detail, in terms of architecture design, AA, and data processing in blockchain-based DM. Finally, we conclude this study and discuss further research directions in Section 6.

2. Architecture design in blockchain-based DM

A blockchain can serve as a decentralized and transparent architecture for DM toward 6G. However, it is not trivial to build DM with black-box use of the blockchain. First, a blockchain essentially deals with maintaining consistent storage and state updates in distributed nodes. As the degree of trust among the nodes can change dramatically in real-world applications, a blockchain can have different architecture designs, with trade-offs between ledger scalability and security. When applying a blockchain to DM, it is also necessary to distinguish among the requirements of different DM use cases. Second, stakeholders can have different capabilities and motivations to participate in DM. Furthermore, there are various roles in a blockchain-based architecture, such as miners and clients. However, it remains unclear how to manage the roles of DM stakeholders in a blockchain.

To address this issue, we explore two essential questions: ① What should the blockchain architecture be for DM, and ② what roles do DM stakeholders play in this architecture? We first review existing blockchain architectures with advantages and limitations for DM, and then discuss two typical use cases for blockchain-based DM: vehicle-to-everything (V2X) [25] and cloud/edge computing.

2.1. Blockchain architectures

Blockchain architectures can be roughly classified into two categories: permissionless blockchains [14] and permissioned blockchains [26]. A permissionless blockchain mainly consists of two kinds of entities: miners and clients [12]. It uses cryptographic currencies to motivate entities to self-organize themselves in public networks. In contrast, a permissioned blockchain is a top–down architecture with three main entities: authorities, miners, and clients. In general, industrial organizations can form a consortium to serve as the supervising authorities of the blockchain. The miners and clients must obtain permission from the authorities before participating in the blockchain. In both architectures, the consensus mechanism is an essential component for maintaining consistency on the ledger.

In terms of consensus protocols, a permissionless blockchain must resist more malicious participants than the permissioned blockchain. A bitcoin blockchain is proven to be secure if the miners possessing the majority of the computational power are honestly following the proof-of-work (PoW) consensus protocol [27]. However, when the number of miners is large, the architecture can suffer from low transaction throughput and high transaction confirmation latency. A permissioned blockchain, such as Hyperledger Fabric [26], relies on the consortium committee to provide

membership management and ordering services. Such top–down architecture incurs fewer restrictions on the consensus protocol, where practical Byzantine fault tolerance (PBFT) and Raft can be implemented. To further improve blockchain scalability, new blockchain architectures have been proposed recently. For permissionless blockchains, Prism [28] and OHIE [29] are two new blockchain architectures to support parallel transaction processing. They separate the single chain into multiple chains and divide the roles of miners into several roles for different tasks.

Although permissionless and permissioned blockchains have different characteristics, most support two attractive functionalities: distributed storage and smart contracts. That is, computer programs can be executed on the distributed environment (the blockchain), which makes blockchains suitable to construct DM platforms toward 6G [30,31].

2.2. Use cases of blockchain-based DM

In the following, we present two exemplary use cases of blockchain-based DM: V2X and cloud/edge computing. Our focus is to summarize how DM stakeholders can participate in the blockchain architecture.

2.2.1. Blockchain-based DM for V2X

V2X communications enable many vehicular applications, such as on-road infotainment and location-dependent services [25,32,33]. To provide more efficient and effective services for pedestrians and drivers in a V2X communication network, V2X service providers must cooperatively communicate with each other and exchange some users' private information. However, this requirement cannot be easily satisfied in the current V2X system, since the vehicle-related data are managed independently by V2X service providers, and inappropriate data sharing may lead to serious privacy information leakage [34] and break the privacy regulations. To bridge the gap between existing V2X services and 6G, blockchains have been introduced into the V2X system, where a large number of V2X service providers can build decentralized trusts. In particular, vehicular information exchange can be recorded onto the blockchain, which allows third-party auditors to trace the information flow and prevent potential privacy leakage. Furthermore, depending on various V2X services, the information written into the blockchain differs; it may consist of vehicle insurance information, driver license information, vehicle velocity, location, and so forth.

A basic blockchain-based V2X communication network includes the following stakeholders: vehicles, roadside units (RSUs), base stations, service providers, edge nodes, and cloud servers. The main difference between DM architectures atop a permissionless blockchain and those atop a permissioned blockchain for V2X services [34] lies in the stakeholders who construct the blockchain. Some existing schemes [30,31] rely on public blockchain platforms as third parties for V2X services. For example, in a public key infrastructure (PKI)-based solution for securing V2X communications that is based on a public blockchain platform [30], vehicles and other stakeholders are conventional clients of Ethereum, which can read/write information on the public ledger and trigger deployed smart contracts. In this setting, the original V2X network architecture and the roles of these stakeholders do not need to be significantly changed, but the stakeholders need to have extra communications with the external public blockchain platform. Although a permissionless blockchain-based DM architecture is considered to be simple and effective, it does not fit well for all V2X services, due to the lack of system scalability and data privacy. A permissionless blockchain platform is public and can be accessed by any party. As a result, some data, such as public certificates and the certificate revocation list (CRL), can be published on the blockchain, while other data, such as personal riding records, should be protected. Moreover, the data-processing delay is high in a permissionless blockchain platform, which makes a permissionless blockchain architecture unsuitable for V2X services with strict latency requirements.

To overcome these issues, many studies incorporate a permissioned blockchain in V2X services [35–42]. In such solutions, the blockchain is maintained by V2X stakeholders themselves, who can be vehicles, RSUs, edge nodes, and cloud servers, depending on V2X application scenarios. For example, mobile edge nodes or RSUs can serve as full nodes for maintaining a permissioned blockchain, since they are sufficiently powerful in terms of computational and storage capabilities. Vehicles usually serve as light nodes, since they have limited resources and high mobility. Compared with the architecture atop a permissionless blockchain, this architecture is more scalable by controlling the number of miners in the blockchain and adopting hybrid consensus protocols at the cost of complicated architecture designs and security models. More specifically, most state-of-the-art architectures have a premise that root trusted authorities exist in V2X services to bootstrap the system.

2.2.2. Blockchain-based DM for cloud/edge computing

Cloud/edge-based DM architecture is established on a centralized model, where a back-end cloud service provider is integrated with front-end interfaces, such as mobile phones, to make data processing and sharing simple and effective. However, the architecture can be vulnerable to internal attacks, due to the lack of procedure transparency at the third-party service provider. Therefore, a more transparent DM framework is essential, in which all data-processing operations can be audited, and even malicious internal attackers can be detected. As a result, a blockchain can be introduced to the cloud/edge-based DM architecture, in order to obtain a transparent DM model with monitoring and auditing capability.

A blockchain is promising for managing multidomain collaborations in a layered edge-computing or joint-cloud architecture [43,44]. Many related solutions have been proposed recently for blockchain-based DM in cloud computing from either permissionless or permissioned blockchains. A basic blockchain-based DM for cloud computing has the following main stakeholders: users, cloud servers, and application service providers, with DM operations including data auditing, data sharing, data integrity checking, and data searching.

Most cloud DM architectures adopt an external permissionless blockchain platform [45–51], without high demands for throughput and latency in data processing. The blockchain is mainly viewed as an honest ledger for storing extra information, while large data are stored in cloud servers with or without privacy protection, according to the privacy requirements. Due to the high cost of processing data on permissionless blockchains, heavy data operations cannot be performed on the chain, although lightweight operations, such as data timestamping and operation record tracking, can be performed. Therefore, off-chain DM operations should only be recorded on the chain after being performed. At the same time, data encryption is a general solution to protect privacy for the data stored on the cloud or the blockchain.

For cloud DM architectures atop a permissioned blockchain, the blockchain is managed by authorized stakeholders, such as cloud servers, edge nodes, and even users [52–56]. The permissioned blockchain can be applied to boost cross-domain trust among different stakeholders. As there are lower on-chain operation costs in a permissioned blockchain, more complex data operations can be done on the chain. In addition, the data privacy protection mechanism is not limited to data encryption. As authorized stakeholders control the on-chain data, they can define access policies for the

data on the blockchain. Although this architecture has many advantages, it relies on the trustworthiness of the authorized stakeholders who serve as blockchain managers. If these stakeholders are compromised, the security and privacy of the architecture cannot be ensured.

A blockchain architecture for DM is shown in Fig. 1. Based on different consensus protocols, distributed ledger storage, and smart contracts, blockchain-based DM can support various V2X and cloud/edge applications. Table 1 provides a summary of the blockchain architectures for DM in two use cases.

3. AA in blockchain-based DM

3.1. AA requirements for DM stakeholders

AA is an indispensable component of blockchain-based DM [57]. In particular, AA addresses two essential questions in DM: Who you are and what you can do. First, there can be multiple participants in DM, such as users, storage nodes, and computing nodes. Authentication helps the DM system to determine unforgeable identities and the exact roles of DM stakeholders. Second, based on their roles, DM stakeholders are authorized to conduct a wide range of operations, such as reading data and modifying the data status. With the above basic functionalities, AA can further help DM stakeholders establish secure and confidential communication channels, which is essential in a distributed blockchain environment. Moreover, the non-repudiability from AA is the key in determining the accountability of DM stakeholders in case of any dispute.

Blockchain-based DM toward 6G has new requirements for AA mechanisms:

(1) **Distributed management:** Without a traditional centralized authority, AA management in DM should be conducted by a set of authorities in a transparent manner.

(2) **Efficiency and privacy:** As the roles of DM stakeholders can change dynamically, blockchain-based AA should support efficient credential update and revocation. Also, the real identities of DM stakeholders should be kept private for certain use cases in order to achieve conditional privacy preservation if necessary. Below, we discuss existing works on achieving transparent, efficient, and privacy-preserving blockchain-based AA for DM.

3.2. Blockchain-based AA

In a complex DM environment, there may exist multiple stakeholders with the right to generate identities for their users and make authorizations for data operations, such as the independent identity management in Fig. 2 [50]. In such a model, cross-domain AA is required due to frequent information exchange between stakeholders. Certificate management can become a hurdle, as each stakeholder has its own certificate authority (CA) for management. Some stakeholders may be compromised and may publish or utilize fake certificates for data operations. To reduce the management costs and the security risks of cross-domain AA, a manager can be introduced to play the role of the centralized identity management shown in Fig. 2 [50], such as a single-sign-on service provider. However, this model requires the DM stakeholders to agree on a single manager, which may not always be practical toward 6G. As shown in Fig. 3, blockchain-based decentralized identity management [50–52] can enable stakeholders to collaboratively manage user identities, authenticate users, and update authorization policies in a distributed and transparent manner. More specifically, the blockchain is managed by a

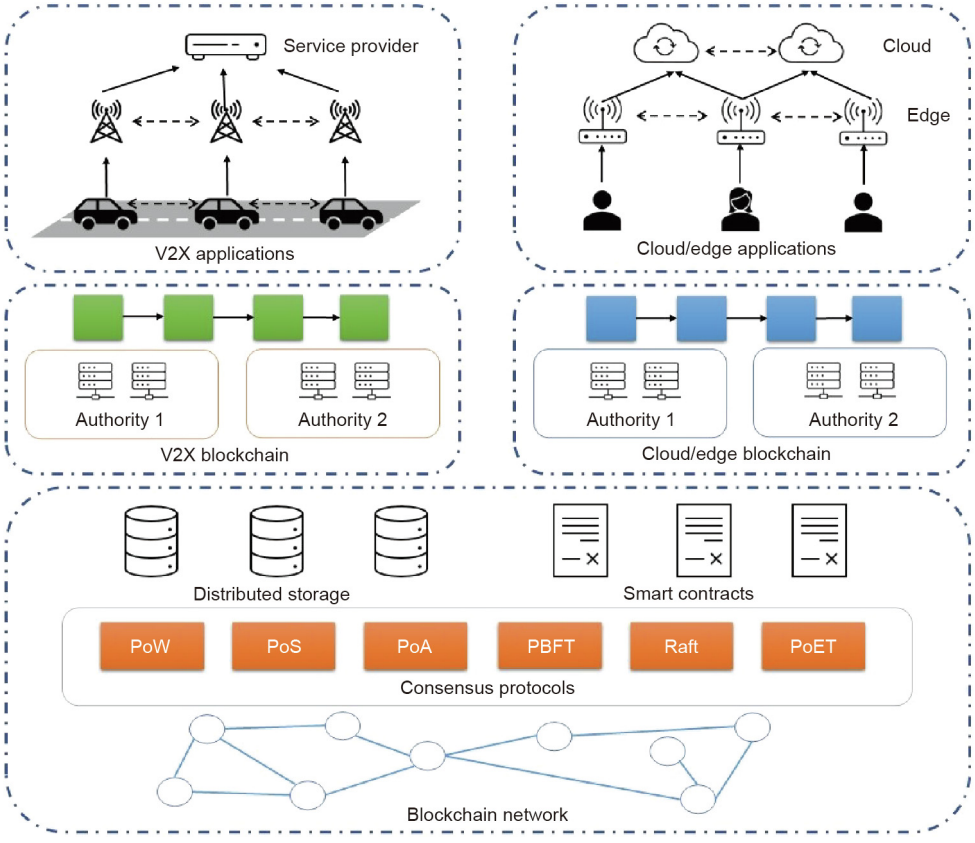

Fig. 1. A blockchain-based DM. PoS: proof of stake; PoA: proof of authority; PoET: proof of elapsed time.

Table 1
Blockchain-based DM architecture: use case.

Use case	Application	Blockchain architecture	Consensus protocol	Maintainer of blockchain
V2X	On-road infotainment and location-dependent services	Permissionless	PoW/PoS	Third-party
		Permissioned	PBFT/Raft	RSU and edge nodes
Cloud/edge computing	Data auditing, data sharing, and data searching	Permissionless	PoW/PoS	Third-party
		Permissioned	PBFT/Raft	User and cloud/edge server

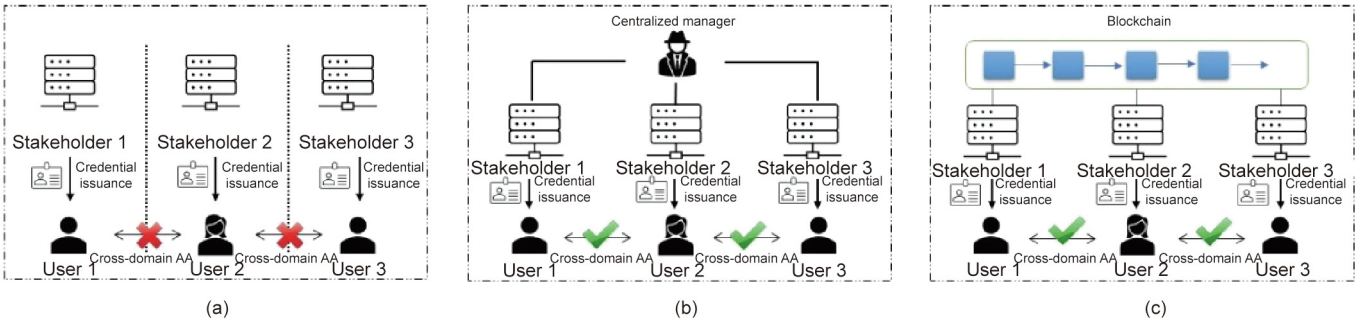

Fig. 2. The evolution of identity management: from independence to decentralization. (a) Independent management in DM; (b) centralized identity management in DM; (c) decentralized identity management in DM.

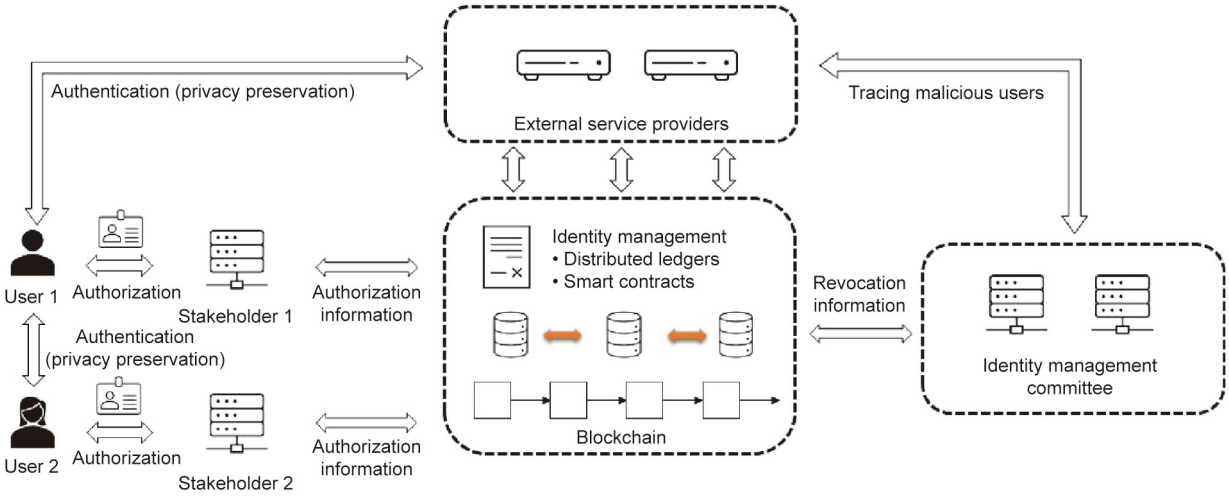

Fig. 3. General procedures of blockchain-based authentication and authorization in DM.

consortium committee and can provide AA services for external service providers. Even if some stakeholders are compromised, all membership updates and revocation operations on the blockchain are still traceable and accountable.

Extensive research efforts have been made to utilize blockchains to enhance AA systems [58]. For traditional certificate-based AA, blockchain-based mechanisms have been proposed to guarantee certificate transparency and revocation transparency for blockchain-based DM [59,60]. More specifically, CAs publish certificates for stakeholders and users, while a group of stakeholders update their certificates on a public blockchain. The validity of the certificates on the blockchain relies not only on the security of the CAs, but also on the group of data stakeholders, as it is necessary for the majority to be honest. Instead of focusing on certificate transparency, other works define authority transparency frameworks to address the issue of auditing AA management among stakeholders [61,62] by viewing the blockchain as public and immutable logs of certificate generation, updating, and revocation.

In contrast to certificate-based AA systems, self-sovereign identity is a blockchain-based identity management mechanism [63]. More specifically, instead of relying on a CA to manage user identities, users themselves can create, store, transfer, and revoke their identity credentials through a blockchain. In this way, the risk of the single-point failure of a centralized CA is much reduced. To achieve fine-grained data access control, attribute-based AA mechanisms such as attribute-based encryption (ABE) can be combined with the blockchain, with user attributes being embedded in the ledgers and smart contracts. Users can access data and retrieve decryption keys based on their attributes [64]. Combined with the blockchain and chameleon hash functions, dynamic attribute updates can be achieved in the blockchain [65].

A blockchain brings many advantages in managing users' identities to satisfy different security properties in the DM. However, it also raises privacy concerns, as all information stored in the blockchain is transparent. Therefore, privacy-preserving mechanisms can be integrated with blockchain-based AA schemes to provide privacy protection. One basic mechanism is based on pseudonyms. Each user can hold a large number of pseudonyms for AA, and the pseudonyms are always denoted by public keys in the blockchain. For example, a blockchain data-sharing system, Ghostor, hides user

identities but allows users to detect integrity violations of remotely stored data [66]. Anonymity is achieved by using a technique named "anonymously distributed shared capabilities." Since the pseudonyms are locally stored by users and are difficult to manage if the number of pseudonyms is larger, other mechanisms, such as group signatures and ring signatures, can help to protect user identity privacy in a blockchain-based DM. A user's anonymous identity generated from group/ring signature schemes can be stored at the user side for multiple uses across different applications. This identity privacy protection mechanism has been adopted by some blockchain platforms, such as Moreno [67].

Group/ring signature schemes are built upon Fiat–Shamir signature schemes, which can be utilized for self-sovereign identity management in blockchains [68]. Anonymous credentials can also be delegated at different levels to inspire more DM applications [69]. Under these circumstances, identity privacy is protected and accountability is guaranteed for tracing malicious users. For example, there are blockchain-based AA schemes in which user identities can be traced under certain stringent conditions [70,71]. One or multiple stakeholders can collaboratively generate anonymous credentials for their users based on the zero-knowledge proof technique. When a user behaves maliciously and needs to be traced, these stakeholders can reveal user identity accordingly. In this way, privacy and accountability can be simultaneously satisfied. For attribute-based access control on the blockchain, a general privacy-preserving approach is to hide the access policies by designing an attribute-hiding ABE [72]. This method is different from previous mechanisms by protecting data attributes and policies rather than user identities.

4. Data processing in blockchain-based DM

Data processing can refer to a wide range of operations in the life-cycle of data items [7,8]. For a blockchain-based DM, we mainly focus on data operations that require interactions between multiple DM stakeholders, including data sharing and collaborative data analytics. In this section, we first discuss privacy and efficiency requirements with general privacy and computation models. Based on these requirements, we summarize the existing literature on blockchain-based data sharing and analytics.

4.1. Privacy requirements and model

For data processing in blockchain-based DM, a general privacy requirement is to restrict data exposure. More specifically, data exposure can be characterized by the following questions:

(1) **What is the sensitivity of the data?** First, data sensitivity can vary dramatically depending on the application. For example, user identity data in financial applications are highly confidential, and can lead to economic loss in case of any exposure. Second, data sensitivity can change with the amount of data. For example, a single exposure of a user location may incur limited damage, while the exposure of consecutive user locations may reveal users' daily routines [73]. Third, data sensitivity can change with time. Many types of data, such as legal files [74], have a "sealing" period, within which the data should not be exposed. After the "sealing" period, the data can be accessed by the public or by certain entities.

(2) **To whom are the data exposed?** Data processing can involve various entities, which can be roughly categorized into internal/external participants and the blockchains. The term "internal participants" refers to the DM stakeholders involved in the data processing. In contrast, the term "external participants" refers to entities that are not involved in the data processing, such as an external attacker. In blockchain-based DM, there is a shared view among blockchain participants. In this case, the blockchain can be modeled as a special entity for data exposure.

From these two questions, privacy requirements in blockchain-based DM can be categorized into four levels:

(1) **Privacy from user anonymity:** This requires user identity information to be separated from the dataset before being processed. However, for a data processor (i.e., an entity that conducts data processing) with strong background knowledge, it is highly possible for the processor to recover user identity information from the dataset.

(2) **Confidentiality for external participants:** Data of less sensitivity can be processed by data processors in plaintext, but cannot be exposed to external participants. This requirement relies on the trustworthiness of the data processor.

(3) **Confidentiality for internal participants:** For data with high sensitivity, data processing should expose as little information as possible to the data processors, including data content, user identity, and data access patterns.

(4) **Confidentiality for blockchains:** Sensitive data should not be directly stored on the blockchain. Similarly, sensitive data operations should not be conducted by smart contracts.

In blockchain-based DM, privacy requirements for different DM applications can change dramatically with the data sensitivity and the roles of data stakeholders. Therefore, GDPR [10] does not provide specific privacy requirements, but rather defines general principles. More specifically, it requires that users have full control over the DM operations on their data.

Internal participants, such as data controllers and data processors, must agree on data usage terms with users and must strictly follow this agreement in the data processing. At the same time, any unauthorized data sharing with external participants is forbidden.

Since privacy requirements can sometimes be vague or ambiguous, it is essential to design privacy models that help users, DM system designers, and regulators to better understand privacy regulations in an executable and implementable manner. A data flow diagram (DFD) is a good way of modeling DM. DFDs are similar to process diagrams in software engineering, and can integrate GDPR elements and data life-cycle events [75]. Unlike models based on data life-cycle events, resource or capability requirements for data stakeholders can be utilized to implement DM with GDPR compliance [76]. Moreover, for blockchain-based DM, executable privacy models can be implemented to automatically regulate the cloud data operations involved in smart contracts [77,78].

4.2. Efficiency requirements and computation model

For data sharing and analytics in blockchain-based DM, a straightforward solution is everything-on-chain, which involves storing the entire dataset on the blockchain and conducting data processing via smart contracts. However, this can require prohibitive storage and place a heavy computational burden on the blockchain participants. To address this issue, it is possible to introduce off-chain storage or computation nodes that can store data or perform data processing more efficiently, and only upload pivotal information onto the blockchain. This paradigm is regarded as an on-/off-chain model [79].

In a general on-/off-chain model, an external data storage provider can store the hash values of the data items onto the blockchain [80]. In this way, the integrity of the off-chain data storage can be ensured, since on-chain hashes cannot be modified. This model can also eliminate the direct exposure of private data to the blockchain. The hash-based approach relies on the trustworthiness of an external data storage provider to perform data operations. In blockchain-based DM, it is desirable to allow weaker security assumptions for the storage provider and to design more expressive on-chain authenticators for verifying the correctness of off-chain

data operations. For example, an aggregation of multiple data records can be computed by an off-chain cloud server, which only sends verifiable computation results to the blockchain. Below, we discuss research works on constructing on-/off-chain models. The main requirement of such a model is to have verifiable off-chain executions, including zero-knowledge succinct non-interactive argument (SNARG) and a trusted execution environment (TEE).

SNARG is a system in which a prover can convince a verifier of the existence of a secret for a public relation. The relation can be represented by an arithmetic circuit for generally verifiable computations [81].

Verification of SNARG is efficient and can be privacy-preserving without directly exposing the inputs and outputs of the computation. As a result, SNARG is widely used to construct an on-/off-chain computation model [24] for blockchain-based DM. However, the verification efficiency of SNARG comes at the cost of a trusted setup of relation-dependent public parameters and expensive prover computation overhead. Therefore, it is critical to properly set universal or updatable public parameters [82] or to use a secure multiparty computation protocol to generate public parameters for SNARG systems. Moreover, SNARG does not naturally provide privacy against internal participants. Data processors must have access to the original data, which is not always desirable for DM applications.

TEE, such as the Intel Software Guard Extension (SGX) [83], provides another way to verify computations. Before execution in TEE, codes are loaded into a secure enclave, which is secure hardware with protected memories. To ensure the loaded codes and data are trusted, SGX provides a remote attestation service: TEE generates an attestation request to a remote attestation service to ensure the integrity and correctness of code executions. Unlike SNARG, TEE does not require a trusted setup of public parameters and is more efficient in generating proof of computations. Therefore, TEE can facilitate the design of on-/off-chain computation models [84,85] by serving as a reliable and authenticated off-chain computation unit. Moreover, with the integration of a key manager, TEE-based solutions can achieve authenticated and encrypted communications between the enclave and external environments to achieve privacy protection against malicious data processors. However, there are some challenges in the practical implementation of TEE. First, a comprehensive and formal security analysis [86] of TEE has recently been discussed. Second, remote attestation strongly relies on the service provider, which can be a single trust point in a blockchain environment.

Besides certifying the computation results with the aid of either SNARG or TEE, another potential path is to adopt game theory to create a competitive relationship between multiple off-chain resource providers in order to eliminate cheating [87]. For example, two cloud servers can be assigned the same computation tasks. By setting proper financial gains and losses, the two cloud servers can be motivated to correctly complete the computing tasks.

4.3. Blockchain-based data-processing mechanisms

An on-/off-chain model based on SNARG, TEE, or the two-server model provides general solutions to data processing tasks. However, for specific tasks, specialized design strategies (e.g., new data structures) are required to fulfill the privacy and efficiency requirements.

4.3.1. Data sharing

When data are collected and stored in a blockchain-based DM, it is important to share or trade the data to enable multiple data-intensive applications [18,88].

Various privacy requirements can be achieved for data sharing using different techniques. Identity privacy for both data owners and receivers can be achieved by pseudonyms [89] or group signature-based anonymous credentials. Data encryption mechanisms with key management techniques can be enforced to achieve on-chain data confidentiality. For fine-grained access control in data sharing, attribute-based or functional encryption can also be used [90,91], in which a data encryption key or a ciphertext can determine access policies. Unlike methods that are based on encryption key management, reputation management [92] can also be integrated into data sharing. In reputation management, data senders and receivers can be enabled to leave reviews for the data-sharing process [93]. The accumulated review score can serve as the criteria for access assessment. For example, a proof-of-collaboration consensus protocol is designed for data sharing at the edge [94], where reputation based on collaboration is quantified. Recently, researchers have also considered GDPR requirements in data sharing [95,96]. More specifically, a blockchain-based solution can enable users to fully control their personal data, which meets the GDPR requirements of consent-based DM.

Data owners often outsource their data to a third-party storage provider, such as a cloud server, and rely on the storage provider to manage their data. In this model, the blockchain can serve as a trusted auditor for the data-sharing process [97]. To relieve data owners of heavy key management overheads, it is desirable to have a reliable key manager for data encryption and decryption. Threshold cryptography, such as (t,n) Paillier crypto (where t is a threshold number and n is the number of secret shares), can be utilized to protect data that are stored on the cloud and shared on the blockchain [98]. At the same time, it is essential to securely choose a set of committee members to manage the keys. The blockchain can also be utilized to manage data modification on the cloud storage [46], where a trusted authority (TA) is integrated with the smart contract to complete the modification process.

Aside from data sharing, data trading can further explore the data value. A blockchain-based digital identity exchange scheme for financial institutions has been proposed [99], in which SNARG is utilized to prove the authenticity of identities in a privacy-preserving manner. TEE can be utilized to build a data-trading platform [100,101] that preserves fairness for both buyers and sellers, and ensures on-chain privacy for data processing [102].

4.3.2. Data analytics

A blockchain can support various data-analysis tasks [103] for intelligent 6G networks. A blockchain-based learning framework is proposed in Ref. [104] to securely compute model parameter updates with a threshold Paillier algorithm. Another important data analysis mechanism is to enable flexible and expressive queries. For data stored on the blockchain, a query should be efficient, and the correctness proof should be verified at a low cost [13], where authenticated data structures can be tailored for both inter-block and cross-block query processing. To maintain data privacy on the blockchain, it is possible to encrypt the data on the chain with a searchable index [105]. Then, a smart contract can be constructed for querying over the searchable indexes, which naturally ensures the verifiability of the search result. For searching over the location-based data, it is desirable to established range-based searchable indexes [106]. When data are stored off the blockchain, data owners can build an on-chain authenticator of the data index from SNARG or cryptographic accumulators. In this way, query operations can be conducted off-chain and the query result can be verified on-chain. More expressive verifiable queries can be supported by integrating database query techniques [107].

A blockchain can naturally serve as a log system [108,109] due to its transparency and immutability. That is, data stored on the blockchain can be utilized to conduct event-driven system debugs and analysis. To support fine-grained data provenance operations,

expressive data indexes can be built atop the original blockchain data [13,110]. At the same time, the blockchain can be utilized to construct log systems for DM applications. A lightweight blockchain logging mechanism is proposed in Ref. [111] with a new log storage structure for data-intensive applications. To achieve communication between different systems, multichain interoperability is considered in Ref. [112]. Compared with direct use of the blockchain for log storage, storing sensitive log data off the blockchain can reduce the on-chain overheads and privacy leakage. In particular, an IoT data provenance scheme is proposed in Ref. [113]. SNARG is adopted to succinctly store provenance data at each network administrator with succinct authenticators on the blockchain for cross-domain network provenance queries. Cryptographic accumulators can be utilized for a single log server to generate proofs of correct log updates, as discussed in Ref. [114] for certificate transparency services. A summary of blockchain-based privacy-preserving data processing is shown in Table 2 [24,84,87,89–91,96,100,104,105,108,115].

5. Research issues and potential solutions

While blockchain-based solutions have great potential for DM toward 6G, many unresolved research challenges still remain. In this section, we discuss research issues and potential solutions in detail, in terms of architecture design, AA, and data processing in blockchain-based DM.

Table 2
Summary of privacy-preserving data processing in blockchains.

Design goal	References	Functionalities	Privacy guarantee
Computation model	[24]	Design a tool chain from SNARG to compile an off-chain program into an Ethereum smart contract	Achieve program execution privacy against the blockchain
	[84]	Design an on-/off-chain computation framework from TEE	Achieve program execution privacy against the blockchain
	[87]	Design a two-server model and use game theory to achieve verifiable computations	NA
Data sharing	[89]	Data sharing on the blockchain	Achieve on-chain data confidentiality and identity privacy for senders/receivers
	[90,91,115]	Data sharing on the blockchain with access control	Achieve data confidentiality and fine-grained access control
	[96]	Data sharing on the blockchain with GDPR compliance	Achieve on-chain data confidentiality and consent-based access control
	[100]	TEE-assisted data trading on the blockchain	Achieve data confidentiality against buyers by only revealing data analysis results
Data analytics	[104]	Blockchain-based learning framework	Achieve confidentiality of local gradients
	[105]	Blockchain-based data search	Achieve on-chain data and index confidentiality
	[108]	Blockchain-based data provenance framework	Achieve pseudonymity for data subjects and on-chain data confidentiality

NA: not applicable.

5.1. Architecture design in blockchain-based DM

Although there are many blockchain architectures for DM, most are designed for applications, and various challenging issues related to DM architecture designs still exist, as follows:

(1) **Incentive and regulation mechanism design:** A permissionless blockchain utilizes financial incentives for its participants, while a permissioned blockchain relies on a consortium committee to regulate its procedures. In practice, DM stakeholders toward 6G are highly heterogeneous, and can have different capabilities, profit considerations, and management frameworks. Therefore, the question of how to design incentive mechanisms for permissionless blockchain-based DM and regulation rules for permissioned blockchain-based DM remains a challenging issue. Multiple technologies, such as game theory and threshold cryptography, can be integrated to offer effective group and organization behavior management.

(2) **Blockchain architecture with network slicing:** Network function virtualization (NFV) enables flexible resource sharing over the same physical infrastructures of a communication network and is envisioned to play an important role in future wireless networks [1]. In NFV, a network slice can contain a set of virtualized functions from multiple physical resource providers and can be managed by local or centralized software defined networks (SDNs) controllers, making DM more complicated. To manage data flows among virtualized functions, the DM architecture design should take into account the roles of new 6G stakeholders, such as a third-party resource provider and a cloud-based slice orchestrator. As the business model and implementation details of NFV-enabled 6G become clearer in the future, their impact on DM architecture design can be studied further.

(3) **Hybrid blockchain architecture design:** Blockchain architectures for DM are designed based on either a permissioned blockchain or a permissionless blockchain. Both architectures have their own advantages and disadvantages; the core component of these architectures is the consensus protocol, which affects the system's scalability and security. To further improve system scalability while simultaneously satisfying the security requirements, a flexible and hybrid blockchain architecture should be utilized, which can support the switching of consensus protocols according to different application requirements in DM. Moreover, as the blockchain plays a critical role in the new information infrastructure for DM toward 6G, blockchain-as-a-service can be a potential solution to provide plug-in DM architecture design [26] that integrates new technologies, such as lightweight clients [116] and stateless blockchains [117].

(4) **Efficient cross-chain interoperability with privacy preservation:** Current DM architectures are designed based solely on a single ledger, without fully considering cross-chain interoperability. With a heterogeneous blockchain architecture for DM [118] that accommodates multiple applications, each application may establish its own sub-chain to manage its own data with privacy preservation. This approach is similar to the concept of private channels in a permissioned blockchain, but does not support cross-chain interoperability, due to privacy concerns. Therefore, a new blockchain architecture for DM with efficient cross-chain interoperability requires further study, especially from the perspective of privacy preservation. Hierarchical blockchain architectures can be designed to manage cross-chain communications at the consensus level. Moreover, it is possible to set broker nodes that operate over multiple chains. With identity management for the broker nodes, cross-chain communications can be securely facilitated.

5.2. AA in blockchain-based DM

Although blockchain-based AA mechanisms have many advantages, they also raise some efficiency and privacy concerns that should be carefully addressed.

(1) **Lightweight AA:** One main difference between blockchain-based AA schemes and conventional AA schemes is that users can self-maintain their identities, with only necessary information being uploaded to the blockchain. With a complicated DM architecture in which multiple stakeholders coexist, a user with limited computing and storage capabilities may have different identity credentials for various use cases. At the same time, blockchain storage and computing resources are expensive in terms of throughput and latency restrictions. As a result, how to achieve blockchain-based lightweight identity management becomes an important issue for DM toward 6G. A potential solution is to integrate with an external credential server for credential management. To enable users to fully control their credentials, additional security guarantees should be achieved, such as verifiable credential updates based on cryptographic accumulators [9] or TEE-based processing.

(2) **Distributed AA with dynamic updates:** To further eliminate trust requirements for any single entity, critical AA operations should be conducted by a set of key managers, such as distributed credential issuance and revocation [70]. Such a model can involve many communications between the key managers, and an effective incentive and regulation mechanism is required to manage their behavior.

Threshold cryptography can be utilized to reduce the computational burdens on the key managers. At the same time, the membership of key managers can change over time and needs to be updated frequently. When the set of key managers changes, the forward and backward security of the identity credentials should also be ensured. That is, the question of how to achieve secure and efficient committee updates becomes a challenging issue. One potential solution is proactive secret sharing [119], in which shared secrets among key managers can be updated frequently. Critical management operations can also be conducted in a secure hardware execution environment.

(3) **Balancing AA privacy and accountability:** Identity privacy can have fine-grained levels in DM by only revealing necessary identity information under privacy regulations, such as an organization membership and stakeholder attributes. For different DM use cases, flexible privacy modeling and execution can be integrated with a smart contract to enforce AA privacy management [77]. However, identity privacy should not be uncompromisable for DM toward 6G. In case of strong dispute, blockchain-based AA should recover the real identities of stakeholders in order to conduct investigations and enforce accountability, which can be achieved using threshold encryption techniques. In this case, it is important to have a clear criterion to decide when and how to recover stakeholder identity. A hierarchical identity management committee can be designed with specialized regulatory frameworks.

5.3. Data processing in blockchain-based DM

There have been extensive studies on blockchain-based DM, from SNARG/TEE-based solutions for general computations to specialized designs for data sharing and data analytics. However, for blockchain-based DM toward 6G, the question of how to balance functionality, efficiency, and privacy continues to pose the following technical challenges.

(1) **On-chain process design:** Blockchain provides a trusted and reliable shared view of certain DM processes among DM stakeholders. As the on-chain storage and computation resources are limited and may cause privacy concerns, DM stakeholders must carefully decide what information to share. There may be very subtle differences between information that should and should not be shared, which may include hash values of original data for integrity checking, DM life-cycle event logs, or just proof of the existence of DM operations. For privacy and efficiency, only pivotal information should be shared, with selective disclosure only to necessary participants [69]. At the same time, there may be cases when on-chain data need to be removed, with redactable blockchain techniques as a potential solution.

(2) **Privacy model design:** Blockchain-based DM is complicated by its dynamic and heterogeneous participants in various applications, which can lead to rapid changes in privacy requirements [74]. As a result, privacy modeling and evaluations under privacy regulations should be considered to enable flexible privacy management on the blockchain [77], where a natural language processing technique can be a potential solution to help smart contracts better understand privacy requirements.

(3) **Modular design for data processing:** Many existing designs can achieve privacy preservation for different DM operations. For example, SNARG can support general arithmetic computations with succinct on-chain verification, TEE is efficient for verifiable hash computations, and searchable encryption can have specialized designs for different query operations. In practice, a DM instance may incur multiple data operations, where a solution based on a single technique cannot meet both the efficiency and privacy requirements. Modular design strategy [120] is a potential solution that decouples DM operation, such as keyword query and identity management [121], with efficient instantiations from different techniques. This strategy requires an overall understanding of different verifiable computation systems in terms of their advantages and limitations. A universal compatible model [122] can be utilized to analyze the systematic security.

(4) **Automation versus transparency and accountability:** In Article 22 of GDPR, users have the right to object to automatic decisions regarding their data, which may contradict the automation property of blockchain [11] and AI-based decision-making. However, it is often difficult to guarantee transparency and accountability during an AI-assisted decision-making process in blockchain-based DM [9]. A potential solution is to design efficient algorithms to directly evaluate outputs from automatic processes. At the same time, users should be given clear explanations of the impacts of the automatic process on their data and granted the right to object in case of any privacy concerns. For collaborative data processing, it is important to enforce the joint accountability of involved DM stakeholders by establishing DM operation provenance and forensic mechanisms.

6. Conclusions

In this paper, we investigated blockchain-based DM for 6G and highlighted its benefits of decentralization and transparency. By identifying efficiency and privacy challenges, we focused on DM architecture design, the AA of DM stakeholders, and blockchain-based data processing.

To explore potential solutions that balance transparency, efficiency, and privacy in decentralized blockchain-based DM, further research can be directed to the following open issues. First, the impact of network virtualization on DM architecture design should be discussed. Blockchain-based DM requires a flexible and versatile architecture with efficient consensus protocols, inter-chain operability, and fast service-oriented configurations. Second, lightweight and distributed AA with dynamic updates should be designed in order to strike a balance between AA privacy and accountability for blockchain-based DM. Third, an executable pri-

vacy model that can accommodate a wide range of privacy requirements in different DM operations should be achieved. Modular integration of privacy-preserving data-processing techniques should be explored under the privacy models.

Acknowledgments

This work was supported by research grants from Huawei Technologies Canada and from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

Compliance with ethics guidelines

Xuemin (Sherman) Shen, Dongxiao Liu, Cheng Huang, Liang Xue, Han Yin,Weihua Zhuang, Rob Sun, and Bidi Ying declare that they have no conflict of interest or financial conflicts to disclose.

References

[1] Shen X, Gao J, Wu W, Lyu K, Li M, Zhuang W, et al. AI-assisted network-slicing based next-generation wireless networks. IEEE Open J Veh Technol 2020;1:45–66.
[2] Wu W, Chen N, Zhou C, Li M, Shen X, Zhuang W, et al. Dynamic RAN slicing for service-oriented vehicular networks via constrained learning. IEEE J Sel Areas Commun 2021;39(7):2076–89.
[3] Dai Y, Xu D, Maharjan S, Chen Z, He Q, Zhang Y. Blockchain and deep reinforcement learning empowered intelligent 5G beyond. IEEE Netw 2019;33(3):10–7.
[4] Dai HN, Wong RCW, Wang H, Zheng Z, Vasilakos AV. Big data analytics for large-scale wireless networks: challenges and opportunities. ACM Comput Surv 2019;52(5):1–36.
[5] Zhou C, Wu W, He H, Yang P, Lyu F, Cheng N, et al. Deep reinforcement learning for delay-oriented IoT task scheduling in space–air–ground integrated network. IEEE Trans Wirel Commun 2021;20(2):911–25.
[6] Shen X, Huang C, Liu D, Xue L, Zhuang W, Sun S, et al. Data management for future wireless networks: architecture, privacy preservation, and regulation. IEEE Netw 2021;35(1):8–15.
[7] Li R, Asaeda H. A blockchain-based data life cycle protection framework for information-centric networks. IEEE Commun Mag 2019;57(6):20–5.
[8] Freund GP, Fagundes PB, de Macedo DDJ. An analysis of blockchain and GDPR under the data lifecycle perspective. Mob Netw Appl 2020;26(2):266–76.
[9] Abiteboul S, Stoyanovich J. Transparency, fairness, data protection, neutrality: data management challenges in the face of new regulation. J Data Inf Qual 2019;11(3):1–9.
[10] General Data Protection Regulation [Internet]. Brussel: European Commission; [cited 2020 Dec 24]. Available from: https://gdpr-info.eu/.
[11] Herian R. Blockchain, GDPR, and fantasies of data sovereignty. Law Innov Technol 2020;12(1):156–74.
[12] Nakamoto S. Bitcoin: a peer-to-peer electronic cash system. Report. Satoshi: Nakamoto Instityte; 2008.
[13] Xu C, Zhang C, Xu J. vChain: enabling verifiable boolean range queries over blockchain databases. In: Proceedings of the 2019 International Conference on Management of Data; 2019 Jun 30–Jul 5; Amsterdam, the Netherlands; 2019. p. 141–58.
[14] Wood G. Ethereum: a secure decentralised generalised transaction ledger. Report Ethereum Project; 2014.
[15] Vo HT, Kundu A, Mohania MK. Research directions in blockchain data management and analytics. In: Proceedings of the 21st International Conference on Extending Database Technology (EDBT); 2018 Mar 26–29; Vienna, Austria; 2018. p. 445–8.
[16] Zyskind G, Nathan O, Pentland AS. Decentralizing privacy: using blockchain to protect personal data. In: Proceedings of 2015 IEEE Security and Privacy Workshops; 2015 May 21–22; San Jose, CA, USA; 2015. p. 180–84.
[17] Deepa N, Pham QV, Nguyen DC, Bhattacharya S, Prabadevi B, Gadekallu TR, et al. A survey on blockchain for big data: approaches, opportunities, and future directions. 2020. arXiv:2009.00858.
[18] Zhang G, Li T, Li Y, Hui P, Jin D. Blockchain-based data sharing system for AI-powered network operations. J Commun Inf Netw 2018;3(3):1–8.
[19] Wu H, Cao J, Yang Y, Tung CL, Jiang S, Tang B, et al. Data management in supply chain using blockchain: challenges and a case study. In: Proceedings of 2019 28th International Conference on Computer Communication and Networks; 2019 Jul 29–Aug 1; Valencia, Spain; 2019. p. 1–8.
[20] Oktian YE, Lee SG, Lee BG. Blockchain-based continued integrity service for IoT big data management: a comprehensive design. Electronics 2020;9(9):1434.
[21] Shi P, Wang H, Yang S, Chen C, Yang W. Blockchain-based trusted data sharing among trusted stakeholders in IoT. Softw Pract Exper 2021;51(10):2051–64.
[22] Xiong Z, Zhang Y, Luong NC, Niyato D, Wang P, Guizani N, et al. The best of both worlds: a general architecture for data management in blockchain-enabled Internet-of-Things. IEEE Netw 2020;34(1):166–73.
[23] Shi S, He D, Li L, Kumar N, Khan MK, Choo KKR. Applications of blockchain in ensuring the security and privacy of electronic health record systems: a survey. Comput Secur 2020;97:101966.
[24] Eberhardt J, Tai S. ZoKrates-scalable privacy-preserving off-chain computations. In: Proceedings of 2018 IEEE International Conference on Internet of Things and IEEE Green Computing and Communications and IEEE Cyber, Physical and Social Computing and IEEE Smart Data; 2018 Jul 30–Aug 3; Halifax, NS, Canada; 2018. p. 1084–91.
[25] Abboud K, Omar HA, Zhuang W. Interworking of DSRC and cellular network technologies for V2X communications: a survey. IEEE Tran Veh Technol 2016;65(12):9457–70.
[26] Androulaki E, Barger A, Bortnikov V, Cachin C, Christidis K, De Caro A, et al. Hyperledger fabric: a distributed operating system for permissioned blockchains. In: Proceedings of the Thirteenth EuroSys Conference; 2018 Apr 23–26; Porto, Portugal; 2018. p. 1–15.
[27] Garay J, Kiayias A, Leonardos N. The Bitcoin backbone protocol: analysis and applications. In: Proceedings of EUROCRYPT 2015; 2015 Apr 26–30; Sofia, Bulgaria; 2015. p. 281–310.
[28] Bagaria V, Kannan S, Tse D, Fanti G, Viswanath P. Prism: deconstructing the blockchain to approach physical limits. In: Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security; 2019 Nov 11–15; London, UK; 2019. p. 585–602.
[29] Yu H, Nikolić I, Hou R, Saxena P. OHIE: blockchain scaling made simple. In: Proceedings of IEEE Symposium on Security and Privacy; 2020 May 18–21; San Francisco, CA, USA; 2020. p. 90–105.
[30] Lin C, He D, Huang X, Kumar N, Choo KKR. BCPPA: a blockchain-based conditional privacy-preserving authentication protocol for vehicular *ad hoc* networks. IEEE Trans Intell Transp Syst. In press.
[31] Li M, Weng J, Yang A, Liu JN, Lin X. Toward blockchain-based fair and anonymous ad dissemination in vehicular networks. IEEE Trans Veh Techol 2019;68(11):11248–59.
[32] Cheng HT, Shan H, Zhuang W. Infotainment and road safety service support in vehicular networking: from a communication perspective. Mech Syst Signal Process 2011;25(6):2020–38.
[33] Li M, Zhu L, Lin X. Efficient and privacy-preserving carpooling using blockchain-assisted vehicular fog computing. IEEE Internet Things J 2018;6(3):4573–84.
[34] Huang C, Lu R, Ni J, Shen X. DAPA: a decentralized, accountable, and privacy-preserving architecture for car sharing services. IEEE Trans Veh Technol 2020;69(5):4869–82.
[35] Aujla GS, Singh A, Singh M, Sharma S, Kumar N, Choo KKR. BloCkEd: blockchain-based secure data processing framework in edge envisioned V2X environment. IEEE Trans Veh Technol 2020;69(6):5850–63.
[36] Jameel F, Javed MA, Zeadally S, Jäntti R. Efficient mining cluster selection for blockchain-based cellular V2X communications. IEEE Trans Intell Transp Syst 2021;22(7):4064–72.
[37] Rawat DB, Doku R, Adebayo A, Bajracharya C, Kamhoua C. Blockchain enabled named data networking for secure vehicle-to-everything communications. IEEE Netw 2020;34(5):185–9.
[38] Yang Z, Yang K, Lei L, Zheng K, Leung VCM. Blockchain-based decentralized trust management in vehicular networks. IEEE Internet Things J 2018;6(2):1495–505.
[39] Su Z, Wang Y, Xu Q, Zhang N. LVBS: lightweight vehicular blockchain for secure data sharing in disaster rescue. IEEE Trans Dependable Secur Comput. In press.
[40] Lin X, Wu J, Mumtaz S, Garg S, Li J, Guizani M. Blockchain-based on-demand computing resource trading in IoV-assisted smart city. IEEE Trans Emerg Top Comput. In press.
[41] Li C, Fu Y, Yu FR, Luan TH, Zhang Y. Vehicle position correction: a vehicular blockchain networks-based GPS error sharing framework. IEEE Trans Intell Transp Syst 20201;22(2):898–912.
[42] Qian LP, Wu Y, Xu X, Ji B, Shi Z, Jia W. Distributed charging-record management for electric vehicle networks via blockchain. IEEE Internet Things J 2021;8(4):2150–62.
[43] Yang H, Liang Y, Yuan J, Yao Q, Yu A, Zhang J. Distributed blockchain-based trusted multidomain collaboration for mobile edge computing in 5G and beyond. IEEE Trans Ind Inform 2020;16(11):7094–104.
[44] Yang H, Yuan J, Yao H, Yao Q, Yu A, Zhang J. Blockchain-based hierarchical trust networking for JointCloud. IEEE Internet Things J 2020;7(3):1667–77.
[45] Xu Y, Zhang C, Wang G, Qin Z, Zeng Q. A blockchain-enabled deduplicatable data auditing mechanism for network storage services. IEEE Trans Emerg Top Comput. In press.
[46] Zhu L, Wu Y, Gai K, Choo KKR. Controllable and trustworthy blockchain-based cloud data management. Future Gener Comput Syst 2019;91:527–35.
[47] Chen L, Lee WK, Chang CC, Choo KKR, Zhang N. Blockchain based searchable encryption for electronic health record sharing. Future Gener Comput Syst 2019;95:420–9.
[48] Zhang Y, Xu C, Cheng N, Li H, Yang H, Shen X. Chronos^{++}: an accurate blockchain-based time-stamping scheme for cloud storage. IEEE Trans Serv Comput 2020;13(2):216–29.
[49] Zhang Y, Xu C, Lin X, Shen X. Blockchain-based public integrity verification for cloud storage against procrastinating auditors. IEEE Trans Cloud Comput 2021;9(3):92337.
[50] Liu Y, He D, Obaidat MS, Kumar N, Khan MK, Choo KKR. Blockchain-based identity management systems: a review. J Netw Comput Appl 2020;166:102731.

[51] Wang J, Wu L, Choo KKR, He D. Blockchain-based anonymous authentication with key management for smart grid edge computing infrastructure. IEEE Trans Ind Inform 2020;16(3):1984–92.
[52] Shen M, Liu H, Zhu L, Xu K, Yu H, Du X, et al. Blockchain-assisted secure device authentication for cross-domain industrial IoT. IEEE J Sel Areas Commun 2020;38(5):942–54.
[53] Yang M, Zhu T, Liang K, Zhou W, Deng RH. A blockchain-based location privacy-preserving crowdsensing system. Future Gener Comput Syst 2019;94:408–18.
[54] Tosh D, Shetty S, Liang X, Kamhoua C, Njilla LL. Data provenance in the cloud: a blockchain-based approach. IEEE Consum Electron Mag 2019;8(4):38–44.
[55] Rahman MS, Omar AAL, Bhuiyan MZA, Basu A, Kiyomoto S, Wang G. Accountable cross-border data sharing using blockchain under relaxed trust assumption. IEEE Trans Eng Manag 2020;67(4):1476–86.
[56] Liang W, Fan Y, Li KC, Zhang D, Gaudiot JL. Secure data storage and recovery in industrial blockchain network environments. IEEE Trans Ind Inform 2020;16 (10):6543–52.
[57] Gilani K, Bertin E, Hatin J, Crespi N. A survey on blockchain-based identity management and decentralized privacy for personal data. In: Proceedings of 2020 2nd Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS); 2020 Sep 28–30; Paris, France; 2020. p. 97–101.
[58] Patsonakis C, Samari K, Roussopoulos M, Kiayias A. Towards a smart contract-based, decentralized, public-key infrastructure. In: Proceedings of International Conference on Cryptology and Network Security; 2017 Nov 30–Dec 2; Hong Kong, China; 2017. p. 299–321.
[59] Wang Z, Lin J, Cai Q, Wang Q, Zha D, Jing J. Blockchain-based certificate transparency and revocation transparency. IEEE Trans Dependable Secur Comput. In press.
[60] Kubilay MY, Kiraz MS, Mantar HA. CertLedger: a new PKI model with certificate transparency based on blockchain. Comput Secur 2019;85:333–52.
[61] Xu R, Joshi J. Trustworthy and transparent third-party authority. ACM Trans Internet Technol 2020;20(4):31.
[62] Chen J, Yao S, Yuan Q, He K, Ji S, Du R. CertChain: public and efficient certificate audit based on blockchain for TLS connections. In: Proceedings of IEEE INFOCOM 2018; 2018 Apr 15–19; Honolulu, HI, USA; 2018. p. 2060–8.
[63] Kondova G, Erbguth J. Self-sovereign identity on public blockchains and the GDPR. In: Proceedings of the 35th Annual ACM Symposium on Applied Computing; 2020 Mar 30–Apr 3; 2020. p. 342–5.
[64] Fan K, Pan Q, Zhang K, Bai Y, Sun S, Li H, et al. A secure and verifiable data sharing scheme based on blockchain in vehicular social networks. IEEE Trans Veh Technol 2020;69(6):5826–35.
[65] Yu G, Zha X, Wang X, Ni W, Yu K, Yu P, et al. Enabling attribute revocation for fine-grained access control in blockchain–IoT systems. IEEE Trans Eng Manag 2020;67(4):1213–30.
[66] Hu Y, Kumar S, Popa RA. Ghostor: toward a secure data-sharing system from decentralized trust. In: Proceedings of NSDI; 2020 Feb 25–27; Santa Clara, CA, USA; 2020. p. 851–77.
[67] Yuen TH, Sun SF, Liu JK, Au MH, Esgin MF, Zhang Q, et al. RingCT 3.0 for blockchain confidential transaction: shorter size and stronger security. In: Bonneau J, Heninger N, editors. Financial cryptography and data security. Cham: Springer; 2020. p. 464–83.
[68] Hardjono T, Pentland A. Verifiable anonymous identities and access control in permissioned blockchains. 2019. arXiv:1903.04584.
[69] Camenisch J, Drijvers M, Dubovitskaya M. Practical UC-secure delegatable credentials with attributes and their application to blockchain. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security; 2017 Oct 30–Nov 3; Dallas, TX, USA; 2017. p. 683–99.
[70] Sonnino A, Al-Bassam M, Bano S, Meiklejohn S, Danezis G. Coconut: threshold issuance selective disclosure credentials with applications to distributed ledgers. 2018. arXiv:1802.07344.
[71] Yu Y, Zhao Y, Li Y, Du X, Wang L, Guizani M. Blockchain-based anonymous authentication with selective revocation for smart industrial applications. IEEE Trans Ind Inform 2020;16(5):3290–300.
[72] Gao S, Piao G, Zhu J, Ma X, Ma J. TrustAccess: a trustworthy secure ciphertext-policy and attribute hiding access control scheme based on blockchain. IEEE Trans Veh Technol 2020;69(6):5784–98.
[73] Zhou L, Du S, Zhu H, Chen C, Ota K, Dong M. Location privacy in usage-based automotive insurance: attacks and countermeasures. IEEE Trans Inf Forensics Secur 2018;14(1):196–211.
[74] Frankle J, Park S, Shaar D, Goldwasser S, Weitzner D. Practical accountability of secret processes. In: Proceedings of the 27th USENIX Security Symposium; 2018 Aug 15–17; Baltimore, MD, USA; 2018. p. 657–74.
[75] Antignac T, Scandariato R, Schneider G. A privacy-aware conceptual model for handling personal data. In: Margaria T, Steffen B, editors. Leveraging applications of formal methods, verification and validation: foundational techniques. Cham: Springer; 2016. p. 942–57.
[76] Labadie C, Legner C. Understanding data protection regulations from a data management perspective: a capability-based approach to EU-GDPR. In: Proceedings of the 14th International Conference on Wirtschaftsinformatik; 2019 Feb 24–27; Siegen, Germany; 2019. p. 1292–306.
[77] Barati M, Rana O, Theodorakopoulos G, Burnap P. Privacy-aware cloud ecosystems and GDPR compliance. In: Proceedings of 2019 7th International Conference on Future Internet of Things and Cloud; 2019 Aug 26–28; Istanbul, Turkey; 2019. p. 117–24.
[78] Corrales M, Jurčys P, Kousiouris G. Smart contracts and smart disclosure: coding a GDPR compliance framework. In: Corrales M, Fenwick M, Haapio H, editors. Legal tech, smart contracts and blockchain. Singapore: Springer Nature Singapore Pte Ltd.; 2019. p. 189–220.
[79] Bowe S, Chiesa A, Green M, Miers I, Mishra P, Wu H. ZEXE: enabling decentralized private computation. In: Proceedings of 2020 IEEE Symposium on Security and Privacy (SP); 2020 May 18–21; San Francisco, CA, USA; 2020. p. 947–64.
[80] Ma Z, Wang X, Jain DK, Khan H, Gao H, Wang Z. A blockchain-based trusted data management scheme in edge computing. IEEE Trans Ind Inform 2019;6 (3):2013–21.
[81] Parno B, Howell J, Gentry C, Raykova M. Pinocchio: nearly practical verifiable computation. In: Proceedings of 2013 IEEE Symposium on Security and Privacy; 2013 May 19–22; Berkeley, CA, USA; 2013. p. 238–52.
[82] Maller M, Bowe S, Kohlweiss M, Meiklejohn S. Sonic: zero-knowledge SNARKs from linear-size universal and updatable structured reference strings. In: Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security; 2019 Nov 11–15; London, UK; 2019. p. 2111–28.
[83] Costan V, Devadas S. Intel SGX explained. 2016. Cryptology ePrint Archive:86.
[84] Cheng R, Zhang F, Kos J, He W, Hynes N, Johnson N, et al. Ekiden: a platform for confidentiality-preserving, trustworthy, and performant smart contracts. In: Proceedings of IEEE European Symposium on Security an Privacy; 2019 Jun 17–19; Stockholm, Sweden; 2019. p. 185–200.
[85] Ayoade G, Karande V, Khan L, Hamlen K. Decentralized IoT data management using blockchain and trusted execution environment. In: Proceedings of IEEE International Conference on Information Reuse and Integration (IRI); 2018 Jul 6–9; Salt Lake City, UT, USA; 2018. p. 15–22.
[86] Pass R, Shi E, Tramèr F. Formal abstractions for attested execution secure processors. In: Coron JS, Nielsen JB, editors. Advances in cryptology—EUROCRYPT 2017; Cham: Springer; 2017. p. 260–89.
[87] Dong C, Wang Y, Aldweesh A, McCorry P, van Moorsel A. Betrayal, distrust, and rationality: smart counter-collusion contracts for verifiable cloud computing. In: Proceedings of 2017 ACM SIGSAC Conference on Computer and Communications Security; 2017 Oct 30–Nov 3; Dallas, TX, USA; 2017. p. 211–27.
[88] Brewster C, Nouwt B, Raaijmakers S, Verhoosel J. Ontology-based access control for FAIR data. Data Intell 2020;2(1–2):66–77.
[89] Bhaskaran K, Ilfrich P, Liffman D, Vecchiola C, Jayachandran P, Kumar A, et al. Double-blind consent-driven data sharing on blockchain. In: Proceedings of 2018 IEEE International Conference on Cloud Engineering (IC2E); 2018 Apr 17–20; Orlando, FL, USA; 2018. p. 385–91.
[90] Li H, Pei L, Liao D, Chen S, Zhang M, Xu D. FADB: a fine-grained access control scheme for VANET data based on blockchain. IEEE Access 2020;8:85190–203.
[91] Koutsos V, Papadopoulos D, Chatzopoulos D, Tarkoma S, Hui P. Agora: a privacy-aware data marketplace. 2020. Cryptology ePrint Archive:865.
[92] Liu D, Alahmadi A, Ni J, Lin X, Shen X. Anonymous reputation system for IIoT-enabled retail marketing atop pos blockchain. IEEE Trans Ind Inform 2019;15 (6):3527–37.
[93] Lone AH, Mir RN. Reputation driven dynamic access control framework for IoT atop PoA ethereum blockchain. 2020. Cryptology ePrint Archive:566.
[94] Xu C, Wang K, Li P, Guo S, Luo J, Ye B, et al. Making big data open in edges: a resource-efficient blockchain-based approach. IEEE Trans Parallel Distrib Syst 2019;30(4):870–82.
[95] Makhdoom I, Zhou I, Abolhasan M, Lipman J, Ni W. PrivySharing: a blockchain-based framework for privacy-preserving and secure data sharing in smart cities. Comput Secur 2020;88:101653.
[96] Truong NB, Sun K, Lee GM, Guo Y. GDPR-compliant personal data management: a blockchain-based solution. IEEE Trans Inf Forensics Secur 2020;15:1746–61.
[97] Zheng X, Mukkamala RR, Vatrapu R, Ordieres-Mere J. Blockchain-based personal health data sharing system using cloud storage. In: Proceedings of IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom); 2018 Sep 17–20; Ostrava, Czech Republic; 2018. p. 1–6.
[98] Zheng BK, Zhu LH, Shen M, Gao F, Zhang C, Li YD, et al. Scalable and privacy-preserving data sharing based on blockchain. J Comput Sci Technol 2018;33 (3):557–67.
[99] Gunasinghe H, Kundu A, Bertino E, Krawczyk H, Chari S, Singh K, et al. PrividEx: privacy preserving and secure exchange of digital identity assets. In: Proceedings of the World Wide Web Conference; 2019 May 13–17; San Francisco, CA, USA; 2019. p. 594–604.
[100] Dai W, Dai C, Choo KKR, Cui C, Zou D, Jin H. SDTE: a secure blockchain-based data trading ecosystem. IEEE Trans Inf Forensics Secur 2019;15:725–37.
[101] Schuster F, Costa M, Fournet C, Gkantsidis C, Peinado M, Mainar-Ruiz G, et al. VC3: trustworthy data analytics in the cloud using SGX. In: Proceedings of 2015 IEEE Symposium on Security and Privacy; 2015 May 17–21; San Jose, CA, USA; 2015. p. 38–54.
[102] Dziembowski S, Eckey L, Faust S. Fairswap: how to fairly exchange digital goods. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security; 2018 Oct 15–19; Toronto, ON, Canada; 2018. p. 967–84.
[103] Liu X, Sun SX, Huang G. Decentralized services computing paradigm for blockchain-based data governance: programmability, interoperability, and intelligence. IEEE Trans Serv Comput 2019;13(2):343–55.

[104] Weng J, Weng J, Zhang J, Li M, Zhang Y, Luo W. DeepChain: auditable and privacy-preserving deep learning with blockchain-based incentive. IEEE Trans Dependable Secur Comput 2021;18(5):2438–55.
[105] Hu S, Cai C, Wang Q, Wang C, Luo X, Ren K. Searching an encrypted cloud meets blockchain: a decentralized, reliable and fair realization. In: Proceedings of IEEE INFOCOM 2018; 2018 Apr 16–19; Honolulu, HI, USA; 2018. p. 792–800.
[106] Nguyen K, Ghinita G, Naveed M, Shahabi C. A privacy-preserving, accountable and spam-resilient geo-marketplace. In: Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information System; 2019 Nov 5–8; Chicago, IL, USA; 2019. p. 299–308.
[107] Zhang Y, Genkin D, Katz J, Papadopoulos D, Papamanthou C. A zero-knowledge version of VSQL. 2017. Cryptology ePrint Archive:1146.
[108] Neisse R, Steri G, Nai-Fovino I. A blockchain-based approach for data accountability and provenance tracking. In: Proceedings of the 12th International Conference on Availability, Reliability and Security; 2017 Aug 29–Sep 1; Reggio Calabria, Italy; 2017. p. 1–10.
[109] Cucurull J, Puiggalí J. Distributed immutabilization of secure logs. In: Barthe G, Markatos E, Samarati P, editors. Security and trust management. Cham: Springer; 2016. p. 122–37.
[110] Ruan P, Chen G, Dinh TTA, Lin Q, Ooi BC, Zhang M. Fine-grained, secure and efficient data provenance on blockchain systems. In: Proceedings of the 45th International Conference on Very Large Data Bases; 2019 Aug 26–30; Los Angles, CA, USA; 2019. p. 975–88.
[111] Tang YR, Xing Z, Xu C, Chen J, Xu J. Lightweight blockchain logging for data-intensive applications. In: Zohar A, Eyal I, Teague V, Clark J, Bracciali A, Pintore F, et al., editors. Financial cryptography and data security. Berlin: Springer Verlag GmbH; 2018. p. 308–24.
[112] Ahmad A, Saad M, Njilla L, Kamhoua C, Bassiouni M, Mohaisen A. BlockTrail: a scalable multichain solution for blockchain-based audit trails. In: Proceedings of 2019 IEEE International Conference on Communication (ICC); 2019 May 20–24; Shanghai, China; 2019.
[113] Liu D, Ni J, Huang C, Lin X, Shen XS. Secure and efficient distributed network provenance for IoT: a blockchain-based approach. IEEE Internet Things J 2020;7(8):7564–74.
[114] Tomescu A, Bhupatiraju V, Papadopoulos D, Papamanthou C, Triandopoulos N, Devadas S. Transparency logs via append-only authenticated dictionaries. In: Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security; 2019 Nov 11–15; London, UK; 2019. p. 1299–316.
[115] Ding S, Cao J, Li C, Fan K, Li H. A novel attribute-based access control scheme using blockchain for IoT. IEEE Access 2019;7:38431–41.
[116] Matetic S, Wüst K, Schneider M, Kostiainen K, Karame G, Capkun S. BITE: bitcoin lightweight client privacy using trusted execution. In: Proceedings of the 28th USENIX Conference on Security Symposium, 2019 Aug 14–16; Santa Clara, CA, USA; 2019. p. 783–800.
[117] Chepurnoy A, Papamanthou C, Zhang Y. Edrax: a cryptocurrency with stateless transaction validation. 2018. Cryptology ePrint Archive:968.
[118] Jiang Y, Wang C, Wang Y, Gao L. A cross-chain solution to integrating multiple blockchains for IoT data management. Sensors 2019;19(9):2042.
[119] Maram SKD, Zhang F, Wang L, Low A, Zhang Y, Juels A, et al. CHURP: dynamic-committee proactive secret sharing. In: Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security; 2019 Nov 11–15; London, UK; 2019. p. 2369–86.
[120] Campanelli M, Fiore D, Querol A. LegoSNARK: modular design and composition of succinct zero-knowledge proofs. In: Proceedings of ACM SIGSAC Conference on Computer and Communications Security; 2019 Nov 11–15; London, UK; 2019. p. 2075–92.
[121] Lim SY, Fotsing PT, Almasri A, Musa O, Kiah MLM, Ang TF, et al. Blockchain technology the identity management and authentication service disruptor: a survey. Int J Adv Sci Eng Inf Technol 2018;8:1735–45.
[122] Canetti R. Universally composable security: a new paradigm for cryptographic protocols. In: Proceedings of 42nd IEEE Symposium on Foundations of Computer Science; 2001 Oct 8–11; Newport Beach, CA, USA; 2001. p. 136–45.