Blockchain and Machine Learning for Communications and Networking Systems

Yiming Liu, F. Richard Yu, *Fellow, IEEE*, Xi Li, Hong Ji, *Senior Member, IEEE*, and Victor C. M. Leung, *Fellow, IEEE*

***Abstract*—Recently, with the rapid development of information and communication technologies, the infrastructures, resources, end devices, and applications in communications and networking systems are becoming much more complex and heterogeneous. In addition, the large volume of data and massive end devices may bring serious security, privacy, services provisioning, and network management challenges. In order to achieve decentralized, secure, intelligent, and efficient network operation and management, the joint consideration of blockchain and machine learning (ML) may bring significant benefits and have attracted great interests from both academia and industry. On one hand, blockchain can significantly facilitate training data and ML model sharing, decentralized intelligence, security, privacy, and trusted decision-making of ML. On the other hand, ML will have significant impacts on the development of blockchain in communications and networking systems, including energy and resource efficiency, scalability, security, privacy, and intelligent smart contracts. However, some essential open issues and challenges that remain to be addressed before the widespread deployment of the integration of blockchain and ML, including resource management, data processing, scalable operation, and security issues. In this paper, we present a survey on the existing works for blockchain and ML technologies. We identify several important aspects of integrating blockchain and ML, including overview, benefits, and applications. Then we discuss some open issues, challenges, and broader perspectives that need to be addressed to jointly consider blockchain and ML for communications and networking systems.**

***Index Terms*—Blockchain, machine learning (ML), distributed ledger technology (DLT), wireless communications, wireless networks.**

Manuscript received December 24, 2018; revised June 6, 2019, October 2, 2019, and January 9, 2020; accepted February 15, 2020. Date of publication February 24, 2020; date of current version May 28, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 61671088 and Grant 61771070, in part by the Chinese National Engineering Laboratory for Big Data System Computing Technology, and in part by the Canadian Natural Sciences and Engineering Research Council. *(Corresponding author: F. Richard Yu.)*

Yiming Liu, Xi Li, and Hong Ji are with the Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: liuyiming@bput.edu.cn; lixi@bupt.edu.cn; jihong@bupt.edu.cn).

F. Richard Yu is with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada (e-mail: richard.yu@carleton.ca).

Victor C. M. Leung is with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China, and also with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: vleung@ieee.org).

Digital Object Identifier 10.1109/COMST.2020.2975911

I. Introduction

RECENTLY, the ubiquitous proliferation of end devices and popular applications brings a huge burden on existing networks. The upcoming fully decentralized and intelligent communications and networking systems for providing innovative secure services trigger the investigation of prospective technologies. Among them, blockchain and machine learning (ML) are considered as promising technologies to support secure and decentralized sharing of data and model as well as the intelligent network operation and management. In this paper, we focus on blockchain and ML, which have a significant potential to promote the development of communications and networking systems.

Blockchain, which is the underlying technique of the digital cryptocurrency, has attracted widespread attention from both academia and industry [1]. Blockchain is essentially a distributed ledger, which is maintained by the network participants in a logical Peer-to-Peer (P2P) network. It establishes a trust paradigm among people and machines and enables applications to be operated without any central controller or any intermediary. Generally, blockchain is included in a number of distributed ledger technologies, which adopt a set of mechanisms for recording and sharing transactions and data across multiple nodes in a decentralized manner. Nevertheless, the structures of distributed ledgers are not only a chain of blocks but also other structures (e.g., directed acyclic graph (DAG) [2]). For simplicity, here we use the term "blockchain technology" to indicate the general classes of distributed ledgers based on the community consensus.

Blockchain has provided a range of promising opportunities for a variety of applications and scenarios by efficiently developing P2P platforms for sharing information [3], [4], enhancing the enforcement of governance [5], [6], and increasing utilization of resources [7], [8]. According to McKinsey, blockchain technology may reach its full potential in any application that uses a centralized solution within the next five years based on its current pace of evolution [9]. Especially, blockchain-based solutions have changed communications and networking systems thanks to their key features compared to traditional solutions. They provide a feasible solution for the dynamic access control, the integrity and validity of exchanged data, and the privacy of mobile users. For example, such a decentralized and distributed blockchain can be applied in ad hoc networks, where smart end devices connect without any central base stations. They may also be applied in fog or cloud radio access networks for maintaining tight synchronization

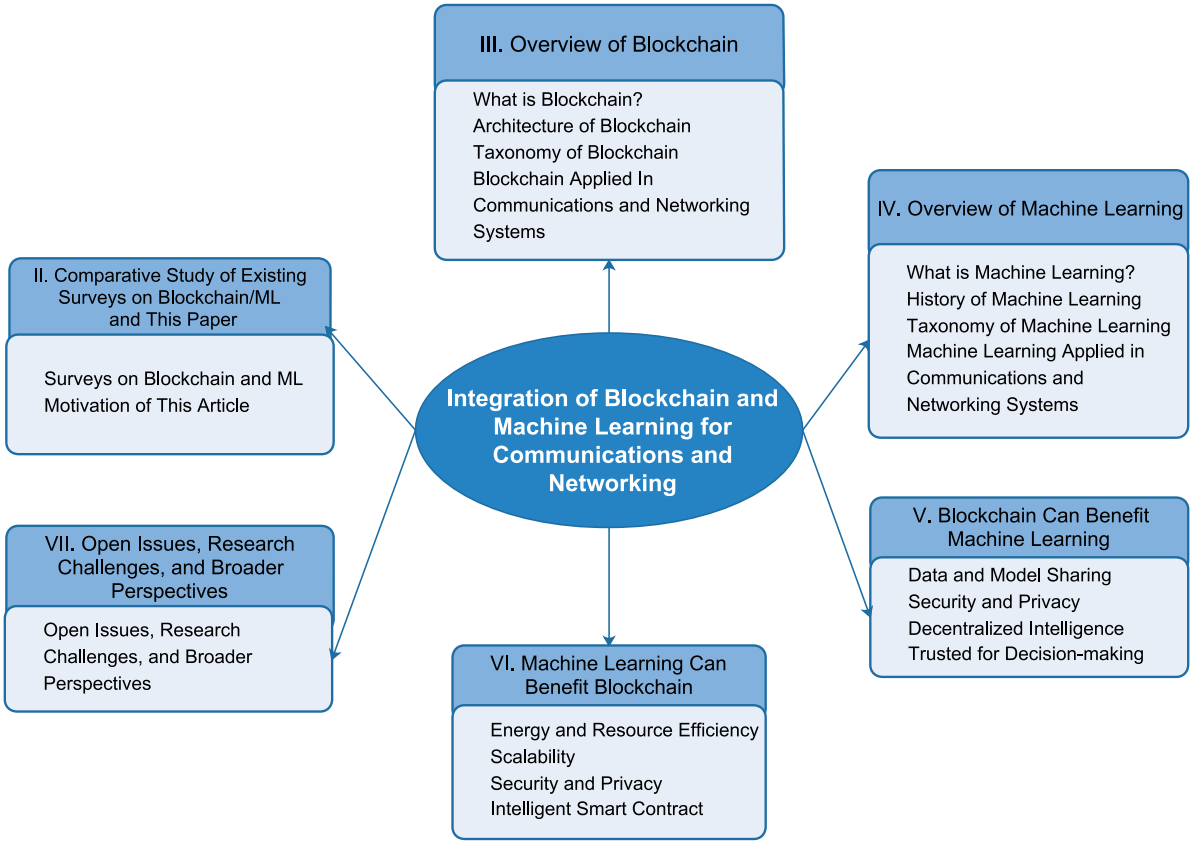


Fig. 1. Outline of the integration of blockchain and machine learning for communications and networking systems.

among different network elements equipped with computing and networking resources [10]. Researchers in [11]–[14] have given the systematic overviews on blockchain and shown the wide-range applications of blockchain. It has been successfully used in a variety of applications, such as vehicle ad hoc networks (VANET) [15], [16], Internet of Things (IoT) [17], [18], healthcare systems [4], [19], content-centric networks [20], [21], reputation systems [22], [23], and security services [24], [25].

On the other hand, ML, which is a sub-domain of artificial intelligence (AI), possesses powerful data processing capabilities and provides feasible solutions to a variety of problems. Current networks typically involve a large number of network elements and end devices. They may produce massive data that can be analyzed to optimize the system performance by ML solutions. Moreover, multi-layer and multi-vendor communication systems become very complicated and need to be managed by more efficient solutions. Compared with conventional techniques, ML-based techniques, which train models and learn from large data sets, can substantially get rid of human intervention and reduce the operation cost for handling large-scale complex systems. Generally, ML techniques are categorized into four classifications: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning (RL). These ML algorithms have been extensively investigated to optimize and enhance various systems and networks in complicated scenarios, such as image recognition [26], [27], virtual reality (VR), [28], [29], and unmanned aerial vehicle (UAV) networks [30], [31].

A lot of researches have shown that the integration of blockchain and ML is powerful and is set to transform many technical sectors, including communications and networking systems [32]–[34]. On one hand, by applying blockchain into ML systems, ML can use the information collected and manipulated by blockchains to sort through problems more quickly than ever before. The authors in [35] propose a blockchain-based incentive mechanism, where large-scale smart end devices in wireless networks are encouraged to collect a large amount of sensing data efficiently. These collected data could be used in ML solutions for learning and training models. Also, blockchain-based ML solutions can process data in a distributed manner, without being restricted to a single entity data set. *HyperNet* [36], which is a decentralized trusted computing and networking paradigm, is proposed to satisfy the requirements for data controlling and sharing. The main features of blockchain (i.e., security and trustiness) also enable ML solutions applied in communications and networking systems to become more trustful [37]. On the other hand, by leveraging ML techniques, blockchain applications could have powerful intelligence for data processing and the execution of computationally intensive applications [38]–[40]. Moreover, ML-based blockchain mechanisms can process various types of transactions and make effective operation possible, especially implementing ML in smart contracts. To expand and diversify the capability of smart

contracts, *Cortex* [41] introduces AI into smart contracts to enable AI functionalities to constrcut intelligent smart contracts and enhanced decentralized applications (DAPPs) with more diverse use cases.

Even though there are a lot of advantages of blockchain and ML, several significant research challenges need to be well investigated before the widespread deployment of the integration of blockchain and ML, including resource management, big data processing, scalability, security and privacy. Big data processing, which makes the design and analysis of the integration of blockchain and ML even more difficult, should be well addressed in future research. Particularly, unlike traditional centralized technologies, a fundamental challenge of the integration of blockchain and ML is how to deal with the scalability issue, especially for large-scale complex systems. Furthermore, resource management, as well as security and privacy also play an important role in enhancing the performance of the integration of blockchain and ML for communications and networking systems.

In this paper, we provide a survey on some of the works that have already been carried out on integrating blockchain and ML for communications and networking systems, and then discuss some related open issues and challenges. An outline of our paper towards the integration of blockchain and ML is given in Fig. 1. As shown in the figure, we identify five aspects of the integration of blockchain and ML, on which we focus: a comparative study of this paper with existing surveys on blockchain and ML, overviews of blockchain and ML, benefits of blockchain and ML for each other, open issues, challenges, and broader perspectives.

To the best of our knowledge, the interrelationship between blockchain and ML for communications and networking systems has not been well addressed in previous works. In essence, the unique characteristics associated with blockchain and ML present a variety of challenges beyond the existing works. The initial steps we take here may help understand how to embed blockchain to improve the performance of ML applications, and how to make blockchain work more effectively through ML for communications and networking systems.

The rest of the paper is organized as follows. Section II provides a comparative study of existing blockchain/ML surveys and this paper. Section III presents an overview of blockchain, including background, structures, taxonomies, and typical applications in communications and networking systems. Section IV presents an overview of the background, history, taxonomies, and various applications of ML technologies in communication and networking systems. In Section V, we explore how blockchain can benefit ML. In Section VI, we discuss that ML can benefit blockchain. Some significant open issues, research challenges, and broader perspectives are addressed in Section VII. Finally, we conclude this paper in Section VIII.

II. Comparative Study of This Paper With Existing Surveys on Blockchain/ML

To give a clear description of the key distributions of this paper, we first briefly discuss the existing surveys on blockchain and ML. Then we illustrate our motivation of this paper through the analysis of limitation of blockchain and ML as well as the benefits of the integration of them.

A. Existing Surveys on Blockchain and Machine Learning

We first introduce several existing surveys investigating the blockchain techniques from different perspectives. The authors in [14] introduce blockchain technology, including the key requirements, evolution, types, consensus mechanisms, and existing blockchain platforms. The authors in [42] investigate the decentralized consensus mechanisms in blockchain systems, such as Byzantine Fault Tolerance (BFT)-based consensus protocols, Nakamoto protocols, virtual mining protocols, hybrid protocols and a series of parallel consensus protocols. The authors in [13] provide a comprehensive survey on the network layer of permissionless blockchains. Then, considering system performance and requirements, such as low cost of participation, anonymity, and topology hiding, they present a design of the network layer of permissionless blockchains to improve the efficiency and security of systems. The authors in [43] investigate blockchain-based approaches in terms of security issues, such as authentication, confidentiality, privacy and access control, data and resource provenance, and integrity assurance. Moreover, thanks to the key characteristics of blockchains like decentralization, tamper-resistance, and security, many researchers have deeply investigated the applications of blockchain and give comprehensive surveys on various scenarios, such as industrial applications [44], energy trading markets [45], smart cities [46], edge computing systems [47], IoT [11]. The application of blockchain brings a lot of benefits in these scenarios, such as increased efficiency and security, enhanced traceability and transparency, and low costs.

On the other hand, many researchers have widely surveyed ML techniques in communications and networking systems. The authors in [48] investigate ML algorithms applied in self-organizing cellular networks. To leverage more robust and intelligent algorithms, they also provide a comparison between the commonly ML algorithms and certain self-organizing networks (SON) metrics for selecting proper ML algorithm for each SON function. The applications of ML techniques have been also comprehensively investigated and surveyed in other network scenarios, such as wireless sensor networks [49], optical communications and networking [50], cognitive radio networks [51], software-defined networks (SDN) [52]. Deep learning (DL) is a powerful tool to add intelligence to communication and networking systems with large-scale devices and complicated network environment. The authors in [53] investigate the applications of DL algorithms for several network layers, such as physical layer modulation/coding, data link layer access control/resource allocation, routing layer path search and traffic balancing. The authors in [54] discuss the DL algorithms applied in IoT for achieving the desired analytics in IoT data and applications. Motivated by the successful applications of ML techniques in many practical tasks like image recognition, a lot of researchers have investigated the applications of ML for specific areas in wireless communication

TABLE I
COMPARATIVE STUDY OF THIS PAPER WITH EXISTING SURVEYS ON BLOCKCHAIN/ML

Subject	Ref.	Use case	Contributions
Blockchain	[11]	IoT	Blockchain techniques for providing a decentralized, secure, and auditable IoT applications
	[13]	Permissionless blockchains	Discussing security issues of permissionless blockchains on the network layer and providing the design of the network layer of permissionless blockchains
	[14]	Blockchain systems	Introducing the blockchain technology, including the key requirements, evolution, types, consensus mechanisms, and existing blockchain platforms.
	[42]	Blockchain networks	Addressing the perspective of building distributed consensus system and incentive mechanism in blockchain networks
	[43]	Security Services	Discussing blockchain-based approaches for several security services, such as authentication, confidentiality, privacy, access control, data and resource provenance, and integrity assurance
	[44]	Industries	Exploring the opportunities, advantages, and open issues of incorporating blockchain in different industrial applications
	[45]	Energy trading markets	Providing a detailed review about the deployment of decentralized transactive energy systems based on blockchains.
	[46]	Smart cities	Investigating blockchain applied in smart cities from different aspects, such as smart citizen, smart healthcare, smart grid, smart transportation, supply chains.
	[47]	Edge computing systems	Integrating blockchain and edge computing for secure access and control of the network, storage, and computation distributed at the edges.
Machine Learning	[48]	Self-organizing cellular networks	ML algorithms applied in self-organizing cellular networks for enabling a fully autonomous and flexible network
	[49]	Wireless sensor networks	ML algorithms applied in wireless sensor networks for adapting dynamic conditions
	[50]	Optical networks	Providing a comprehensive survey on the ML applied to optical communications and networking
	[51]	Cognitive radio network	Discussing the ML algorithms applied to cognitive radio networks
	[52]	Software defined networking	Providing ML algorithms applied to SDN from the different aspects, such as traffic classification, routing optimization, resource management, and security
	[53]	Wireless networks	Performing a survey on the applications of DL algorithms for different network layers
	[54]	IoT	Discussing the emerging DL techniques for IoT data analytics
	[55]	Network traffic classification	Providing a systematic review of ML-based traffic classification and the requirements of ML algorithms in the traffic classification field.
	[56]	Internet traffic classification	Discussing the applications of ML techniques to IP traffic classification with DM techniques
	[57]	Network traffic control systems	Providing DL applications for various network traffic control scenarios
	[58]	Mobile and wireless networking	Discussing the DL solutions applied to mobile systems
	[59]	Cyber security systems	Discussing ML and DM methods in intrusion detection applications for cyber analytics
	[60]	Intrusion detection	Providing detailed ML techniques in detecting intrusive activities
Blockchain and ML	[61]	AI applications	Investigating blockchain for AI applications
	Our work	Communications and networks	Integrating blockchain and ML in communications and networking systems

and networks. The authors in [55], [56] introduce systematic reviews on ML-based traffic classification and discuss the advantages, disadvantages, and necessary requirements that apply ML in the traffic classification. The authors in [57], [58] provide overviews of the current DL algorithms for various network traffic control aspects. To address the security issues, the authors in [59] provide a survey on ML and data mining (DM) methods applied in intrusion detection system (IDS), which is a critical component for monitoring various threats. They also address the complexity of ML/DM algorithms and the challenges for applying ML and DM technologies for cybersecurity systems. The authors in [60] also provide a comprehensive discussion of several ML techniques in detecting intrusive activities.

Since blockchain and ML are considered as two of the most promising and powerful technologies, there are some research works on integration of blockchain and ML in various applications. The authors in [61] provide a brief survey on blockchain applications for AI, which includes ML and other intelligent techniques. They discuss the existing blockchain applications, platforms, and protocols targeting AI area. However, this survey mainly address some issues that blockchain for improving AI performance and did not address the advancement of AI techniques to blockchain. Although blockchain and ML techniques have been applied in various domains, no existing works focus on the integration of blockchain and ML in the domain of communications and networking systems. Thus, in this paper, we provide a comprehensive survey on the integration of blockchain and ML techniques for communications and networking systems. Table I shows a brief comparison of this paper with existing surveys on blockchain and ML.

B. Motivations of Integrating Blockchain and Machine Learning in Communications and Networking Systems

Although blockchain and ML are promising technologies applied in communications and networking systems, there are still a lot of open issues and problems. In this subsection, we first investigate the limitations of blockchain and ML applied in communications and networking systems. Then, we give a brief overview of the benefits of blockchain and ML for each other in communications and networking systems.

1) The Limitation of Blockchain Applied in Communications and Networking Systems: Despite the enormous potential of blockchain, there are still some open issues that restrict its widespread application in communications and networking systems. The trilemma points out that blockchain systems can only at most have two of the three properties (scalability, decentralization, and security) [62]. Specifically, decentralization enables the system to be fault tolerant, attack and collision resistant, security guarantees the immutability and the attacks resistant, and scalability addresses the ability to process transactions.

With the ever-increasing number of transactions, the scalability issue becomes a severe bottleneck and limits the practical development of blockchain. A lot of blockchain applications such as Bitcoin and Ethereum process on average 7-20 transactions per second and consume huge energy resources. To address the scalability issues, increasing the block size, sharding, pruning, and off-chain are proposed for enabling practical applications (As discussed in Section VI-B). However, these methods still have their respective problems, for example, where to execute the off-chain and which transactions to be pruned. In addition, building and managing the communications and networking systems based on blockchain usually need to consume much time and resources. Furthermore, although blockchain is almost impossible to hack, its further layers and applications are not secure. With the increasing number of personal data stored in blockchain-based communication systems, privacy leakage becomes another critical issue. Current blockchain applications often require transactions and smart contracts to produce metadata, which may be used to disclose some information, even if the data itself is encrypted.

2) The Limitation of Machine Learning Applied in Communications and Networking Systems: ML technologies bring significant benefits across a range of fields and applications, however, there are still some challenges to enable the intelligent and flexible managements of complex communications and networking systems, which involve massive users with diversified quality of service (QoS) requirements.

The ML-based solutions usually require large amounts of training data, which is often implemented at a centralized network controller with sufficient storage resource and computational capability for information collection. However, accessing a large number of data for training ML models may be not available in current communication systems. Besides, aggregating data in heterogeneous networks for ML training is also a challenge. To deal with these issues, a distributed ML architecture where a central server controls the process of training model and the distributed users train the models with their datasets independently is proposed [63]. Nevertheless, the inefficient management of data sharing and the overhead of information exchange among distinct multi-parties are crucial bottlenecks of the development of ML technique.

Both security and privacy preservation are critical factors for using the ML-based methods in communications and networking systems. Currently, ML systems often adopt a centralized architecture, which is prone to hacking, as any malicious node only needs to break into a single system to manipulate the instructions. Since the training data usually involves a large amount of personal information, the data breaches may lead to the privacy concerns of personal data. How to protect the training data put into ML models from hackers is an important issue. In addition, the centralized ML system, where a single entity needs to have a global vision of the resources for collecting large datasets, creating highly sophisticated models and solving various types of tasks, are often non-scalability, ineffective operation, and easy to suffer failure problems. The centralized ML systems are inadequate for communications and networking systems evolving large-scale complicated parameters. Furthermore, the lack of trust would severely limits the performance of ML solutions. To address this issue and make sure the ML algorithms work effectively, the trust mechanisms and audit processes should be well designed for ML solutions.

3) The Benefits of Blockchain and Machine Learning for Each Other: To give a broader perspective of blockchain and ML technologies in communications and networking systems, we provide a brief overview of the benefits of blockchain for ML and the benefits of ML for blockchain, respectively.

On one hand, blockchain can benefit ML for data and model sharing, security and privacy, decentralized intelligence, and trusted decision-making. Specifically, blockchain-based system can store the large-scale data in a secure and tamper-resistant manner and ensure data confidentiality and auditability of the collaborative training process and the trained ML model via cryptographic techniques. Moreover, blockchain solutions enable a decentralized infrastructure to ensure secure access control without trusting external central entities. Leveraging blockchain, ML could learn, train, and derive decisions on local devices securely in decentralized and distributed networks. In particular, smart contracts and DAPPs may provide new opportunities to model the interactions among different entities in a decentralized ML application. At last, blockchain techniques enable the transparent immutable records of the training data and variables used by ML algorithms for their decision-making results to be reviewed and audited at any time by authorized nodes with access to the system. In this way, the processes of the ML algorithms are easily audited and the collaborative dynamics in decentralized ML applications are greatly optimized.

On the other hand, ML can benefit blockchain for energy and resource efficiency, scalability, security and privacy, and intelligent smart contracts. Specifically, leveraging the training data, ML-based mining algorithms may manage tasks in a more intelligent manner rather than adopting the brute force approach. Since ML algorithms can predict and speedily calculate data, it would also provide a feasible way for miners to select more important transactions to perform. In addition, the integration of blockchain and ML may revolutionize the traditional energy sector to make it become much more efficient and smarter. Leveraging ML techniques, blockchain applications can support predictive analytics to ensure the requirements for energy and resources to be accurately met and improve the efficiency of blockchain operation. To address the scalability issues, ML techniques can benefit the blockchain for optimizing data maintenance and storage by offering more efficient data sharding or pruning solutions. ML techniques

can also enable more efficient off-chain solutions or to adjust the block size dynamically [33] to make the blockchain-based system more scalable. Furthermore, the application of ML in blockchain-based communications and networking systems can detect malicious behavior on the blockchain by deploying the trained models and algorithms. In this way, ML could assist blockchain systems in identifying and preventing theft, fraud, and illicit transactions on the chain. Lastly, ML technique provides a feasible way to create and execute complicated smart contracts and make them more effective. Natural language processing (NLP) techniques are beneficial for smart contract negotiation and construction. With the aid of NLP, self-writing smart contracts can make exchanges of money, property, shares anything of value in more safe and cost-effective manner. It is conceivable that ML could be used on smart contracts for its creation and also for enhanced verification.

In this paper, we focus on the study of both blockchain and ML in communications and networking systems, including background, benefits of employing blockchain in ML algorithms and benefits of applying ML in blockchain systems. We also discuss some open issues, challenges, and broader perspectives that need to be addressed to consider blockchain and ML jointly for communications and networking systems.

III. Overview of Blockchain

In this section, we give a brief overview of the basic concepts, features, structure, and taxonomy of blockchain. Then the applications of blockchain in communications and networking systems are presented in detail.

A. What Is Blockchain?

Blockchain is essentially a kind of distributed ledger that involves a set of immutable transactions among the untrusted parties without any centralized controller or human intervention. These transactions on blockchain are grouped and stored in a persistent, immutable and tamper-proof ledger. Each of them is verified by using consensus mechanisms before attaching to the chain. Blockchain offers new opportunities for coordinating multiple untrustful parties and enabling decentralized governance in current networks. The main features of blockchain are summarized as follows, including decentralization, transparency, immutability, security, auditability, anonymity, and autonomy.[1]

- Decentralization: Thanks to P2P communications among the nodes, each transaction is generated, recorded, and validated by numerous nodes on blockchain without any central controller [64].[2] In this way, blockchain can significantly break through the bottlenecks of the operation of central servers.
- Transparency: A full copy of all transactions ever executed is stored and is transparent to all the nodes on public blockchain or for those permissioned on private/consortium blockchain [65]. The transparent transactional data makes blockchain more credible.
- Immutability: All the records on blockchain are irreversible and nonrepudiable thanks to the one-way cryptographic hash functions. All the transactions after the verification are recorded, and impossible to be changed unless someone controls the majority of the miners (voters) at the same time.[3]
- Security: Every transaction in blockchains will be broadcasted, checked, validated, and linked together by nodes in the whole distributed network. By using cryptography, any attempt to change the distributed ledger is nearly impossible and will be detected easily.
- Auditability: All the nodes on public blockchain or the permissioned nodes on private/consortium blockchain can audit, trace, and verify the current transactions iteratively via accessing to all timestamped transaction records [46].
- Autonomy: Based on the consensus mechanisms, each node on blockchain sends or receives transactions independently. By using public/private key pairs, the transactions are operated and managed without the need for human interaction or any trust third-party. It also prevents the blockchain from conflicting or double records.
- Pseudonymity: In blockchain systems, each node interacts with a generated pseudonymous address to avoid identity exposure. By exploiting pseudonyms, blockchain can provide pseudonymity and be suitable for some use cases that require highly privacy [64], [66].

These good features of blockchain may bring a lot of benefits to apply ML techniques in communications and networking systems. Due to its decentralized feature, blockchain can potentially neutralize the disadvantages of centralized ML, such as privacy-intrusive, single point of failure, and non-scalability. With the aid of blockchain, from training to optimization, every single step in the lifecycle of ML models can be improved with certain degrees of decentralization [67]. Specifically, blockchain could record processing and training data and models utilized by decentralized learning agents in a secure and private manner, providing a highly trusted data source for ML systems. Blockchain also facilitates decentralized ML to provide a reliable and permissioned medium for the exchange of the model parameters and maintain an association between the model parameters and the mini-batches used to train those parameters. Moreover, the transparent accountability and auditability features of blockchain could guarantee the trustworthiness of the data as well as of the ML models and provide a traceable ML decision-making processes. In blockchain-enabled ML systems, blockchain provides a simple way for auditing the decisions which are recorded on a datapoint-by-datapoint basis and cannot be tampered with [68]. Meanwhile, the

[1]Each type of blockchain places a different level of importance on decentralization, transparency, immutability, efficiency, and anonymity. For example, the public blockchain prioritizes decentralization, transparency, anonymity, and immutability over efficiency. Whereas private and consortium blockchains value immutability and efficiency over decentralization, transparency, and anonymity.

[2]In public blockchain, all the nodes are responsible for the maintenance of the blockchain. Meanwhile, in private or consortium blockchains, only some selected nodes could generate, record, and validate the transactions on the blockchain.

[3]Miners or voters are nodes on blockchain that are responsible for recording transactions, checking their accuracy, and adding them to blocks by solving a complex computational math problem or staking their assets.

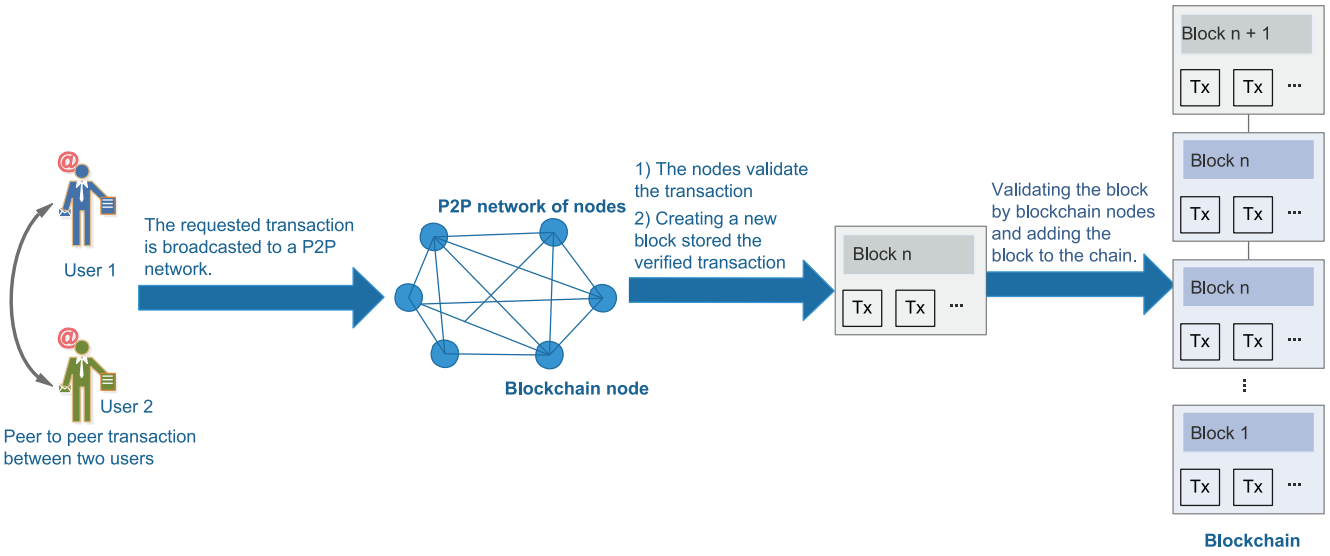


Fig. 2. A typical processing procedure of blockchain network. Tx stands for Transaction.

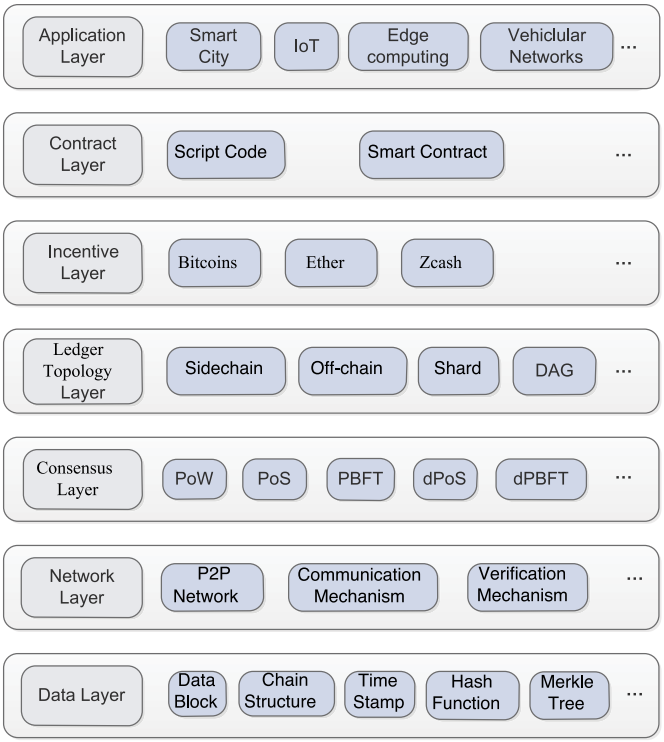


Fig. 3. A general architecture of blockchain.

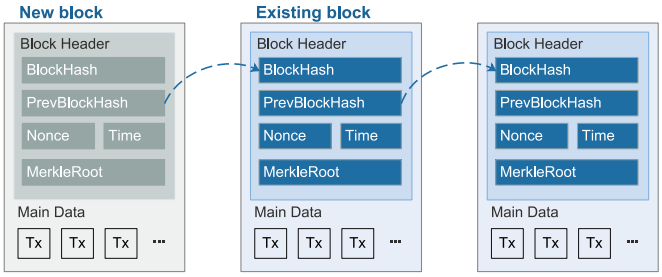


Fig. 4. Illustration of a chain of blocks.

immutability and distributed consensus models of blockchain technology intrinsically introduce a level of trust and enables collaborative dynamics in decentralized ML applications. Blockchain also enable decentralized ML applications to operate in autonomous and decentralized way through smart contracts without any central third-party.

To give a better understanding of blockchain techniques, an overview of the working procedure of blockchain technology is shown in Fig. 2. When a transaction is executed, it is hashed and broadcasted to each node on blockchain. Based on relevant consensus protocols, the transactions that are considered valid will be disseminated, stored, and grouped into a block by special nodes (i.e., miners or voters).

B. Architecture of Blockchain

To provide a clear presentation of blockchain techniques, based on the studies in [46], [69], a typical blockchain system is decoupled into six main layers, as shown in Fig. 3. The description and functions of these layers, including data layer, network layer, consensus layer, incentive layer, contract layer, and application layer are discussed in detail as follows.

The data layer mainly involves the transactions and blocks which store the transactional data generated from different applications. Each block includes a number of transactions and is linked to the previous block, forming an ordered list of blocks. Generally, as shown in Fig. 4, the block includes two parts, including the block header and main data. The block header specifies the metadata, including block version, the hash of previous and current blocks, timestamp, Merkle root [70], and other information. Explicitly, block version stores the relevant version of blockchain system and protocol. Based on the hash of the previous block, all blocks can be linked together to form a chain. Since the new block is constantly being superimposed on the old blocks, the latest block contains the hash pointer of its previous block. Utilizing the hash value of the Merkle tree root in blockchain, all the

TABLE II
TYPICAL CONSENSUS ALGORITHMS COMPARISONS

Property	PoW	PoS	DPoS	PBFT	dBFT
Node identity management	open	open	open	permissioned	permissioned
Energy saving	no	partial	partial	yes	yes
Tolerated power of adversary	$< 51\%$ computing power	$< 51\%$ stake	$< 51\%$ validators replicas	$< 1/3$faulty	$< 1/3$ byzantine voting power
Resource needed	Computation power	Wealth or stake	Wealth or stake	None	None
Reward miner	Yes	No	No	No	No
Scalability	Excellent	Excellent	Excellent	Poor	Good

recorded transactions on the current block could be checked easily and quickly. The block generation time is recorded in the timestamp. The main data of the block stores all transactions ever executed. The data type depends on what service on blockchain. Nevertheless, instead of blocks, the new transaction in the DAG reference previous transaction directly. There are no blocks of transactions in DAG networks.

The network layer provides the specific networking mechanism used in blockchain for broadcasting, forwarding, verifying, and auditing data generated from the data layer. The network is generally modeled as a P2P network, where nodes distribute the transactions and blocks in a decentralized way.

The consensus layer determines which the consensus algorithm is adopted to achieve consensus on some information among untrusted parties in decentralized systems [71]. Currently, there are a lot of consensus protocols being applied in blockchain systems, which could be roughly divided into consensus protocols with proof of concept (e.g., Proof of Work (PoW) [1], Proof of Stake (PoS) [72], Delegated Proof of Stake (DPoS) [73], Proof of Authority (PoA) [74]) and Byzantine fault-tolerant replication protocols (e.g., Practical Byzantine Fault Tolerance (PBFT) [75], Delegated Byzantine Fault Tolerance (dBFT) [76], Ripple [77]) [42]. According to the different types of blockchains, consensus protocols are selected differently. For example, due to the loose control and poor synchronization, the consensus protocol in permissionless blockchain (public blockchain) widely selects the incentive-based consensus schemes, such as PoW and PoS. On the contrary, the permissoned blockchains (private blockchain and consortium blockchain) usually adopts the Byzantine Faulty-Tolerant (BFT) consensus protocols such as PBFT and dBFT. Due to the limited space, we discuss some typical consensus mechanisms in detail. The comparison among these typical consensus mechanisms is listed in Table II.

PoW, which has been applied successfully in Bitcoin, is a probabilistic distributed consensus protocol. It requires a complicated computational process, where all the nodes (miners) that spend their computational resources are competing to solve a "puzzle". Specifically, all the nodes on blockchain use different nonces to calculate hash value continuously until the derived value is equal to or smaller than a given target. When one node gets the required value, all other nodes on blockchain would confirm the correctness of the value easily and quickly. Then the transactions in the current block are validated and recorded on the ledger. In this way, any change and modification to the block is nearly impossible once the block is added into blockchain system. It could also prevent Sybil attacks [78] by enforcing nodes to spend many resources.

Considering the huge energy consumption and scalability issue of PoW, PoS has been proposed to perform consensus process based on the users' assets ("stake"). Specifically, in PoS mechanism, only nodes that hold assets ("stake") could participate in the consensus process to generate the blocks. However, the security of the systems that adopt the PoS mechanisms may become another important issue since the mining cost is nearly zero.[4] To guarantee security while reducing the resource consumption, some blockchains jointly adopt PoW and PoS for performing the consensus process. For example, Casper is a consensus mechanism combining PoS algorithm and BFT consensus theory [79].

DPoS is also a consensus protocol based on the nodes' stakes. As compared to the direct democratic POS mechanisms, DPoS is representative democratic since the nodes elect their delegations to generate and validate the block. In this way, the transactions on the block can be verified and confirmed more quickly since the number of nodes that validates the block are much fewer. For instance, Bitshares [73] is implemented DPoS to perform the consensus process more effectively and offers a stack of financial services on a blockchain.

PBFT, which is successfully applied in Hyperledger Fabric [80], is a replication algorithm to tolerate Byzantine faults [81]. It can handle up to 1/3 malicious byzantine replicas. Different from PoW and PoS, every node in PBFT systems needs to know each other in the network. A new block is generated in a round without running the hashing process. Based on some rules, a primary is selected to be responsible for ordering the transactions in each round. Specifically, the consensus process includes three phases, namely, pre-prepared, prepared, and commit phases. If a node receives votes from over 2/3 of all the nodes, it will go to the next phase. However, PBFT consensus protocol is not scaleable and inefficient for large networks. This is because each node must exchange messages to other nodes to keep the network secure, which significantly increases the communication cost as a large number of nodes scale upwards. Unlike PBFT, in dBFT, combining the characteristics of dPoS, some professional nodes are elected to reach consensus and create a new block, while the other nodes

[4]Mining is the process of adding transactions to the distributed ledger (blockchain) of existing transactions.

TABLE III
COMPARISONS AMONG PUBLIC BLOCKCHAIN, PRIVATE BLOCKCHAIN, AND CONSORTIUM BLOCKCHAIN

Property	Public blockchain	Private blockchain	Consortium blockchain
Participants	Free Anonymous, could be malicious	Permissioned Identified and trusted	Permissioned Identified and trusted
Consensus determination	All miners	One organization	Selected set of nodes
Read permission	Public	Public or restricted	Public or restricted
Immutability	Yes	Partial	Partial
Efficiency	Low	High	High
Centralized	No	Yes	Partial
Consensus process	Permissonless	Permissioned	Permissioned

will act as ordinary nodes to receive and verify blocks. It is implemented in Antshares [73] and NEO [76].

The ledger topology layer determines the ledger topology for storing the data produced by the consensus layer. Mostly blockchain applications adopt the chain of blocks storing the ledger of the system. Considering the scalability issues, there are some new chain topologies, such as DAG, sidechain, and off-chain. DAG is a directed graph data structure in which individual transactions linked to multiple other transactions. Sidechains are blockchains that enable digital assets or data exchange from one blockchain to be used securely in a separate blockchain and subsequently returned to the original chain [82]. Off-chain refers to transactions that are not processed on the main chain.

The incentive layer introduces economic incentive to make the nodes to contribute their efforts to verify data in blockchain systems. It plays an important role in maintaining the decentralized blockchain system without any centralized authority.

The contract layer enables programmability in blockchain systems. Various script codes, smart contracts, and other programmable codes can be utilized to enable more complex programmable transactions. For instance, smart contract is written in a Turing-complete language as a digital contract that extends the semantics of transactions and implements complex business rules [83]. Smart contracts are essential scripts embedded on blockchain, and each of them has a unique address. Smart contracts enable the operators and the providers to precisely specify the conditions and define the business rules, possibly adding penalty mechanisms [84]. Currently, a lot of blockchain platforms support smart contracts. For example, Ethereum is a blockchain platform that supports smart contracts.

The application layer involves various applications, including IoT, smart cities, edge computing, security system, digital identity, etc. These applications may revolutionize these fields and provide efficient, secure, and decentralized management and optimization.

C. Taxonomy of Blockchain

Generally, there are three different types of blockchains, i.e., public blockchain, private blockchain, and consortium blockchain. The summarizes of three types of blockchains are listed in Table III.

1) Public Blockchain: In a public or permissonless blockchain network, anyone can join and participate in the network freely and all the transactions are completely open. Every node on the public blockchain can send, receive, check, and verify the transactions, participating in the process of consensus. In a public blockchain, publishing a new block usually needs either computationally expensive puzzle solving, or staking one's own cryptocurrency. This prevents the public blockchain from being hacked since it would be too costly to tamper its contents. However, although the transactions are anonymous, the openness of public blockchain may bring the concern about the privacy for transactions. Like Bitcoin [1], Ethereum [85], and Litecoin [86] are public blockchain.

2) Private Blockchain: In private or permissioned blockchain, only nodes that have the authority can participate in this blockchain. Private blockchain provides a feasible way for some groups and participants sharing and verifying transactions internally. It is generally applied in some applications with the scalability requirements and data privacy issues. However, since blocks are published by delegated nodes within the network, some rules or even data could be modified or tampered by the authority, and the organization may choose to roll back their blockchain to any point in the past [11], [64].

3) Consortium Blockchain: A consortium or federated blockchain usually operates by a group of pre-selected nodes. Different from public blockchain, not all the nodes have the authority to send, receive, check, and verify transactions in the consortium blockchain. The contractions may be public or restricted to read by some authorized nodes, which are selected in advance. To make blockchain much more scalable and provide more privacy, the consensus process may be operated by some preauthorized nodes instead of all the nodes. Similar to private blockchains, a consortium blockchain is usually not computationally expensive to publish new blocks. Considering the limited number of nodes in consortium blockchain, the data may be tampered and the blockchain would be rebuilt as long as the majority of the participants have reached a consensus [87], [88].

D. Blockchain Applied to Communications and Networking Systems

The exponential growth of smart devices and the fast expansion of applications put significant challenges for the current communications and networking systems, including management inefficiency and malicious hacking attack issues [89], [90]. First, along with the increase in devices, managing these devices in the existing centralized model will

bring great issues to the network maintenance and management. In addition, communications and networking systems involve massive amounts of data traffic. These data are often routed through critical centralized servers which may suffer malicious parties attacks. Last, the users may worry about the privacy of their data and are reluctant to provide valuable data for processing and analysis.

To support enormous smart devices/applications and enable decentralized, self-managed, and secure networks, the application of blockchain in communications and networking systems brings significant benefits. First, blockchain-based systems provide direct communication among users in a P2P way without relying on any intermediary agents, which reduces the communication cost and improve security and privacy. In addition, more useful data restored on blockchain could be provided over the communication networks for optimizing the network operations since the blockchain is trustworthy, secure, and cannot be tampered with. Last, the decentralization of blockchain also enables flexible management and achieves the desired level of security and trustworthiness in the future heterogeneous network connected together by 5G, satellite network, ocean-based networks (supporting networking and communications in ocean), and so on. A lot of works about blockchain have been done in a myriad of scenarios. In the following, we present blockchain applications in some typical network scenarios, including IoT, vehicular networks, cloud/edge computing systems, and wireless networks.

1) Blockchain Applied in IoT: IoT, which involves the ubiquitous interconnection of various devices with networking and computing abilities, provides a promising way for data collection, analysis, and sharing. It becomes one of the most fundamental architectures for various applications such as smart healthcare, intelligent transportation, smart home, and industrial IoT. The traditional IoT systems rely on centralized servers may suffer security attacks and high operational cost. The application of blockchain in IoT can facilitate the sharing of resources and information and build IoT services marketplace where various devices share their data and resource in a secure and verifiable manner [91]. However, the authors in [92] state that PoW is not practical for IoT networks since it requires high computational, bandwidth, and energy. Considering the huge resource requirements of blockchain and the inadequate performance of IoT devices, the authors in [93] propose *LightChain* for IoT which implements a synergistic multiple proof consensus mechanism to set multiple standards and dynamic difficulty for each node to save the computing power consumption of mining. In addition, an unrelated block offloading filter and a lightweight data structure are proposed to reduce the occupied storage resources. Moreover, smart contracts can automate the procedures and interactions among IoT network operators and subscribers to further reduce operational cost [94]. The work of [95] provides a blockchain-based system to improve the operational and service capabilities of IoT nodes. In the energy sector, the authors in [96] propose a secure energy trading system in IoT for building P2P energy trading without any trusted intermediary. They also propose a credit-based payment scheme in the blockchain-based energy system to reduce the transaction confirmation delay and enable fast and frequent energy trading.

2) Blockchain Applied in Vehicular Networks: Due to the high mobility and variability of vehicular networks, the security, privacy, and trust management problems are still open issues for facilitating safe, efficient, and intelligent transportation. Since blockchains guarantee the data consistency and tamper-resistance, the authors in [97] propose a secure consortium blockchain-based data sharing and storage system with the digital signature in vehicular ad-hoc network (VANET). The roadside unit (RSU) deploys smart contracts for setting the condition of data sharing and storing the replica sensor data in a distributed way. To enable trustless, traceable, and privacy-aware vehicular networks, the authors in [98] propose blockchain-based vehicular public key infrastructure (VPKI) to support membership establishment and privacy. They use off-chain solution and propose a fragmented ledger that stores the key data relevant to vehicles for performing post-accident analysis to minimize storage and processing overhead. The authors in [15] propose a blockchain-based anonymous reputation system, which considers the trustworthiness of messages based on the reputation score and utilizes a lexicographic Merkle tree to improve the efficiency of the blockchain-based authentication. To guarantee the information security and privacy in VANETs, the authors in [99] propose a blockchain-based traffic event validation framework which uses proof-of-event consensus mechanism to ensure the reliability of confirming the event occurrences. In terms of security and traceability of data sharing, the authors in [100] propose an enhanced DPoS consensus scheme with two-stage soft security solution, i.e., miner selection and block verification, for secure vehicle data sharing in blockchain enabled Internet of Vehicles (IoV). To provide efficient data storage and sharing, the authors in [101] propose a secure reputation-based data sharing system with consortium blockchain and smart contracts in vehicular computing and networks. The proposed reputation-based scheme mainly takes three-weight subjective logic model into account, including interaction frequency, event timeliness, and trajectory similarity, to improve the credibility and ensure high-quality data sharing.

3) Blockchain Applied in Mobile Edge Computing: Mobile edge computing (MEC), which extends data processing to the edge of the network, is initiated by European Telecommunications Standards Institute (ETSI) as a promising edge-cloud computing paradigm in mobile networks. Since September 2016, ETSI renamed MEC as Multi-access edge computing to extend its applicability into heterogeneous networks including WiFi and some fixed access technologies [102]. The decentralized blockchain and the distributed MEC network may work together and bring considerable benefits in wireless networks. The authors in [47] present a survey on the integration of blockchain and edge computing for secure access and control the network, storage, and computing resources at the network edge. The authors in [103] propose economic edge computing resource management based

on Stackelberg game model for mobile blockchain.[5] In edge computing enabled mobile blockchain networks, edge servers could provide computing resources and services for enabling the blockchain applications of mobile users [104]. To address the mining issues in mobile blockchain networks, the authors in [105] propose an auction-based edge computing resource allocation scheme, where mobile users (miners) could offload the mining tasks to an edge server. To maximize the social welfare and guarantee the truthfulness, the edge resource allocation for mining offloading is modeled as a combinatorial auction problem considering allocative externalities, i.e., the competition among the miners and the network effects in the blockchain network. To address the massive demands for the computational capability in the mining process, the authors in [106] propose an MEC-based computational tasks offloading scheme with blockchain which includes two offloading modes, i.e., offloading to the nearby access point and offloading to a group of nearby users. To handle the frequent handover and significant overhead issues in fog networks, the authors in [107] propose a dynamic mobility management mechanism based on blockchain, which could resolve hierarchical security issues without affecting the network layout and also satisfy fully distributed security requirements. The authors in [20] propose a blockchain-based scheme with encrypted cloud storage to deal with the privacy issues in content-centric mobile networks. In the proposed scheme, the users are able to select miners to maintain the public ledger expediently and perform access control for managing the transactional data.

4) Blockchain Applied in Other Wireless Network Scenarios: Aside from the discussions of the blockchain applied in specific network scenarios, the authors in [10] explore the blockchain applied in wireless networks, such as computing, fog radio access network (RAN), grant-free uplink access, IoT. Specifically, the consensus mechanism of blockchain may be applied to manage tight synchronization among networking resources in different locations. Blockchain could also be applied to avoid the potential collision in RAN. To address the frequency authentication requirements in ultra-dense networks (UDN), the authors in [108] propose a blockchain-based security authentication scheme, which adopts PBFT based consensus mechanism to reduce the authentication frequency when the user moves among the access points and improve the operation efficiency. In cognitive radio networks, the authors in [109] propose a blockchain-based decentralized database to provide secure spectrum sharing and introduce a virtual currency for performing the spectrum sharing and access. The authors in [110] propose a blockchain RAN (B-RAN) architecture to enable network authentication and access control in a decentralized, secure, and efficient manner among a number of trustless network entities.

Although blockchain technique provides a feasible way to guarantee data immutability and protect user/data privacy for securing communications and networking systems. Employing blockchain mechanisms in communication networks has some technical challenges and limitations. For example, how to efficiently and reliably achieve interoperability between different blockchains, and how to reduce the resource consumption and improve the management efficiency and when using blockchains still need to be studied and addressed. On top of these challenges, the intelligence of blockchain in communication and networking systems is still in its infancy and requires designing a concrete consensus for the same. To address these issues, ML can benefit blockchain in communications and networking systems. Specifically, leveraging the training data, ML-based algorithms may provide a feasible way to operate different blockchains and optimize the data storage and maintenance solution for blockchain system. Considering the resource cost issues, leveraging ML techniques, blockchain applications can support predictive analytics to ensure the requirements for resources to be accurately met and to improve the efficiency of blockchain operation. ML-based algorithms may manage mining tasks (e.g., PoW-based blockchain) in a more intelligent manner rather than adopting the brute force approach. Furthermore, the application of ML in blockchain-based communications and networking systems can detect malicious behavior on blockchain by deploying the trained models and algorithms. In this way, ML could assist blockchain systems in identifying and preventing theft, fraud, and illicit transactions on the chain. Lastly, ML technique provides a feasible way to make the blockchain more effective and intelligent. In the next section, we will provide an overview of ML to show the advantages of ML for blockchain-based applications. Also, we will discuss the applications and requirements of ML-based for communications and networking systems, which may be well addressed by blockchain techniques.

IV. Overview of Machine Learning

In this section, we give an overview of the basic concepts, history, and taxonomies of ML. Then we present a wide variety of ML algorithms applied in communication and networking systems.

A. What Is Machine Learning?

ML, which was first coined by Samuel, in 1959, is defined as "the field of study that gives computers the ability to learn without being explicitly programmed" [111]. Subsequently, E. Tom Mitchell gave a better ML definition as "a computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E". Particularly, ML is a compelling tool for providing solutions to the problems and improving the performance of the developed system based on data sets.

As shown in Fig. 5, a typical workflow of ML framework commonly contains the training process and test process.[6] Initially, in the training phase, the raw data are pre-processed to provide usable data for the next step. Then the features and

[5]In Stackelberg game model, one player (leader) has announces its strategy first, and then other followers make them choice accordingly.

[6]The presented workflow is a typical procedure of ML framework, however, there are still some other ML algorithms have different steps for practical constraints or given tasks.

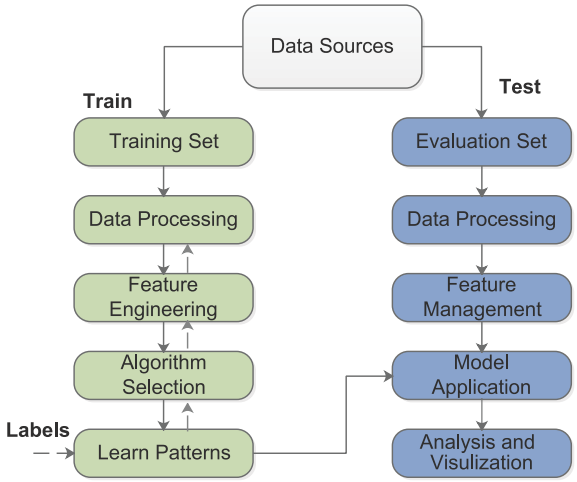


Fig. 5. A workflow of machine learning model construction and test.

patterns of these data can be extracted and processed for the given tasks. Support Vector Machines (SVM) [112], Hidden Markov Model (HMM) [113], Principal Component Analysis (PCA) [114], etc., are widely used for pattern recognition and feature extraction. The key steps of algorithm selection and learning patterns are generally performed in tandem. As shown in Fig. 5, in learn patterns part, different learning algorithms have different patterns. For example, supervised learning uses labeled training datasets to construct the model representing the learned relationship while unsupervised learning algorithm uses training datasets without labels. According to the diverse requirements of different tasks, a variety of feasible results are derived. In the test phase, using the test data, the outcomes of the ML models are mapped to representative knowledge (e.g., a particular pattern), and insights of them are delivered either to a dashboard or other related software components.

B. A Brief History of Machine Learning

ML is a subfield of AI where computer machines are used to autonomously learn from data and information. In this subsection, we present a brief history of ML about several important discoveries and inventions.

In 1950, Alan Turing created the world-famous Turing Test to estimate whether a computer has real intelligence. In this test, if a computer can convince a human that it is a human and not a computer, the computer passes the test. In 1952, the first computer learning program - the game of checkers, was written by Arthur Samuel. In 1957, the first neural network for computers, called Perceptron, was designed by Frank Rosenblatt, simulating the thought process of the human brain. In 1967, The "nearest neighbor: algorithm" was written, allowing computers to use simple pattern recognition. In 1981, Gerald Dejong introduced the concept of Explanation-Based Learning (EBL), in which a computer analyzed training data and created a general rule it can follow by discarding unimportant data. In 1997, IBM's Deep Blue beat the world champion Garry Kasparov at a game of Chess. From then on, many projects and models in the ML field were widely developed. At the start of the 21st century, much more ML projects were developed in various applications and fields. For example, *GoogleBrain*, which was a deep neural network created by Google in 2012, focused on pattern detection in images and videos. *DeepFace*, which also was a deep neural network created by Facebook in 2014, could recognize people with the same precision as a human can. *Toolkit*, which was developed by Microsoft in 2015, enabled the distribution of ML problems across multiple computers. Moreover, *AlphaGo*, which was developed by Google DeepMind, beat a professional player at the Chinese board game. In conclusion, ML technologies have developed and undergone various transformations over a long time. Currently, ML technologies are considered as a very promising solution in some areas to build intelligent machines, although they are computationally very expensive.

C. Taxonomy of ML Technology

The ML techniques can generally be classified into four different areas, i.e., supervised learning, unsupervised learning, semi-supervised learning, and RL. As depicted in Fig. 6, we present a brief discussion of the four types of ML techniques.

1) Supervised Learning Algorithm: Supervised learning algorithm is a type of labeling learning technology which uses labeled training dataset to construct the model representing the learned relationship between the input, and then output for predicting output values. After sufficient training, the supervised learning algorithm can provide targets for any new input, compare its output with the correct, intended output and find errors to modify the model sequentially. In the context of communications and networking systems, supervised learning can be applied in a large variety of fields, such as mobility prediction, resource allocation, and load balancing. For different applications and objectives, there are several supervised learning algorithms, including k-nearest neighbor, decision tree, neural network, SVM, Bayesian' theory, and HMM.

2) Unsupervised Learning Algorithm: Different from supervised learning, the unsupervised learning algorithm is a type of ML algorithm that uses training dataset without labels. The unsupervised learning algorithm generally aims to classify the sample set into different sets by exploring the similarity among them. The unsupervised learning techniques are widely used in clustering, anomaly detection, data aggregation. In the context of communications and networking systems, the popular unsupervised learning algorithms, including k-means, self-organizing maps, and anomaly detectors, have been applied in several fields, such as handover management, fault detection, network operational configuration, and energy efficiency management.

3) Semi-Supervised Learning Algorithm: Semi-supervised learning is a type of learning technology in which most of the training samples are unlabeled and a few of them are labeled. By utilizing the unlabeled data with a small number of labeled data, semi-supervised learning can efficiently improve the learning accuracy over unsupervised learning, while do not need much time and costs required by supervised learning. To

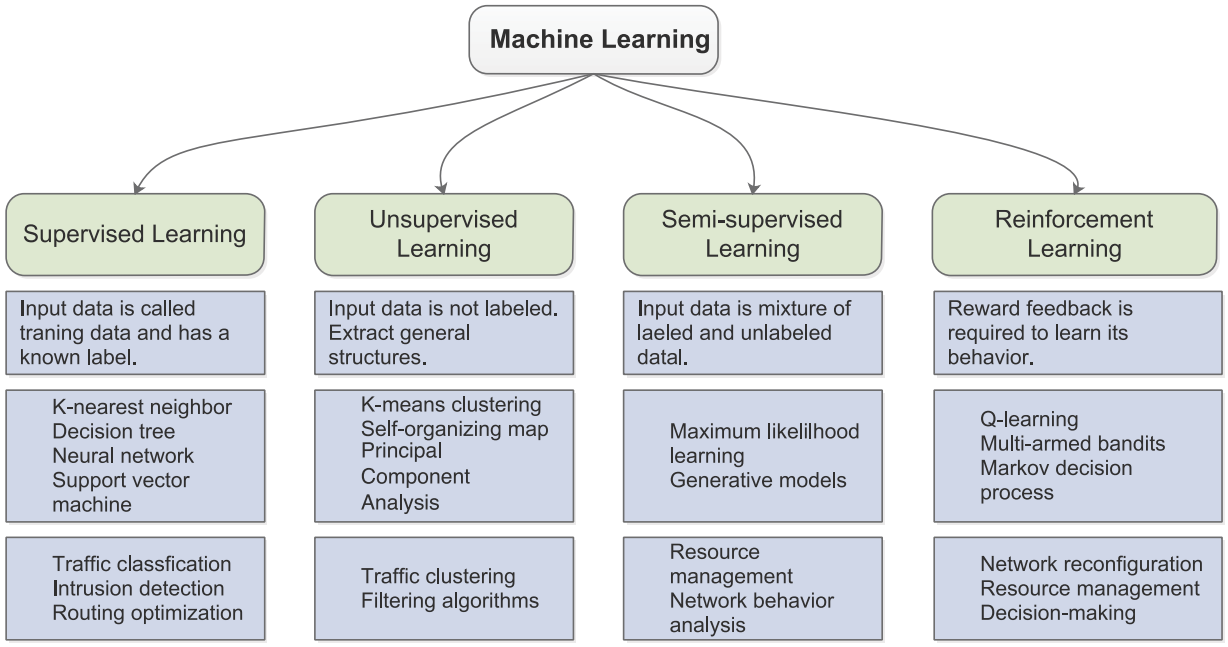


Fig. 6. Common machine learning algorithms applied to communications and networking systems.

effectively utilize unlabeled data, some the underlying distribution of data structure has to be assumed, such as cluster assumption and manifold assumption. By utilizing different structure assumptions, the commonly semi-supervised learning methods such as Expectation Maximization (EM), co-training, transductive SVM, and graph-based methods, are widely used in a lot of domains, such as speech analysis, Web content classification, and caching management.

4) Reinforcement Learning Algorithm: RL is a type of learning technology that trains algorithms using a system of reward and punishment mechanisms. It enables agents to learn by exploring the available actions and refining their behavior within a specific environment, to maximize its long-term reward. In other words, agents not only consider the immediate reward but also evaluate the consequences of their actions on the future. RL is usually modeled as a Markov decision process (MDP), which involves a set of environment and agent states S, a set of actions A of the agent, the probability of state transition, the immediate reward, and rules. The agent interacts with the environment by exploring actions to derive the sequence of state-action pairs that maximizes the expected discounted reward, i.e., the optimal policy. The Q-learning is the most popular RL algorithms in which Q-function is used to learn a table storing all state-action pairs and their long-term rewards. In communications and networking systems, RL algorithms are quite popular and have been applied in a variety of applications, such as coverage and capacity optimization, handover parameters optimization, load balancing, and resource optimization.

D. Machine Learning Applied in Communications and Networking Systems

ML technologies bring significant benefits across a range of fields and applications, including image/video recognition, natural language and text analysis, robotics and autonomous vehicle. The same potential can be expressed to enable the management and operations of complex communication and networking systems that provides various services and applications for a great number of users with the distinct QoS requirements. On one hand, with ML algorithms, communications and networking systems can become much more intelligent, autonomous, and efficient in improving network resource utilization and optimization. On the other hand, the users can enjoy much better services, as enabled by the learning of the operating network environment and the continuous adaptation of communication and network parameters as the observed conditions evolve. In the following, we present some typical ML applications in these fields, including network traffic classification, routing optimization, resource management, and network security.

1) Network Traffic Classification: Network traffic classification is crucial for operators to perform network operation and management. Traditionally, network traffic classification has mainly adopted port-based, payload-based, and host-based approaches. However, there are several severe limitations of them for current complicated networks. For example, Port-based approaches are usually unreliable and antiquated, payload-based approaches are computationally intensive and complicated, and host-based approaches are highly susceptible to routing asymmetries. To address these issues, ML-based traffic classification has been proposed as a promising solution for the network traffic classification, prediction and knowledge discovery [56]. The authors in [55] present a comprehensive survey on the ML applications in network traffic classification. They investigate several problems of the IP traffic classification using ML algorithms, ranging from data collection, feature engineering, algorithm selection, to model deployment. They also identify some

challenges at each stage, in order to optimize the performance of ML approaches for traffic classification.

Supervised learning, such as kernel estimation [115], [116], neural network [117] and SVM-based ML techniques [118], are widely used for network traffic classification and could achieve high accuracy. Considering the practical deployments, neural network-based traffic classification has been proposed and has also achieved higher accuracy [119], [120]. In addition, the authors in [121], [122] propose multi-class SVM-based traffic classification mechanisms with joint consideration on the accuracy of traffic classification and the applicability to large datasets. On the other hand, unsupervised learning, such as hard and soft clustering, has also been deeply explored for robust network traffic classification using flow features [123]–[125]. While the hard clustering may not be able to handle the flow features that exhibit high similarity, the soft clustering has a lower training time and achieves the required granularity with density-based clustering technique.

Considering the high accuracy of supervised ML approaches and the robustness of unsupervised ML approaches, the combination of supervised and unsupervised ML for network traffic classification has drawn a lot of attention from academy [126]–[128]. Moreover, to address the data volumes and algorithm-driven applications, DL algorithms are able to provide powerful tools that learn highly complicated patterns and gain a higher accuracy in network traffic classification since they can automatically extract and select features through training. Considering the privacy issues, the authors in [129], [130] investigate DL-based methods for encrypted classification of mobile services, such as classification task definition, data preparation, pre-processing, model input design, pre-training design, and model architecture.

2) Routing Optimization: Routing optimization significantly affects the performance of communications and networking systems. Adaptive routing and shortest path routing algorithm are two types of most popular routing optimization approaches. However, the high computational complexity of the adaptive routing algorithm limits its performance in current networks. Meanwhile, shortest path routing is a best-effort routing protocol and may result in low resource utilization. To address these limitations, ML algorithms have been widely proposed for handling different routing issues, which could derive the near-optimal routing solutions quickly.

Supervised learning algorithms are widely investigated to train the model to derive the optimal heuristic-like routing solutions. Specifically, the network and traffic states are often modeled as the input and the routing solutions of heuristic algorithms are often derived as output. The authors in [131] propose a supervised ML-based framework to deal with the dynamic routing issue in real time. In [132], the authors propose a supervised ML-based approach for routing optimization in named data networking. The proposed approach uses the compressed forwarding information base (FIB) data structure and trains artificial neural networks from the routing information base in an offline manner. The authors in [133] propose a dynamic routing framework called *NeuRoute*, which uses the network state and the estimated network traffic as the input and the routing solution derived by heuristic algorithms as output to trains the neural network model and achieve the optimal routing solution.

RL algorithms are also investigated to optimize routing solutions. In RL-based routing mechanisms, the network control entity acts as an agent, the network is the environment, the state space contains the network and traffic states, the action is the routing solution, and the reward is defined based on optimization metrics [134]. The authors in [135] propose an RL-based hierarchical protocol to build a Q-value table based on the traffic flow in neighbor grids for the grid selection. The authors in [136] propose a distributed RL-based routing protocol in SDN, which determines the optimal data transmission paths based on the network status and users' requirements. The authors in [137] propose a deep RL (DRL)-based routing optimization mechanism, which uses neural networks to optimize the customizable routing by intelligently adjusting the reward function in SDN.

3) Resource Management: The application of ML technique into communication and networking systems would revolutionize the resource sharing and scheduling schemes and satisfy various application requirements by learning from the network environment. The authors in [138] propose an improved genetic algorithm for improving energy efficiency and users' QoS in fog computing-based IoT networks. For the complex network environments, much more requirements such as bigger datasets and faster learning algorithms, are carried out in ML systems.

To address these issues, RL-based resource managements in communications and networking systems have drawn a lot of attention and been applied in various networks. To maximize energy efficiency, the authors in [139] propose an enhanced actor-critic RL scheme to investigate the resource allocation problem for live streaming in edge-cloud networks. In [140], the authors investigate the joint caching, computing, and radio resources problems and propose a natural actor-critic DRL model to solve them in a fog-enabled IoT networks. In [141], the authors propose an intelligent resource allocation framework (iRAF), which learns the network environment and derives resource allocation solution automatically in a collaborative MEC network. They also propose a multi-task DRL algorithm to solve the complex resource allocation problem to reduce latency and power consumption.

Moreover, network virtualization and network slicing are two emerging concepts for optimizing resource utilization and providing cost-effective services in the future wireless networks. The authors in [142] propose a DRL-based scheme for efficiently handling complex and dynamic service function chain (SFC) in IoT. In the proposed scheme, SFC is an ordered combination of virtual network functions (VNFs) that are related to each other based on the logic of IoT applications. On the other hand, network slicing is considered as a promising solution to support network operators to dynamically allocate the customized slices to different tenants. The authors in [143] investigate the application of DRL for addressing several typical resource managements with network slicing, such as radio resource slicing and priority-based core network slicing. In

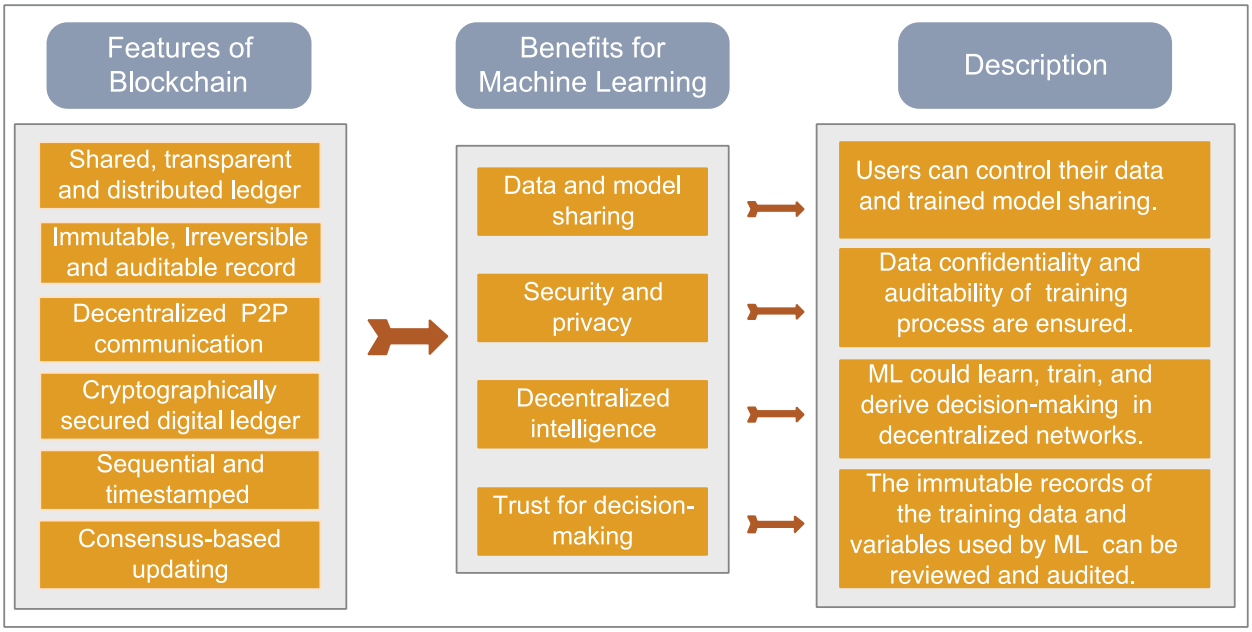


Fig. 7. Key features of blockchain that can benefit machine learning.

addition, they present an in-depth discussion about the advantages and disadvantages of DRL applied to network slicing scenarios.

4) Network Security: Network security is always a crucial aspect of network operation and management. Currently, ML algorithms have been widely used in IDS for preventing intrusions. The authors in [59] present a comprehensive survey on the application of ML for three types of IDSs, such as misuse-based, anomaly-based, and hybrid detection mechanism. Moreover, IDS is also often embedded with some data analysis methods such as DM and classification to improve their effectiveness. The authors in [127] propose an incremental network traffic classification algorithm based on semi-supervised particle swarm optimization (PSO), to enhance the performance of IDS. By applying ANN classifier, the authors in [144], [145] propose hybrid magnetic optimization algorithms based on PSO to improve the accuracy of the network attack and intrusion detections. For the distributed architectures, a distributed intrusion detection framework based on PSO and SVMs is proposed in [146] to improve the accuracy of the detection. A DL approach is proposed for developing an efficient and flexible architecture for network IDS using self-taught learning (STL) [147]. To efficiently detect attacks, the authors in [148] propose a reservoir computing based attack detection strategy for robustly detecting attacks in smart grids.

However, contemporary ML systems are generally built based on a centralized paradigm, where a single entity needs to have a global vision of the problems, as well as all the necessary knowledge and resources for collecting large datasets, creating highly sophisticated models and solving various types of tasks. However, the centralized ML systems currently may no longer be adequate for evolving large-scale complicated parameters and may pose severe limitations in communications and networking systems. The centralized approaches are often non-scalable, operationally ineffective, and easy to suffer failure problems. In addition, the lack of trust would severely limit the performance of ML solutions. How to design decentralized training models with proper parameters and structures to address security, privacy, and trust issues is still challenging. To address these issues, blockchain can benefit ML in communications and networking systems. Specifically, the decentralized feature of blockchain enables a decentralized ML solution, which could learn, train, and derive decisions on local devices in decentralized and distributed networks without any central server or intermediaries. Moreover, these collaborative training and decision-making processes can be recorded on blockchain in a secure and tamper-resistant manner, which ensures data confidentiality, privacy, and auditability via cryptographic techniques. By leveraging blockchain technology, those transparent immutable records of all the data, variables, and processes used by ML applications for their decision-making processes can be reviewed and audited at any time by authorized nodes with access to the system. In this way, it provides a feasible solution to audit the processes of the ML applications and significantly improve the trustworthiness of the data and ML models. In the next section, we will explore the benefits of blockchain for ML in communications and networking systems.

V. Blockchain can Benefit Machine Learning in Communication and Networking Systems

Thanks to the main features of blockchain, including decentralization, immutability, and transparency, blockchain provides a new opportunity for ML algorithms applied in communications and networking systems. In this section, as shown in Fig. 7, we show that blockchain can benefit ML, including data and model sharing, security and privacy, decentralized intelligence, and trustful decision-making. The summaries are listed in Table V.

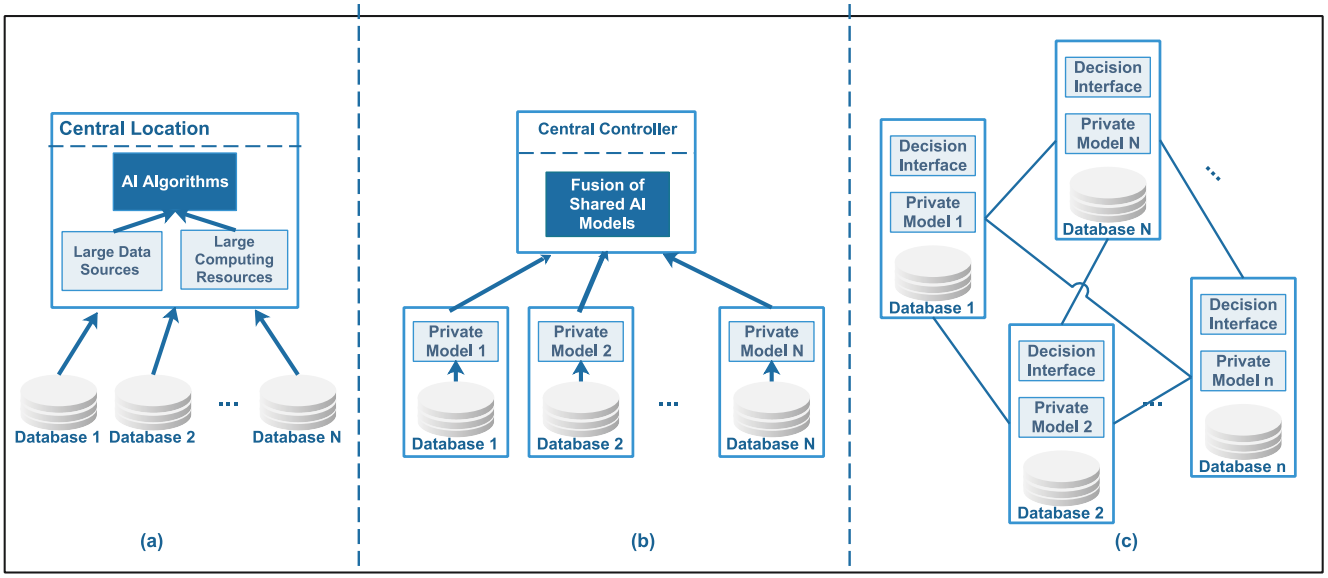


Fig. 8. The illustration of different machine learning architectures: (a) Centralized ML architecture where a central server processes the collected data using large storage and computing resources; (b) Distributed ML architecture where a central server controls the process of training model and the distributed users train the models with their datasets independently; (c) Decentralized blockchain-based ML architecture without any central controller for data and model sharing.

TABLE IV
KEY ROLES IN OCEAN ECOSYSTEMS

Role	belong to which layer	Responsibility	Rewards
Data/service provider, data owner		Providing data, algorithms, compute, storage, etc	Ocean Tokens for providing data/service
Data/service consumer		Accessing to publisher data/services Giving signals to curators	
Marketplaces, data commons		Running metadata store, secret store	Transaction fees
Data/service curators	Curation layer	Signaling the relative value of data/service	Ocean Tokens for curating
Data/service verifier	Verification layer	Data/service verification	Ocean Tokens for verification
Keepers	Keeper layer	Correctly running nodes in network	Ocean Tokens for chainkeeping

A. Blockchains can Benefit Machine Learning for Data and Model Sharing

Communication and networking systems applying ML solutions generally involve a large volume of data, such as training on large datasets or high-throughput stream processing. With more data to analyze, the prediction and decision-making of machines are considered more correct, and the trained models and algorithms are more reliable. As shown in Fig. 8(a), a central server collects and processes data from various databases to train the model effectively. However, accessing to a large amount of data for training ML models may be not available in current communication systems. In addition, aggregating data in heterogeneous networks for ML training is also a challenge. Specifically, in the traditional centralized ML solutions, the bandwidth constraints and the communication cost significantly limit the aggregation of data in heterogeneous networks. The severe overhead incurred in getting this data lies in collecting, organizing and auditing the data for accuracy. A distributed ML architecture where a central server controls the process of training model and the distributed users train the models with their datasets independently is shown in Fig. 8(b). Nevertheless, the inefficient management of data sharing among distinct multi-parties is a crucial bottleneck of the development of ML technique. Heterogeneous devices have different levels of willingness to participate in and share their data considering the privacy issue. While, some heterogeneous devices may be malicious and able to infer sensitive information from shared parameters, which potentially negates privacy preservation. These malicious participants may also attack the ML-based systems, which brings security issues in communications and networking systems. To deal with these issues, as illustrated in Fig. 8(c), blockchain technology encourages individuals to share data to a decentralized and distributed ledger, and then it provides more datasets and more available data for training ML models.

Blockchains enable data sharing and storage without entrusting any third-party by exploring some existing technologies, including time-stamping of transactions, cryptography, P2P networks, etc. Specifically, data and information could be timestamped in a decentralized and tamper-proof manner. Cryptography offers secure data-transmission and enables records immutability in decentralized P2P networks. For instance, data are hashed and the hash can be incorporated into a transaction stored in the blockchain, which provides a secure proof of the exact time at which that data existed. Moreover, a blockchain is an immutable read-only data structure, where

TABLE V
A SUMMARY OF BENEFITS FOR ML APPLICATIONS PROVIDED BY BLOCKCHAIN

Benefits	Ref.	Use case	Year of publication	Purposes	Contributions
Data and model sharing	[149]	AI systems	2019	Sharing AI data and model in a safe, secure and transparent manner	Developing a system for privacy-preserving data sharing with blockchain
	[150]	MLsystems	2019	Offering quality data for ML professionals to buy or rent	Dividing the network into multiple Interest Group (IG) as sharding solutions
	[151]	Healthcare systems	2018	Building a decentralized personal health data ecosystem for drug discovery, biomarker development, and preventative healthcare	Enabling patients to control and profit from their personal data as well with the incentives to undergo constant health monitoring
	[152]	Healthcare systems	2018	Utilizing transaction data for disseminating model and other meta information	Preserving the privacy of data and improving the healthcare prediction model
	[154]	Robotic systems	2018	Building a secure, decentralized and computationally efficient data and model sharing among multiple robot units	Leveraging blockchain and federated learning for performing ML operations in a secure, decentralized way
	[155]	IoT	2018	Facilitating processing of personal data in a transparent, unambiguous, and auditable manner	Utilizing blockchain to support the integrity, the non-repudiation and the versioning of consents in a public verifiable way
	[156]	ML system	2018	Enabling the trustworthiness of data for ML applications	Leveraging blockchain to store the results of an ML application
Security and privacy	[166]	ML system	2019	Building a secure and reliable data sharing platform among multiple data providers	Designing secure building blocks and constructing a secure SVM training algorithm to ensure the confidentiality of the sensitive data
	[167]	DL system	2019	Supporting a secure and privacy-preserving DL framework among multiple workers	Designing a blockchain-powered privacy-preserving DL framework
	[168]	ML system	2018	Providing data confidentiality, computation auditability, and incentives for collaborative training	Using cryptographic techniques and the tool of non-interactive zero-knowledge proof to provide auditability of the collaborative training process
	[169]	ML system	2018	Protecting data holder's privacy and guaranteeing the resilience of the system	Designing a decentralized SGD algorithm to learn a general predictive model over the blockchain
	[170]	IoT	2018	Running an ML model in a decentralized way for personalization of IoT devices in a smart home	Performing distributed association rule mining to generate the rules of user activity from the device logs and stored in the blockchain s
	[171]	IoT	2017	Providing a dynamic, optimized and self-adjusted security services	Proposing a dynamic and fully distributed security scheme for access control
	[172]	D2D networks	2018	Performing the access control	Proposing a blockchain-based scheme to verify the authenticity of channel state information (CSI)
Decentralized intelligence	[183]	SingularityNET platform	2017	Decentralized ML for the improvement of ML and AGI	Discussing decentralized ML and present a case study
	[173]	Decentralized ML	2017	Providing a solution for implementing decentralized ML algorithms	Proposing DML protocol and creating a distributed ML system
	[185]	ML system	2019	Decentralizing ML by leveraging blockchain	Utilizing the large scale data generated by users to train ML models
	[186]	Blockchain	2018	Providing decentralized ML applications	Building a scalable network of decentralized applications
	[187]	Database services	2018	Accessing to high-quality models on a public blockchain	Proposing blockchain-based protocol for evaluating and purchasing ML
	[67]	ML services	2018	Enabling decentralized training mechanisms in ML models via blockchain	Proposing blockchain-based framework for hosting and training ML models on Ethereum blockchain
	[188]	ML system	2018	Enabling the decentralized ML without any trusted third-party	Proposing a crowdsourcing blockchain-based scheme for decentralized ML
	[189]	DL system	2018	Enabling a decentralized DL applications	Proposing a cooperative decentralized DL architecture
Trustful decision-making	[197]	ML system	2018	Enabling trustable ML applications	Proposing a trustable ML framework by using blockchain and smart contracts
	[199]	healthcare systems	2018	Establishing a decentralized trust between parties involved in a disease modeling process	Allowing users to track, validate and verify ML task and results
	[200]	Wireless sensor network	2019	Improving the routing security and efficiency for wireless sensor network	Proposing a trusted routing scheme using blockchain and RL
	[201]	API marketplace	2018	Ensuring fair and trusted usage of hosted APIs and averts malicious behavior	Designing a blockchain based decentralized trustless API marketplace
	[202]	Mobile networks	2019	Defending against unreliable model updates in mobile networks	Introducing reputation as the metric to measure the reliability and trustworthiness of the mobile devices

new blocks get appended onto the end of the ledger by linkage with the previous block's hash identifier. In this way, blockchain supports secure data sharing and storage in trustless networks.

The Ocean Protocol [149] is a decentralized platform which utilizes blockchain technology for sharing AI data and model in a secure and transparent manner. The Ocean decentralized network has three significant layers: 1) the Keeper

layer that manages service agreements, low-level access control, accounts, balances, and the incentive schema (or block reward); 2) the Verification layer that introduces cryptographic challenges to improve the integrity and security of the services; 3) the Curation layer that serves as a discovery mechanism as well as signalling and governance aspects. Table IV outlines the roles participating in the Ocean network. In the platform, data providers can upload their data and earn revenue by publishing data while the data consumers can discover and buy data for creating innovative applications. The keepers implement service agreements (SAs) as smart contracts that dis-intermediate between data providers and consumers. An incentive mechanism is proposed to encourage participants/users to run keepers. In particular, the data consumer needs a provably correct model execution on the purchased data. To address this issue, verifiers in the verification layer are responsible for verifying and enforcing the services and conditions in the SAs via posting challenges for provers and verifying the returned proof. The type of challenge depends on the service offering (i.e., data integrity and computation integrity.)[7] For data integrity, they use a data integrity hash in an on-chain smart contract and use the checksum information to compose this hash for avoiding the malicious nodes change or delete the data on blockchain. For computational integrity, a verifier sends a task or a function and input to a prover that executes the computation on behalf of the verifier then returns the output along with a short proof. Ocean Curators take into accounts more subjective measures such as opinion and signalling. In essence, tokenized curation is a way to make opinion a scarce resource and to facilitate the discovery of assets through signalling that opinion. It's noteworthy that each layer is independent of the end-to-end integration of the stack and can activate new applications. Moreover, considering the privacy and permission issues, the data is stored off-chain and managed by the different tribles.

The authors in [150] propose a permissionless blockchain-based platform for offering high-quality data rather than quantitative data for ML professionals to buy or rent. To make the data sharing more scalable, the network is divided into multiple Interest Group (IG) as sharding solutions, where users are encouraged to gather important and relevant information about a topic that interests them. Each IG has a unique dataset, which is the amalgamation of all the data generated by each node of the IG. These members of the IG can get the rewards based on the size and relevance of the data they offer to the IG. Similar to PoW, they propose a consensus protocol called the Proof of Common Interest (PoCI), where nodes use DL and topic modeling[8] to add nodes into the IG which shares the same interests with the other members of the same IG and determines whether the data of a prospective IG member is of relevance and aligns with the interests of the IG. In this way, the proposed PoCI protocol ensures that data is relevant and provides high-quality data for ML applications.

The integration of blockchain and ML technologies brings great opportunities for health data sharing and analysis ecosystems. The authors in [151] propose a blockchain-enabled personal health data sharing system, which assesses the value of time and the combined value of personal data in the context of an AI-mediated health data exchange on blockchain. In the proposed scheme, patients can get several rewards through sharing their personal data to other healthcare institutions, such as pharmaceutical companies, hospitals, and research institutions. Particularly, all these shared data are encrypted on the user side before uploading to cloud storage using a threshold encryption mechanism for guaranteeing the user/data security and privacy. Considering the limited storage capability of blockchain, they use off-chain solution with cloud storage to keep encrypted data.

The authors in [152] propose a decentralized privacy-preserving online ML framework on a private blockchain, called *ModelChain*, to utilize transaction data for disseminating model and other meta information ((i.e., flag of the model, hash of the model, and error of the model). In the proposed framework, each of the participated nodes controls the access to a subset of local data and utilizes these data to train and transfer the ML models separately. Specifically, each block only contains one transaction which has a unique timestamp. The flags in each block including four types: initialize, update, evaluate, and transfer, which indicate the action a site has taken to a model. Considering the incentive mechanism for mining nodes, they propose a proof-of-information (PoI) consensus protocol to decide the order of the online ML on blockchain for improving efficiency and accuracy. Similar to the concept of Boosting [153], in the proposed PoI protocol, the node which contains data that cannot be predicted accurately using current partial model probably contains more information to improve the trained model than other nodes should have a higher priority to update the trained model in next epoch. Therefore, *ModelChain* can improve the security and robustness of distributed privacy-preserving healthcare predictive models.

To improve the learning process and the interaction capabilities, robotic systems need to access a number of private data and support knowledge sharing among multiple robots, which may increase the privacy concerns in current systems. To address the privacy issue in Human-Robot Interaction (HRI), combining ML and blockchain technologies, the authors in [154] propose a *RoboChain* - a learning framework for secure, decentralized, and computationally efficient data and model sharing among multiple robot units. In *RoboChain*, users could check whether the generated information compromises their privacy during the interaction with the robot. Leveraging federated learning, *RoboChain* could perform ML operations in a distributed way without the need of storing the data in a centralized location. Specifically, in *RoboChain* network, robots connect to local hubs, which are interconnected in a sparse network. Once a robot derives a new candidate ML model validated by the network, the robot is responsible to send a transaction to blockchain including the

[7]Computational integrity is defined by correctness, soundness, and zero-knowledge, where correctness means that prover can convince verifier concerning a true statement and soundness means that prover cannot convince verifier of any false statement.

[8]Topic modeling is a text-mining tool for discovery of hidden semantic structures in a text.

key information such as timestamps, hash of the new model, and encrypted data, to verify the consensus process.

The authors in [155] propose a framework called *ADVOCATE*, to facilitate handling personal data in IoT environments. Specifically, the proposed framework gathers and analyzes the policy data and recommend rules for making decisions and developing user-centric solutions via ML technologies. By leveraging blockchain technology, the proposed framework could assist the data controllers and processors for managing data in a transparent, unambiguous, and auditable manner. All the non-repudiation and the versioning of consents would be digitally signed by contracting parties to ensure the non-repudiation and the hashed version of the consents would be submitted to a blockchain infrastructure to ensure their integrity and anonymity.

On the other hand, the trustworthiness of data also plays an important role in ML algorithms. In the exploration of data resources, ML approaches need better data sources for training models to solve the problems more effectively. However, the goal of high-accurate and privacy-aware data sharing for ML algorithms is still challenging in current complicated and trustless networks. For the privacy and reputation issues, most users are not willing to share their data with the public. Taking advantages of blockchain, it becomes possible for solving the trustworthiness of big data issues. Specifically, every transaction on the blockchain is checked, verified, and stored based on the one-way cryptographic hash functions in the distributed networks. These ever executed transactions are irreversible and nonrepudiable after the verification. Thanks to these main features of blockchain, it enables provenance on training data and models and significantly improves the trustworthiness of the data and models. *CognitiveScale* [156] leverages blockchain technology to securely store the results of an ML algorithm that it builds for regulatory compliance in the financial markets world.

B. Blockchains can Benefit Machine Learning for Security and Privacy

Both security and privacy preservation are key factors for using ML-based methods in communications and networking systems. Currently, ML systems generally adopt a centralized architecture, which is prone to hacking, as any malicious node only needs to break into a single system to manipulate the instructions. Moreover, since the training data usually involves a large amount of personal information, the data breaches may lead to the privacy concerns of personal data. How to protect the training data put into ML models from hackers is an important issue.

Thanks to the cryptography embedded in blockchain, data stored in blockchain are highly secure. It is ideal for storing the highly sensitive and personal data on blockchain which provides the encrypted storage of data on a distributed ledger. Malicious data cannot be recorded in the blockchain by applying efficient consensus protocols. In addition, blockchain provides a promising solution to track the history of information modifications that is characterized by interesting features, such as transparency and auditability. Recently, blockchain-based approaches for secure data management have gained significant attention and been well investigated in various scenarios, such as healthcare systems [19], [157], IoT [158], [159], cloud storage [160], [161], intelligent transportation systems [162], and identity management systems [163]. The authors in [43] investigate blockchain-based solutions for security services, including authentication, confidentiality, privacy, data and resource provenance, and integrity assurance. These services play an important role in distributed applications, especially for the network that involves a large volume of processed data.

Training an ML model usually requires a number of data generated by multiple entities, which may increase great privacy concerns. For the privacy issues, in blockchain systems, each node interacts with a generated pseudonymous address to avoid identity exposure. By exploiting pseudonyms, blockchain can provide pseudonymity and be suitable for some use cases that require highly privacy [64], [66]. In addition, based on cryptographic mechanisms, the privacy of data/model owners is preserved and the secrecy of data/model sharing among multiple service providers is guaranteed [164]. Hyperledger Fabric integrates zero-knowledge proof (ZKP) into a wider range of applications to accommodate privacy-preserving asset management with audit support (also known as zero-knowledge asset transfer, or ZKAT) to address the privacy issues [165]. ZKAT is built on top of anonymous authentication mechanisms offered by Identity Mixer, which leverages ZKP to offer anonymous authentication for clients in their transactions.

To address the data privacy issues, the authors in [166] utilize blockchain technology to build a privacy-preserving SVM training scheme, called *secureSVM*, where IoT data is encrypted locally and recorded on a distributed ledger. By implementing the homomorphic public-key crypto-system, Paillier,[9] they designed secure building blocks and constructed a secure SVM training algorithm without the need of any trusted third party. The proposed scheme can ensure the confidentiality of the sensitive data and the SVM model parameters for data analysis. To address the security and robustness vulnerabilities, the authors in [167] design a blockchain-powered privacy-preserving DL framework. Then they propose two algorithms, including initial benchmarking and privacy-preserving collaborative DL, to deal with fairness in collaborative DL. During initial benchmarking, each party trains a local differentially private generative adversarial network (DPGAN) and publishes the generated privacy-preserving artificial samples for other parties to label, based on the local credibility and transaction points. In the proposed framework, they also propose a three-layer onion-style encryption scheme to guarantee accuracy, privacy, and system robustness.

Federated learning, which supports the different devices upload and shares the training and knowledge of the model, brings a number of concerns on privacy issues. To address security and privacy issues in DL applications with federated learning, the authors in [168] propose a secure and decentralized blockchain-based framework with cryptographic technique, called *DeepChain*, for privacy-preserving distributed

[9]Paillier is an efficient additive homomorphic cryptosystem.

DL. Specifically, *DeepChain* can securely aggregate local intermediate gradients from untrusted parties through creating transactions and introduce an incentive mechanism for enabling workers (an entity in *DeepChain*) to process the transactions and update local training and parameters. Meanwhile, by using cryptographic techniques, *DeepChain* ensures data confidentiality and auditability of the collaborative training processl. Moreover, *DeepChain* integrates the tool of non-interactive ZKP to provide auditability of the collaborative training process, timeout-checking and monetary penalty mechanism of blockchain to push participants to behave fairly.

The authors in [169] explore the blockchain technology and propose a decentralized privacy-preserving and secure ML system, called *LearningChain*, to prevent the data privacy from malicious attacks. In the system, a decentralized Stochastic Gradient Descent (SGD) mechanism, which considers differential privacy and *l*-nearest aggregation algorithm, is designed to learn a general predictive model over the blockchain for protecting the data privacy and protecting the system from potential Byzantine attacks. Specifically, the proposed systems contain two types of nodes, including data holders and computing nodes. Among them, data holders first create pseudo-identities and calculate the local gradients based on the current global model. Then they use a differential privacy scheme to perturb their local gradients, encapsulate them with other related information, and broadcast the messages to all the computing nodes. Similar PoW protocol, the computing modes compete to get the authority of creating a new block by solving a mathematical puzzle. The winning computing node adopts a Byzantine attack tolerant aggregation scheme to the received local gradients in its memory pool, calculates the global gradient to update the model parameters, and appends a new block to the chain. In this way, the data holders can collaboratively train a predictive model while protecting their data security and privacy in trustless networks.

With the fast development of smart end devices, several ML models may be trained and run on the end devices with limited computing capabilities. The authors in [170] propose a secure blockchain-based system to run an ML model in a decentralized and collaborative way among various nodes in a smart home. In the proposed system, IoT devices generate raw data and store the data in the temporally segmented InterPlanetary File System (IPFS) files.[10] Then, ML algorithm could learn the personalization parameters of each user and device type and generate the user device profiles, which are stored in the blockchain. The hash of the device user profile information stored on the IPFS is recorded in Ethereum blockchain. In this way, the proposed system could customize IoT devices based on users' individual preferences derived by ML algorithms and guarantee the security and privacy in smart home.

Data protection aside, access control management is also critical for providing security services in communications and networking systems. The authors in [171] propose a dynamic and fully distributed security scheme for access control in IoT. On one hand, by utilizing the blockchain, the proposed scheme could ensure a totally distributed infrastructure for access control without trusting external central entities. On the other hand, by utilizing "online learning" mechanism of RL algorithms, the proposed scheme could provide a dynamic, optimized and self-adjusted security services.

The authors in [172] propose a blockchain consensus-based scheme for performing the access control in D2D underlying cellular network, where users can act as blockchain nodes to use the consensus protocols and cryptographic mechanisms to maintain the blockchains. Specifically, based on the authentication of channel state information (CSI), there are two types of blockchains, namely, integrity chain (I-chain) and fraud chain (F-chain). In the same cell, each user collects its CSI and broadcasts the CSI to other users. Then, all the users use consensus algorithms to check the authenticity of broadcasted CSI message. When a CSI message is verified to be authentic, it will be stored on the I-chain by signing it. Otherwise, the unauthentic message is stored on the F-chain. For example, the users who intentionally advocate their CSI for a larger amount of allocated resource could be recorded in F-chain. If the CSI of a user appears on the F-chain, the mobile user can be suspected to be a fraud user (Byzantine node). In particular, to enhance the network performance and enable efficient control, a user access control scheme is proposed to maximize the area spectral efficiency with the consideration of the authenticity of CSI with blockchain technology.

C. Blockchain can Benefit Machine Learning for Decentralized Intelligence

The contemporary ML algorithms are generally built based on a centralized paradigm, where a single entity needs to have a global vision of the problem as well as all the necessary knowledge and resources for collecting large datasets, creating highly sophisticated models and solving various types of tasks. However, the centralized ML systems currently may no longer be adequate for evolving large-scale complicated parameters and may pose severe limitations in communication and networking systems. In addition, centralized approaches are often non-scalable, operationally ineffective, and easy to suffer from failure problems. Recently, decentralized ML has emerged as a solution paradigm for addressing the complicated computing and networking problems by enabling the sharing, selecting, and aggregation of distributed heterogeneous resources [151], [174]. Without depending on any centralized controller, decentralized ML could significantly improve the efficiency of training and reduce latency and power consumption. However, the current decentralized ML framework may suffer severe privacy issues. Some researchers show that the intermediate gradients can be used to infer important information about the training data [175], [176]. For example, the authors in [177] demonstrate that a curious parameter server can learn private data through generative adversarial network (GAN) learning. The authors in [178] exploit the intermediate gradients to launch linkability attack on training data since the gradients contain sufficient data features.

The decentralized nature of blockchain provides the fundamental protocols to enable decentralized ML algorithms

[10]IPFS is a protocol and network used to create a content-addressable, P2P method of storing and sharing hypermedia in a distributed system.

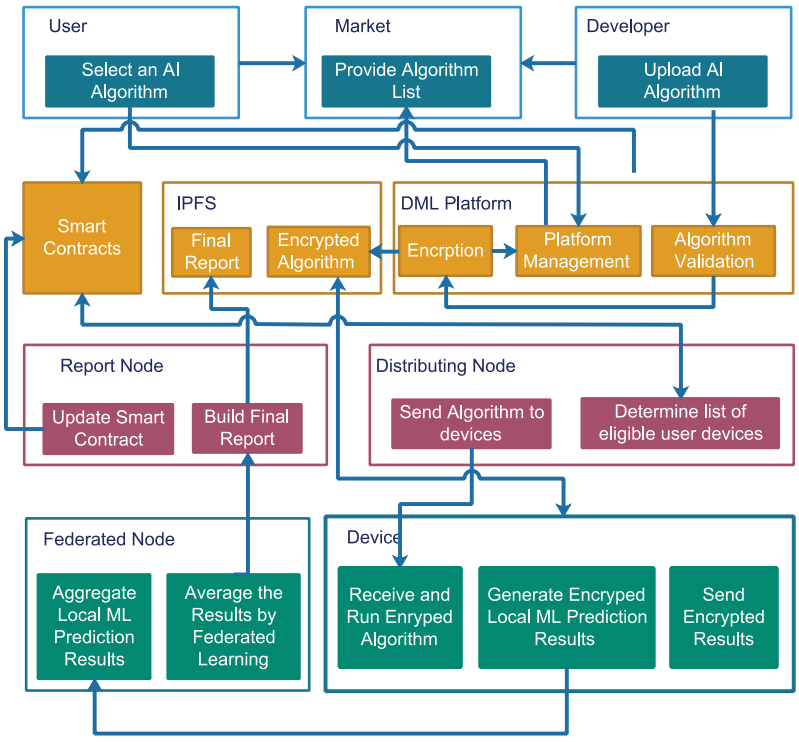


Fig. 9. The flow chart of DML Protocol based on smart contracts [173].

taking security and privacy into consideration. By applying blockchain, ML could learn, train, and derive decision-making on local devices in decentralized and distributed networks. Leveraging blockchain technology, from training to optimization, every single step in the lifecycle of ML models can be improved with certain degrees of decentralization [67]. On one hand, data held on blockchain are thus a natural fit for processing and training models by decentralized learning agents in secure and private manner. Specifically, blockchain allows verifying tampering and origin of data, making blockchain a highly trusted data source. The shared ledger, which holds strictly structured, classified, and trusted data, makes data simultaneously available at multiple locations. Additionally, the blockchain permissioning constrains the agents to appropriate data partitions based on identity and granted access rights. In this approach, the data are made available for learning, while the operators can satisfy the required privacy, regulatory and other requirements. On the other hand, blockchain facilitates decentralized ML to provide a reliable and permissioned medium for the exchange of the model parameters and maintain an association between the model parameters and the mini-batches used to train those parameters. The immutability and distributed consensus models of blockchain technology intrinsically introduce a level of trust and enables collaborative optimization in ML applications. By subjecting agents to random verification (or audit), poisoning attacks can also be discouraged leading to better trust in the global optimized model. Utilizing the OpenMalaria framework, the authors in [179] demonstrate the capability of blockchain to store, share and maintain auditable logs and records of each step in ML and computational simulation process. While decentralized ML remains a nascent and still experimental market, there are several initiatives that have gained some notoriety beyond the research phase, such as healthcare systems [180], smart home [170], Industrial IoT (IIoT) [181], vehicle networks [182], and so on.

The *SingularityNET* [183] implements on Ethereum blockchain and creates a protocol shared by each AI product on the network to build an AI-as-a-Service marketplace and enable decentralized intelligence. To maintain network validity, these agents can be incentivized to perform tasks optimally, through the implementation of a reward-based benefit rating system based on some factors. In addition, to make transactions as simple as possible, *SingularityNET* employs application programming interface (API) to incorporate standard AI services, like image and language processing, and other types of services into smart contracts. In this way, both parties have the security of knowing that the other parties fulfill their part of the agreement. The authors in [184] take *SingularityNET* platform as a case study and discuss that decentralized ML for the improvement of ML. In particular, smart contracts and DAPPs may provide new opportunities to model the interactions between different entities in an ML application.

By leveraging blockchain and smart contracts, the authors in [173] propose decentralized machine learning (DML) protocol and build a distributed ML system that connects participants directly without the need of any centralized third party. As shown in Fig. 9, in this system, individual devices run the ML algorithms independently and directly via using their processing power and private datasets. According to the ML requirements stated in the smart contracts, distributing nodes will identify appropriate data owners and distribute the algorithms, which are homomorphically encrypted, to

appropriate data owners' devices as identified. Leveraging federated learning, the federated nodes aggregate the encrypted prediction results, formulate and validate the comprehensive prediction, and send the encrypted aggregated results to the report node which shares and stores the final report in IPFS as well as automatically updates the smart contract. They also introduce an incentive scheme to reward participated nodes in collaborative training process automatically. The proposed system provides a feasible solution for implementing DML algorithms effectively and protects personal data from privacy leaks.

The *OpenMinded* [185] is also a recent blockchain-powered project that aims to decentralize ML for providing valid data which improves the actual model. The *OpenMinded* mainly contains four parts, including Capsule, Syft, Mine, and Sonar. Specifically, Capsule is responsible for generating public and private keys and ensuring that Sonar neural network stays encrypted. Capsule receives the trained model from the Sonar along with the organization's public key and initializes the trained model by generating the homomorphic keys which will be stored on the IPFS. Syft is a homomorphically encrypted DL library and contains neural networks that can be trained in an encrypted state. In *OpenMined*, each miner owns its Mine which represents the data source that can be used to improve the ML model. The miner is constantly checking with the Sonar running on a blockchain to find out new neural networks to contribute. If a new neural network is found by the miner, it downloads the model from the IPFS, trains the model locally, submits the gradients to IPFS, and receives an address. Then the miner will send the IPFS address of gradient to Sonar. Sonar with the help of Capsule, sorts encrypted loss corresponding to different gradients it received and decides how each gradient to be weighted. Using this information about the actual contribution of each miner, the reward for miners will be executed on the smart contract of Sonar. So the owner of the actual data can be benefited for every useful contribution. Once training is complete, the encrypted model is sent back to the Capsule where it is getting homomorphically decrypted using the private key and then encrypted with the public key of the organization which initiated the campaign and sent to the organization. Thus, combining DL, federated learning, blockchain, and homomorphic encryption, the *OpenMinded* could enable data owners to keep their data private during the decentralized ML model training process and allow models to be trained, encrypted, and shared within insecure and distributed environments.

The *Effect.AI* [186] platform leverages EOS blockchain,[11] which uses dBFT consensus and smart contracts, to provide decentralized runtime to AI applications. In the *Effect.AI*, AI developers access to a large workforce to train AI algorithms. Once a worker completes a task, it can get rewards. Each algorithm has its own wallet to allow for the acceptance of transactions. In addition, the algorithms are able to communicate and collaborate with other algorithms and purchase services from each other. They also use a reputation-based mechanism to avoid malicious users submitting the work with bad quality. Workers with a higher reputation score will be able to apply for higher rewarding tasks than workers with a lower score.

The *Algorithmia* [187] is a blockchain-based protocol for evaluating and purchasing ML models on a public blockchain such as Ethereum. It provides high-quality ML models to be shared and accessed by individuals. In particular, *Algorithmia* offers these services in two forms: 1) Serverless AI Layer for setting up models in the cloud; 2) Enterprise AI Layer for setting up the model in any public or private cloud. *Algorithmia* introduces a protocol called Danku (Daniel + Kurtulmus) contract protocol to allow users to solicit ML models. It establishes a marketplace for exchanging ML models in an automated and anonymous manner for participants. On *Algorithmia*, users can request an AI/ML model that performs a specific function while ML developers can download the sample data to train their models and then submit the models back to the blockchain. It's noteworthy that ML models are executed and evaluated on the blockchain and the reward is paid when a suitable model is presented.

AI researchers from Microsoft proposed *Decentralized & Collaborative AI on Blockchain (DCAI)*, which is a framework to host and train ML models on Ethereum blockchain and leverages smart contracts for enabling decentralized training mechanisms in ML models [67]. Specifically, DCAI structures the process of adding data/training to an ML model based on three main components: 1) The Incentive Mechanism, which is designed to encourage the contribution of high-quality data and is responsible for validating the transaction; 2) The DataHandler, which stores data and metadata on blockchain for enabling the data is accessible for all future uses; 3) The Model, which encapsulates a specific ML model, which is updated according to predefined training algorithms. Leveraging blockchain technology, DCAI framework reliably executes an incentive-based system to encourage participants to contribute data that will help improve a model's performance and offers participants a level of trust and security. With the appropriate blockchain programming, all steps from data entry to conclusions can be observed, and the observing party will be sure that this data has not been tampered with.

Considering the computational intensive of the ML programs, the authors in [188] propose a crowdsourcing blockchain-based scheme to enable the decentralized ML without the need of any trusted third-party. To address the interaction of workers (blockchain nodes) and free-riding problems in crowdsourcing system, a non-cooperative game-theoretic solution with two workers that audit each other and a cryptographic tool of commitment are proposed. Moreover, the expensive and randomized computation is crowdsourced via the application layer to be "asynchronously" executed and separated from the consensus layer. The full nodes/miners can simply put the output into the coming block as the result is submitted, and do not need to wait for the execution to finish during mining.

Leveraging blockchain technology, the authors in [189] propose an asynchronized cooperative decentralized DL architecture, where the contributors can train DL models with

[11]EOS is a blockchain platform and cryptocurrency designed to build a scalable network of decentralized applications.

private data and share them to the cooperative data-driven applications. Shared models are fused together by learning-model fusion mechanism, in which the feature vectors to be fused present features with the highest level of abstraction, to obtain a better model. Ethereum blockchain is applied to provide much higher, collective processing power and grants access to large amounts of data for devising a decentralized and efficient DL mechanism. They also design a ZKP-enabled encryption interface for providing homomorphic encryption to further improve the security and privacy.

D. Blockchain Can Benefit Machine Learning for Trusted Decision-Making

As ML algorithms become much smarter to handle complicated communication and networking problems through training and learning, it will become extremely difficult for humans to understand how these programs come to specific conclusions and decisions. Despite the advances made by ML in many fields, the lack of trust would severely limit the performance of ML solutions. To address this issue and make sure the ML work effectively, the trust mechanisms and audit processes should be well designed for ML solutions.

Recently, blockchain's decentralized, transparent and cryptographic features provide a feasible way to record the decision-making process of an ML on a blockchain with confidence that the records have not been tampered. Through the use of blockchain technology, those transparent immutable records of all the data, variables, and processes used by ML applications for their decision-making processes can be reviewed and audited at any time by authorized nodes with access to the system [68]. In this way, it provides a feasible solution to audit the processes of the ML applications and significantly improve the trustworthiness of the data and models for gaining public trust. Blockchain provides a route that traces back the machine decision process and helps improve the ML effectiveness by providing a secure means for sharing data as well as models [190]. Therefore, ML techniques that utilize blockchain can facilitate a trust and secure sharing of knowledge and decision outcomes across a large number of autonomous agents, which can contribute, coordinate, and vote on further decisions [191], [192]. Recently, a lot of researchers have proposed a lot of blockchain-based trustable ML mechanisms, where blockchain technologies hold the promise of enabling trusted decision-making in ML-based algorithms.

The authors in [193] investigate a set of fundamental aspects in blockchain-based trust management systems, including transparency, provability, and explainability. *TrustChain* [194] is proposed to establish trust in an IoT ecosystem, which involves four entities, namely, manufactures, retailers, customers, and data analysis companies. By utilizing blockchain and smart contract technologies, *TrustChain* develops a trustworthy end-to-end trading platform to manage devices and data for IoT ecosystems. For intelligent vehicle networks, the authors in [15], [195] propose feasible blockchain-based trust mechanisms to deal with the security and privacy issues. The authors in [4] discuss the trust issues for blockchain-based

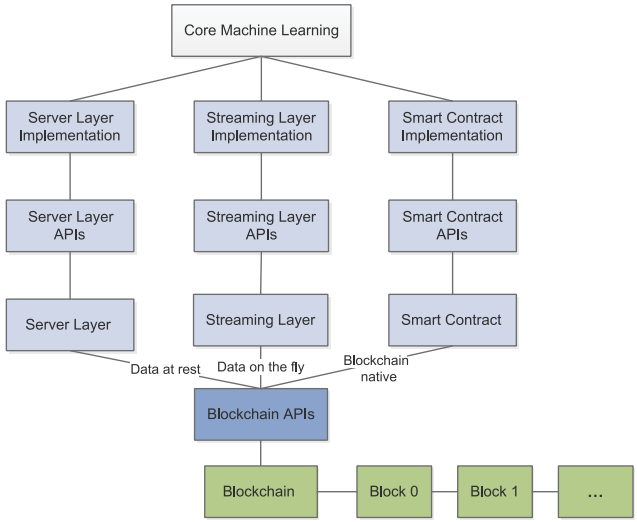


Fig. 10. The analytical framework for trustable machine learnings [197].

methods in healthcare systems. Using smart contracts, which run inexorably and are not open for breach or discretionary compliance, a data movement framework for building the trust in cloud is proposed in [196]. With the appropriate blockchain programming, all data and models processed by ML solutions can be observed, and all the entities in the networks will be able to audit these data and models.

The authors in [197] propose a trustable blockchain-based ML framework. As shown in Fig. 10, the proposed analytical framework for trustable ML mainly involves five parts: core ML that is responsible for model initialization, training, validation, scoring, evaluation, serialization, and cleanup; server layer and streaming layer that are computing environment hosting one or more ML activities and managing data records and requests; smart contract layer that hosts one or more ML activities in an automated and trustable way; and blockchain that is an appended-only list of immutable blocks. In the unified framework, a computer is able to translate core ML implementation from a single thread on a single machine to multiple threads on multiple machines running with blockchain, which store the data and records in a permanent and immutable manner. In particular, they use association rule mining (ARM) [198] to evaluate how trustable ML can be done with blockchain.

The authors in [199] present a blockchain-enabled system that establishes a decentralized trust among parties involved in a disease modeling process. Leveraging blockchain technology, the system allows users to track, validate, and verify ML task and results, such as trained models, outputs, policies, model parameters, and other meta-data associated with the learning process. In addition, the system has added the validation mechanism for providing trusted service that the system uses to validate and verify worker outputs. Specifically, the process engine sends the computation results to the validation engine which chooses a set of nodes, called endorsers. These endorsers recompute the report and endorse the computation results when the recomputation results is within an acceptable tolerance. The validated computations are stored on blockchain ledger, where other agents could simply review

and verify the records to ensure the consistency of the ledger.

The authors in [200] propose a trusted routing scheme based on blockchain and RL to improve the routing security and efficiency for wireless sensor networks. By leveraging blockchain technology, the packets including routing information, which are forwarded by the routing nodes, must be verified by the server nodes using PoA consensus protocol. In addition, the scheme uses an RL algorithm to learn the dynamic, reliable, and extensive routing information on blockchain and generate a dynamically updated RL model in each routing node through the updated reward value brought by the action (scheduling) of each state (packet location). Thus, routing decisions are optimized to select reliable and efficient routing links.

The authors in [201] present a design of blockchain-based decentralized trustless API marketplace that ensures fair and trusted usage of hosted APIs and averts malicious behavior of all three stakeholders - cloud vendors, API providers, and API consumers. In particular, the proposed system divides the storage and execution of models across multiple collaborating cloud vendors such that no single vendor has full knowledge of the models invoked during an API request. The system is designed to incentivize all three stakeholders to record transactions related to their actions on blockchain. All transactions related to the storage of models components across cloud vendors, invocation of API request by API consumers, execution of specific model components held by individual cloud vendors, and receipt of the final output by API consumers are stored on the blockchain. The system hosts smart contracts to automatically enforce correct execution of functionality and provide evidence for dispute resolution by a trusted arbitrator.

The authors in [202] propose a reputation-based worker selection scheme for measuring the reliability and trustworthiness of the mobile devices in mobile networks. To provide reliable federated learning, they use a multi-weight subjective logic model and consortium blockchain to store and manage the reputation for workers in a decentralized manner. To encourage mobile devices to provide high-quality data, an effective incentive mechanism combining reputation with contract theory is proposed for collaborative federated learning.

E. Key Lessons Learned

In this subsection, we briefly summarize the above discussion about blockchain-based solutions for ML and present the key lessons learned from these works and investigation.

1) Data and Model Sharing: Blockchain technology encourages individuals to share data to a secure, decentralized and distributed ledger, and then it could provide much more datasets and more available data for training ML models. The blockchain-based data and model sharing systems also enable users to have the ownership of their data and trained model. In particular, considering the operation cost and privacy issues, the ever-increasing data in communications and networking systems could be stored off the chain, while only the keyword tags or data references are stored for checking the authenticity and accuracy of the off-chain data.

2) Security and Privacy: Thanks to the crypto-modules embedded in blockchain, it is ideal for enhancing the security and privacy of ML solutions in communications and networking systems. Using cryptographic techniques, blockchain solutions ensure data confidentiality and auditability of the collaborative training process and the trained ML model. To provide auditability of the collaborative training process, some timeout-checking and monetary penalty mechanisms of blockchain can be also applied in ML solutions. Moreover, the blockchain solutions support a totally distributed infrastructure to ensure secure access control without any central third-party.

3) Decentralized Intelligence: The decentralized nature of blockchain provides the fundamental protocols to enable decentralized ML applications. By applying blockchain, ML could learn, train, and derive decision-making on various end devices in decentralized and distributed networks. In particular, smart contracts and DAPPs may provide new opportunities to model the interactions between different entities in an ML application.

4) Trusted Decision-Making: Blockchain is considered as a feasible solution to enable trusted ML applications in communications and networking systems. A general idea is that blockchain techniques enable the transparent immutable records of the training data, variables, and processes used by ML applications for their decision-making processes to be reviewed and audited at any time by authorized nodes with access to the system. In this way, the processes of the ML applications are easily audited and the performance of the system is significantly improved.

In conclusion, these blockchain-based solutions have the potential to act as the backbone for building a decentralized, transparent, secure, and trusted ML-based communications and networking systems. However, they are discussed based on reasonable concepts and are still in infancy. Some technological issues, such as the scalabilities and incentive issues still need further investigation.

VI. Machine Learning Can Benefit Blockchain Applications in Communication and Networking Systems

Although blockchain is a promising technique, it still suffers from some inherent challenges and limitations. In this section, we discuss the key challenges of blockchain and introduce the ML to enhance blockchain in some aspects, including energy and resource efficiency, scalability, security and privacy, and implement of smart contracts. The summaries are listed in Table VI.

A. Machine Learning Can Optimize Energy and Resource Efficiency

Due to the PoW consensus mechanism which is still applied in many blockchain applications, energy is a major factor in validating and sharing transactions on blockchain. On one hand, the cryptographic hashing process of every block is computationally intensive and consumes huge energy. On the other hand, to create a block, miners need to solve a "puzzle"

TABLE VI
A SUMMARY OF BENEFITS FOR BLOCKCHAINS PROVIDED BY MACHINE LEARNING

Benefits	Ref.	Use case	Category of ML	Year of publication	Purposes	Contributions
Energy consumption	[208]	Blockchain network	Supervised learning	2018	Avoiding complicated hash operation and monopoly	A node selection algorithm based on ML for saving resource and guaranteeing security
	[209]	Mobile blockchain	Supervised learning	2018	Maximize the revenue of service provider	Developing an optimal auction based on DL for the edge resource allocation
	[210]	Cloud network	Reinforcement learning	2017	Reducing energy cost	Developing a blockchain-based decentralized framework and embedding RL to minimize the energy cost
	[213]	Electricity system	Supervised learning	2017	Maintaining automated and secure energy trading	A flexible P2P energy trading with low transaction costs
Scalability	[33]	IIoT systems	Reinforcement learning	2019	Addressing the scalability issues and improving the throughput	Proposing a DRL-based performance optimization framework for blockchain-enabled IIoT systems
	[216]	Hiring system	Supervised learning	2018	Automating the hiring process	Using ML on the decentralized, encrypted Blockchain to develop an automate hiring system
	[217]	Smart ledger platform	Supervised learning	2019	Allowing the individual agents to communicate and share information efficiently	Proposing the resource lane as a sharding scheme and enabling multiple transaction chains in parallel
Security and Privacy	[221]	Blockchain system	Unsupervised learning	2017	Identifying blockchain breaches automatically	Proposing PFM with unsupervised ML to perform two-phase validation process
	[222]	Bitcoin system	Supervised learning	2017	Detecting the cybercrime activities in Bitcoin system	Estimating the proportion of cybercriminal entities using supervised ML algorithms
	[223]	Firewall system	Unsupervised learning	2017	Decentralizing firewall for malware detection	Developing a decentralized firewall system based on blockchain
	[224]	Blockchain network	AI	2019	Protecting crypto assets and detecting vulnerabilities	Building an AI power blockchain security platform
Intelligent Smart Contracts	[229]	Energy markets	Machine learning	2018	Improving the operation efficiency and automating the auction among buyers and sellers in energy markets	Proposing an AI-based blockchain smart contracts in energy markets
	[230]	Supply chain	Machine learning	2018	Managing the logistics involved and enabling intelligent smart contracts in a supply chain	Selecting the efficient and optimized smart contract via ML technology with transaction history
	[231]	Blockchain network	Reinforcement learning	2017	Enabling blockchain-driven self-aware agents-assisted contracts	Presenting a cross-organizational blockchain-agnostic framework for P2P collaboration
	[41]	Blockchain network	Reinforcement learning	2017	Supporting AI smart contract and AI inference	Using GPU/FPGA to execute the AI smart contracts which create a transparent audit trail to improve the trustworthiness of the results
	[236]	Ethereum	Supervised learning	2018	Detecting compiler bugs of smart contracts	Using CNN methods to enhance the security of smart contract
	[237]	Contract mechanisms	Supervised learning	2019	Enabling safer smart contracts	Designing a LSTM model to detect new attack trends in smart contracts

competitively. That process of solving "puzzle" usually consumes lots of energy. Although several works are going to explore more efficient consensus mechanisms with ML training process, such as proof of learning [203], proof of DL [204], proof of training quality [181], and proof of useful work [205], the huge energy and computing resource consumption are still challenging. Furthermore, building and managing the communications and networking systems based on blockchain usually need to consume much time and resource. In communications and networking systems, energy efficiency is often defined as the ratio of achievable capacity and total power consumption [206]. In terms of blockchain-based systems, it can be measured as a saving or reduction against a baseline of consumption or expense [207]. Leveraging the training data, ML-based mining algorithms may manage tasks in a more intelligent manner rather than adopting the brute force approach. Since ML algorithms can predict and speedily calculate data, it would also provide a feasible way for miners to select more important transactions to perform. Moreover, ML algorithms, which may learn the process and the architecture of the blockchain network, allow the transactions on blockchain to be executed much faster. Therefore, implementing ML in the blockchain-based system could improve the efficiency of blockchain operation.

To address the energy issues, the authors in [208] propose an energy-saving consensus protocol - Proof of AI (PoAI) and a node selection algorithm to ensure the efficiency, decentralization, and safety of a blockchain system. In the proposed PoAI protocol, there are two types of nodes, including super nodes and random nodes. The super nodes are those nodes with more powerful computational capability and the random are nodes apart from super nodes, which are still qualified to participate in transaction records and guarantee the fairness of the network. To ensure the decentralization and security of the network, they propose a convolutional neural network (CNN) framework and a dynamic threshold for selecting the super nodes to reach the consensus. Specifically, the proposed CNN framework exploits nearly complementary information of each node, calculates as well as predicts the average transaction of each node and makes statistics of threshold values of average transaction number of nodes. Based on the threshold values, all mining nodes are divided into super nodes and random nodes, which shorten the cycle of reaching the consensus.

Considering the huge energy resources consumption in the mining process of mobile blockchain, the authors in [209] proposed an optimal auction based on DL mechanism for edge computing resource allocation. In the proposed mechanism, edge computing service provider could support offloading the mining tasks from the mobile devices, i.e., miners, in the mobile blockchain environment. Through learning the miners' bids and resource occupation information, multi-layer neural network architecture is constructed and updated autonomously to perform the resource allocation and payment solutions for the miners. The proposed scheme can quickly converge to a solution at which the energy and resource efficiency of the system are highly improved.

To reduce the energy consumption in cloud data centers, the authors in [210] embed a RL-based algorithm in smart contracts and propose a blockchain-based decentralized resource management framework. In the proposed framework, the energy cost minimization algorithm implemented by a smart contract would be triggered by every incoming request. Specifically, in each learning iteration, according to the historical migrating decisions of energy cost, data centers first select an action (migrated requests and their virtual machines to data centers) and obtain the new state and reward (data centers' load and energy cost) for learning. Then, the agent broadcasts the current learned parameters which are considered as some resources recorded by transactions, and other data centers do not need to learn them repeatedly.

Aside from the above discussions, the integration of blockchain and ML may revolutionize the traditional energy sector to make it become much more efficient and smarter. On one hand, blockchain combined with IoT devices supports enables consumers to trade and purchase energy directly from the grid rather than from retailers. On the other hand, ML techniques increase efficiency and stability to numerous energy sources through analyzing large datasets in a short frame of time. They can also predict the energy consumption. Currently, serval works have investigated the efficient energy trading issues with blockchain and ML in various scenarios, including Internet of Electric Vehicles [211], smart grid [212]. Using cutting-edge ML technology, *Verv* - a P2P energy trading platform, is proposed in [213] to identify interconnected patterns of energy consumption for providing accurate predictions of consumption and optimization of energy trading. *Verv* uses neural networks, which are trained to disaggregate electricity data stored on blockchain to appliance user information, provides better predictions of energy generation, consumption, and battery activity. *Verv* could effectively improve the efficient utilization of electrical infrastructure and enable energy trading much more intelligent.

Combining ML and blockchain in communications and networking systems, devices which demand on the energy and resource could be operated by ML to reduce energy cost. These demands and transactions could also be recorded by blockchain for supporting ML to analyze and optimize the utilization efficiency of energy and resource. This has led to the development of smart grids, which are designed to handle multiple energy sources (conventional gas and renewables) at the same time efficiently.

B. Machine Learning Can Optimize Scalability

With the significantly increasing number of transactions, the scalability of blockchain becomes a key bottleneck and limits various aspects of the development of blockchain. It not only needs to reduce the time required for confirming and validating a transaction but also needs to reduce the cost required for generating a confirmed transaction. Currently, most blockchain applications such as Bitcoin cannot handle a large number of transactions in a short time and low cost [214]. It results in a series of problems for initialization of new nodes and verification of new transactions, especially in communications and networking systems, where a massive number of users need to be served and the network scales up fast. Considering the dynamic networking systems, where the nodes may frequently send updated transactions, the scalability issues should be well addressed [43].

To deal with this issue and support practical applications, some approaches like several efficient consensus protocols, the lightning protocol, sharding, and pruning are proposed. The comparison between these approaches on their abilities to overcome scalability limitation and highlighting major problems in these approaches are well investigated in [215]. In lightning protocol, the light nodes just need to store a subset of the blocks, and the full nodes are responsible for operating the complete blockchain. Sharding is also a prevalent technique to increase the scalability of distributed systems. Blockchain sharding [62] splits the entire state of the network into partitions called shards, where each shard contains its own independent state and transaction history, and nodes that manage a shard maintain information only on that shard in a shared manner without the need for loading the information on the entire blockchain. This would essentially split up the network into smaller sub-chains. In pruning solutions, the older blocks have to be dropped from the memory while the index and block database entry created for these blocks remain when newer blocks are generated.

Since each block contains a lot of transaction data, ML could work perfectly with the growing sophistication of blockchain scalability. By analyzing these large-scale data produced by blockchain system, ML can assist the blockchain in making better decisions for more efficient sharding or pruning solutions. In addition, some decentralized learning systems such as federated learning, which utilizes computing and storage resources on the user's device, can be used to enhance scalability. The local copies of the model on the device eliminate network latencies and costs incurred due to continuously sharing data with other nodes.

To address the scalability issues and improve the throughput in industrial IoT (IIoT), the authors in [33] propose a DRL-based performance optimization framework for blockchain-enabled IIoT systems. Considering the four-way tradeoff, i.e., scalability, decentralization, security, and latency, they quantify the performance of blockchain systems. Then, to improve the performance and facilitate a wide range of applications in blockchain-enabled IIoT systems, a DRL-based algorithm is designed to select and adjust the block producers, consensus algorithm (Byzantine fault-tolerant replication protocols), block size, and block interval. However, the consensus protocols with proof of concept are not considered in this modulable blockchain system.

TalentSnap is proposed in [216] for developing a highly decentralized, secure, and scalable P2P hiring system with decentralized AI technologies. *TalentSnap* has three layers, including blockchain layer, AI layer, and Web layer. The AI layer of *TalentSnap* is to generate scalable models on top of the blockchain, which consists of independent AI agents that work with encrypted data in a decentralized way. Leveraging federated learning and homomorphic encryption technologies, *TalentSnap* supports the users to train these AI models on their secure data without the need to decrypt it and get the results proceeding via sending them the AI models. To create a more scalable model, *TalentSnap* introduces a blocker, which is treated as a data structure. It takes in a search query and quickly prunes a large number of candidates that are unlikely to fit the search criteria. The search space would be restricted and the models become much more scalable via differential AI models used by the blocker.

Fetch [217], which is a decentralized self-organizing scalable ledger platform, allows the individual agents to communicate and share information. The underlying smart ledger combines elements of blockchain and DAG technologies with built-in AI and a consensus mechanism called useful Proof of Work (uPoW). Unlike traditional blockchains, the uPoW computations are "redeployed to perform the ML tasks that give the ledger its intelligence". To address the scalability issues, *Fetch* introduces the resource lane as a sharding scheme and enables multiple transaction chains (resource lanes) in parallel with a cross-chain synchronization mechanism. Being different from the traditional sharding, in *Fetch*, a transaction may be assigned to several different resource lanes simultaneously. It uses a pre-evaluation module to select the resource lanes for the transaction. The efficiency of this process can be improved by using ML algorithms trained on historical data. Although transactions are separated, their original order is saved. Thus, transactions need to be parallelized as much as possible while staying in a consistent order. Rather than solving hashing problems to mine new blocks, miners on *Fetch* network compete to solve the transaction lane puzzle. The node that finds the optimal way to parallelize the transactions across the transaction lanes gets the reward. The system's capacity is directly proportional to the number of available resource lanes. Then, *Fetch* groups transactions by hashing the resource identifiers to minimize resource correlations.

C. Machine Learning Can Improve Security and Privacy

Although blockchain is almost impossible to hack, its further layers and applications are not secure, i.e., the DAO [218], Mt Gox [219], Bitfinex [220]. Moreover, with the increasing number of personal data stored in blockchain-based communication systems, privacy becomes another critical issue. In particular, current blockchain applications often require transactions and smart contracts to produce metadata, which may be used to disclose some information, even if the data itself is encrypted.

The authors in [221] propose probability based factor model (PFM) that is implemented over the blockchain with unsupervised ML to automatically identify breaches that may result in damage and have a high probability for recurrence. Using the factor analysis and stochastic modeling, the PFM performs two-phase validation process based on historical data in issue a court injunction. First, the significance and the probability of a breach are evaluated to determine whether the breach has the potential to create substantial damage. It the evaluated value is high, PFM invokes a transaction and executes a function in a smart contract that results in the issue of a court injunction. Therefore, the proposed model could benefit enterprises that view a breach of contract as a limiting factor for implementation of smart contract in cyber-physical system or IoT.

To address the cybercrime activities in Bitcoin system, the authors in [222] estimate the proportion of cybercriminal entities using supervised ML algorithms. The methodology has three main parts, including the data pipeline, the classifier selection and assessment, and the final output production. Initially, the clustered addresses, categorized and uncategorized data are retrieved in the data pipeline using pre-process techniques for producing ready-to-use datasets for the ML models. Then a number of different classifiers are tested in the classifier selection and assessment to form a list of top four classifiers. Furthermore, the classifier trained with categorized observations predict the category of uncategorized observations and produce charts with the resulting labels. The prototype could produce an analysis and investigations filtering the list of suspicious addresses by a number of categories based on the priorities of investigators. Moreover, the prototype can flag entities that are more likely to participate in cybercriminal activities and be used to predict future cybercrime trends.

Combining blockchain and DL technologies, the authors in [223] propose a decentralized blockchain-based firewall system powered by a malware detection engine. The proposed

system models the files as grayscale images and uses a deep belief neural network (DBN) and restricted Boltzmann machine to classify the images into the malicious or benign. In addition, the proposed system uses a PoW-based consensus mechanism to obtain the final result of whether to allow or block a file. The combination of blockchain with the ML-based detection engine guarantees full security over a variably sized network.

Anchain [224] is an AI-enabled blockchain security platform for providing proprietary AI, knowledge graph and threat intel on blockchain transactions. There are two basic components, namely, Situational Awareness Platform (SAP) and Smart Contract Auditing Platform (CAP). The SAP is responsible for detecting and even predicting vulnerabilities and threats before and after they occur. The CAP is a cloud-based smart contract auditing sandbox that scans most known vulnerabilities, such as re-entrance and overflow, in an automatic and fast manner. They ca access to the cloud and connect to professional auditing experts. The sandbox, which is designed as "Virtual Machine", allows smart contract developers (currently Solidity) to quickly scan their smart contracts for known vulnerabilities in a fully automated manner.

The application layer vulnerability aside, data encryption mechanism also plays an important role in blockchain systems. Blockchain systems could be more secure by creating robust ciphers via computational intelligence improved system attack-defense process. An emerging field of ML is designing algorithms which are capable of processing data while it is still in an encrypted state, such as *OpenMinded* [185]. In this way, the security risks could be greatly reduced since any part of a data process is exposing encrypted data.

D. Machine Learning Can Make Smart Contracts Smarter

With the advancement of Turing-complete blockchains, smart contracts are considered as automation tools for many business-related processes that involve complex transactions. Residing on blockchain, smart contracts are programs that can be automatically executed based on parameters agreed to by two or more counter-parties without any intermediary. When the conditions in the contracts are met, relevant actions could be triggered actively. However, smart contract technology is still impossible to implement as the basis of real-world agreements to date due to the following reasons.

1) Legal Viability: Current smart contracts are still unavailable for enforcing contracts like the actual legal contract. In addition, there is still an issue for transforming the agreements of the smart contracts into the physical documents written in a natural language.

2) Flexibility Management: Since smart contracts are programs built on code immutably, the items and conditions are not subject to real-world variability and constructability. Existing smart contracts cannot handle the large volumes of data or complex business transactions.

3) Difficult Implementation: According to users' demands, smart contracts have to encompass the entire lifecycle of the contract, allowing the actions to branch into different scenarios. However, applying in more complex use cases needs to construct more complex smart contracts. Then the sophistication of the program may result that smart contracts are not easily understood and adjusted with various scenarios.

To deal with the above issue, ML techniques bring a set of opportunities to the implementation of smart contracts. ML provides a feasible way to create and execute complicated smart contracts and make them more effective. NLP techniques, such as shallow semantic parsing [225], named entity recognition [226], co-reference resolution [227], word sense disambiguation [228], and a host of other techniques, are beneficial for smart contract construction. ML and NLP could provide a promising solution to generate the smart contracts, negotiate and agree to terms on behalf of people. On one hand, NLP technique, which generates and interprets message or information without a human, could apply to construct self-writing smart contracts to make exchanges of money, property, shares anything of value, really seamless, safe, and more cost-effective. It is conceivable that NLP could be used to create partially or even whole self-executing codes and parameterization by parsing a traditional "human" contract to generate the smart contract. On the other hand, these contracts may be programmed to negotiate terms for price and quality of certain goods using ML-based game-playing algorithms. Parameters can be established for certain gap filler terms such as ranges of price and range of quality that can be adjusted dynamically, and fixed inputs by users for the type of goods. ML-powered negotiation agents on behalf of real people, as offer and acceptance are both the cornerstones of contracts. Therefore, the interaction of blockchain and NLP has widespread implications for the future of legal agreements.

The authors in [229] propose an AI-enabled smart contracts applied to electricity infrastructure and the array of networked things. By collecting and processing the data generated by billions of IoT devices, the proposed AI-based solutions could analyze data sets from a massive number of variables and construct them into weighted relationships. In particular, using ML computer-based neural networks, the data patterns among these variables are better understood in the electricity infrastructure. The AI-enabled smart contract has significantly improved the efficiency and automated the auction among buyers and sellers in energy markets.

The authors in [230] propose a smart logistics solution which manages the logistics involved in a supply chain. The main components of the proposed solution include three modules, namely, smart contracts, logistic planner, and condition monitoring using distributed ML framework. Smart contract provides a recommended list of suppliers for a particular item that needs to be procured. Once the list of suppliers has been selected, smart contract systems send notifications to the suppliers and initialize the negotiation process between the consumer and the supplier. A smart contract is established when the negotiation process is accomplished. The logistics planner supports the design and execution the optimal plan to fulfill smart contract. Condition monitoring using distributed ML framework deploys the intelligence and processes data at multiple levels and locations. Leveraging ML-based approach, it could select the efficient and optimized smart contract based on transaction history from past smart contracts.

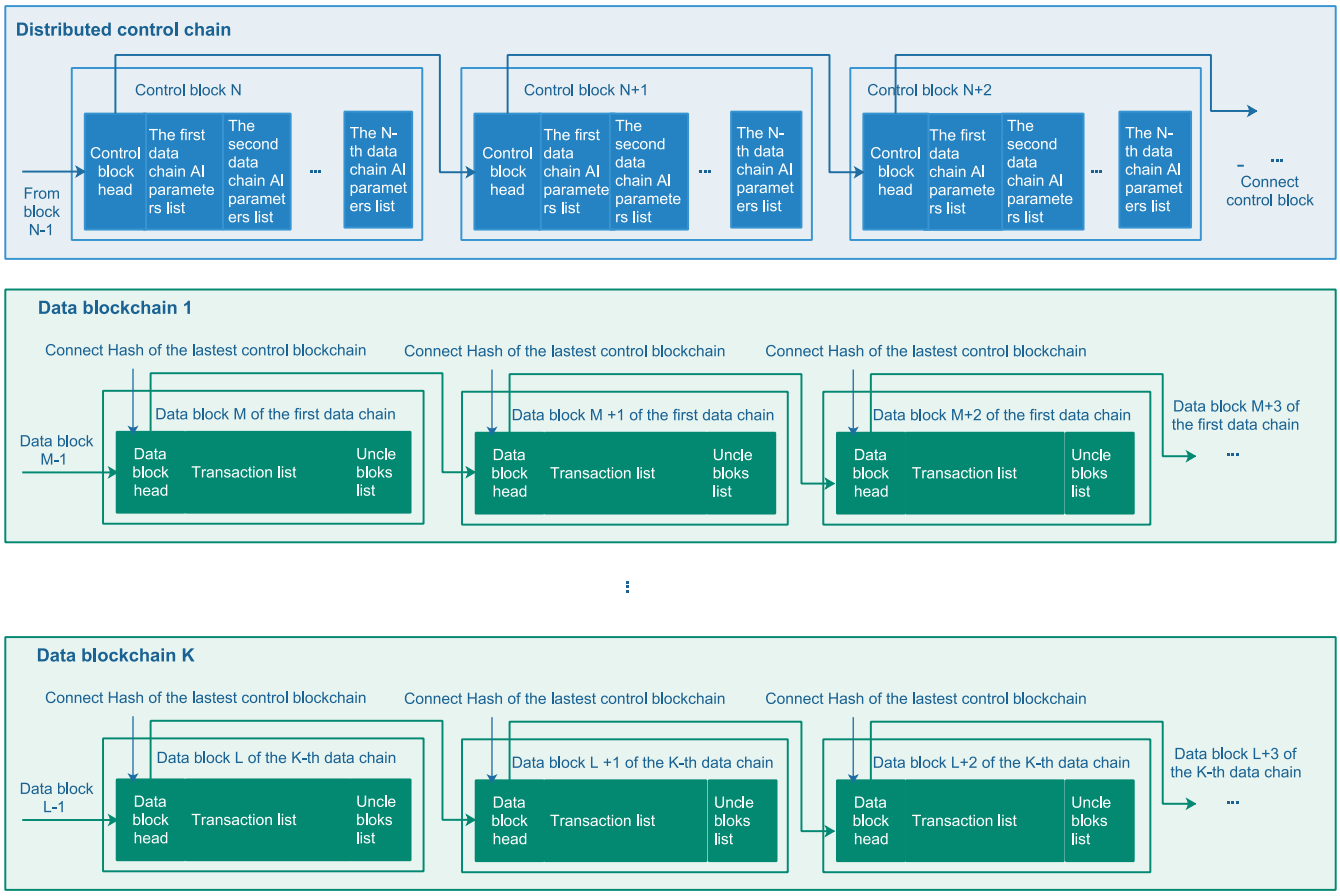


Fig. 11. A multi-chain structure combined with ML solution [231].

Recently, the rapid development of DL technology proves the feasibility of automatic code generation. *MATRIX* [231] combines DL and blockchain technologies to enable automatic generation of smart contracts. In *MATRIX*, users input the core elements (e.g., input, output, and transaction conditions) of a contract with a scripting language which could be automatically converted into an equivalent program by code generator with a CNN. Specifically, a code generator trains a CNN with a large number of labeled samples and put the script about the purpose of a contract into CNN to identify the underlying transaction patterns and data attributes. Then the discovered patterns are ordered in a sequence and then put into an recurrent neural network (RNN) which has its parameters trained with typical patterns of smart contracts. The RNN transforms the input patterns into a program based on code library containing a large number of codes for various design patterns of smart contracts. In addition, *MATRIX* enables a multi-chain platform, which includes one control chain and multiple data chains as shown in Fig. 11. It allows the multi-chain integration and interoperability of public chains being completely public and private chains being coordinated with a multitude of security access and control mechanisms. Furthermore, *MATRIX* embeds an RL-based parameter optimization engine for adjusting the blockchain parameters (block size, permission algorithm, and consensus mechanisms). The optimization paradigm ensures dynamic updating of parameters for near-optimal performance without the risk of incurring hard fork.

To expand and diversify the capability of smart contracts, *Cortex* [41] is a decentralized AI-based platform that supports AI smart contract and AI inference. AI developers can upload their models to the blockchain, smart contracts and DAPP developers can then access these AI models for constructing various types of smart contracts and DAPPs by paying CTXC, which is the *Cortex* token. Considering the complexity of AI algorithms, the *Cortex* public blockchain uses graphics processing unit (GPU) or field programmable gate arrays (FPGA) on full nodes to support AI inference. The *Cortex* virtual machines (CVM)[12] use GPU/FPGA to execute the AI smart contracts which also create a transparent audit trail to improve the trustworthiness of the result.

Meanwhile, ML techniques enable these automated processes of executing smart contracts to be orchestrated intelligently and flexibly, such as simplified negotiation, automate document assembly, visualize scenarios, contracts verification, etc. The authors in [232] present an automated contract generation toolkit without any human involvement. Furthermore, neural networks and the min-max algorithm are well applied for automatic negotiations and smart contract

[12]The CVM is fully compatible with EVM but the entire infrastructure runs on GPU.

controls [233], [234]. According to the analyzed results of existing contracts and the negotiations among different parties, ML solutions can assist the construction and updating of the smart contracts. Moreover, by introducing ML in smart contracts, computers and machines can learn from the data sources and detect the flaws in smart contracts. The authors in [235] develop an execution tool called *Oyente*, to discover the potential vulnerability in Ethereum's smart contracts based on static analysis. However, the static analysis method is less applicable because of the irreversible and immutable nature of the development of smart contracts.

The authors in [236] present an analysis of potential attacks methodology and utilize CNN tools for automatic feature extraction as well as learning and detecting compiler bugs in smart contracts. Specifically, they translate the byte-code of Solidity into RGB color code. Then, they transform them into a fixed-sized encoded image. Finally, the encoded image is fed to CNN for automatic feature extraction as well as learning and detecting compiler bugs in Ethereum smart contract. Because of the inability of smart contracts to be repaired after deployment, the early detection of relevant vulnerabilities plays a fundamental role in repairing smart contracts and reducing the security issues without deploying source code before deployment. The authors in [237] propose a long-short term memory (LSTM)-based security approach for smart contracts, which could detect new attack trends relatively quickly and improve the security of smart contracts. The objective of LSTM learning model is to perform a two-class classification and minimize the detection loss function, in order to maximize classification accuracy and detect if any given smart contract contains security threats. Smart contract threat detection, like many sequence learning tasks, involves processing sequential opcode data. As smart contracts become available in sequential order, they could be used to update the LSTM model for future contracts at each point in time. Namely, the sequence model learns the errors through the loss (e.g., logarithmic loss, softmax, and squared loss) provided by each training data point. Therefore, the proposed model based on DL could effectively detect smart contract security threats.

E. Key Lessons Learned

In this subsection, we summarize the above discussion about the ML-based solutions for blockchain and present the key lessons learned from these works and investigation.

1) Energy and Resource Efficiency: The ever-increasing data size on blockchain brings the need for more storage resources. Offering more efficient data sharding or pruning solutions, ML techniques assist the blockchain in making better decisions with maintenance and data storage. In addition, leveraging ML techniques, blockchain applications can support predictive analytics to ensure the requirements for energy and resources to be accurately met and then improve the efficiency of blockchain operation.

2) Scalability: Scalability is an essential requirement for communications and networking systems that involve large-scale devices and data. With an explosion in data and increasing user requirements on low latency and high data rate, ML solutions can support intelligent learning systems such as federated learning, to provide optimal data sharding and pruning solutions and make the system more efficient. In addition, ML techniques can enable more efficient off-chain solutions or adjusting the block size dynamically [33] to make the blockchain-based system become more scalable.

3) Security and Privacy: The application of the ML in the blockchain-based communications and networking systems can identify and prevent theft, fraud, and illicit transactions in the blockchain, by developing and training ML algorithms that can detect anomalous behavior effectively.

4) Intelligent Smart Contracts: ML provides a feasible way to create and execute complicated smart contracts and make them more effective. NLP techniques, such as shallow semantic parsing [225], named entity recognition [226], co-reference resolution [227], word sense disambiguation [228], and a host of other techniques, are beneficial for smart contract construction. It is conceivable that ML could be used to smart contracts for its creation and also for enhanced verification.

Although these ML-based solutions have been presented to promote the development of efficient, secure, and intelligent blockchain-based communications and networking systems, they are still in their infancy. Many ML-based solutions are presented based on reasonable concepts, and lack practical development and quantitative analysis.

VII. Open Issues, Research Challenges and Broader Perspectives

While research on integrating blockchain and ML technology for communications and networking systems is still emerging, a lot of open issues and challenges that need to be addressed carefully by future efforts. In addition, many other techniques are affecting the developments of blockchain and ML. At the same time, both of them may also have significant impacts on each other. In this section, we first discuss some of the most important open issues and research challenges to improve the performance of blockchain as well as ML. Then, we present some research opportunities in related areas with a broader horizon.

A. Open Issues and Research Challenges

1) Big Data Processing: As a critical factor in current networks, big data processing is exceedingly useful for a wide range of applications, including cybersecurity, fraud detection, IoT, and medical informatics. Processing collected data can make data incredibly more valuable in the near future. Big data is widely involved in blockchain and ML applications, where large-scale data are collected, grouped, and stored in blocks and processed as training sets for ML models.

Blockchain technology provides a new opportunity for storing data more securely with protection from modification, deletion or other attacks. In addition, blockchain makes it possible for users to monetize the data, by which users can actually own and control their data. Meanwhile, DL mechanism is a promising tool for the analysis and process of a large number of raw data. It learns and extracts complex abstractions

as data representation through a hierarchical learning process, where complex abstractions are learned at a given level. By supporting transactions of data, blockchain technology enables DL/ML algorithms to learn and develop based on large datasets for training through decentralized data marketplaces.

Several works have investigated data sharing and processing issues with blockchain and ML in healthcare systems [238], network traffic system [239], smart home [170]. Combining Ethereum blockchain and DRL, the authors in [240] propose a blockchain-enabled efficient data collection and secure sharing scheme, where DRL is used to achieve the maximum amount of collected data, and the blockchain is used to ensure security and reliability of data sharing. Each node senses nearby point-of interests and enables data sharing to achieve maximum data collection amount, geographic fairness, and minimum energy consumption by DRL technology. However, considering different data sharing and collection applications have different tasks and requirements. How to design DL/ML models for each node to execute multiple tasks simultaneously is a key issue.

In addition, data collection and processing in which raw data is usually unlabeled and un-categorized is a significant challenge and needs further exploration for training ML models and deriving solutions for a series of problems. Futhermore, large data sets usually require large storage and enormous amounts of computing resources to process them. There are some issues with moving data and computing power onto blockchain, including how to store these large-scale distributed data in a decentralized manner, how to discover and use data in huge data markets, and how to incentivize consumers to share their data on blockchain. Thus, as demand for more powerful processors increases, big data processing is a complex field, making it more difficult for researchers to adopt ML and blockchain solutions.

2) Scalability of Integration of Blockchain and ML Applications: As we discussed in Section VI-B, the scalability of blockchain becomes a key bottleneck and limits various aspects, especially in communications and networking systems, where thousands of users need to be served and the network scales up fast. For instance, in the case of achieving consensus in blockchain-based IoT networks, it not only requires high computing capabilities for handling large-scale transactions but also needs to have huge storage to record these transactions. To address the scalability issues, *Fetch* [217] introduces the resource lane as a sharding scheme and enables multiple transaction chains (resource lanes) in parallel with a cross-chain synchronization mechanism. The more transaction lanes the *Fetch* network can maintain, the faster the network can obtain. In *Fetch*, transactions are distributed among the resource groups, and their computing power is partitioned into resource lanes which independently operate transactions. Rather than solving hashing problems to mine new blocks, miners on the *Fetch* compete to solve the transaction lane puzzle. The one that finds the optimal way to parallelize the transactions across the transaction lanes gets the reward. *Fetch* combines elements of blockchain and DAG technologies with built-in AI, a consensus mechanism, and a data architecture that "can support millions of agents transacting together". However, there is no silver bullet that will solve scalability trilemma problems. It is a kind of hybrid blockchain (Smart Ledger), hybrid consensus (DAG PoS + uPoW) and sharding (Resource Lanes). However, the process of scaling, solutions on how the nodes and these technologies operate efficiently is still an open issue. For example, how to make ML model be trained among different resource lanes in a decentralized way for selecting proper resource lane for transactions should be well addressed.

Meanwhile, the small capacity of blocks may result in huge delay for many transactions. Increasing block size will also slow down the propagation speed and lead to blockchain branches. It is difficult for handling the mass of transactions and balancing the storage size and process time. As a solution, by analyzing large-scale data produced by blockchain system, ML can assist blockchain in making better decisions for more efficient sharding or pruning solutions. However, current ML approaches, i.e., reactive, centralized management, one-size-fits-all approaches, and traditional data analysis solutions that have limited capabilities (space and time) are not able to satisfy the diversified service requirements of the heterogeneous and complicated networks. Moreover, the current ML models often require the creation of the custom dataset with specific variables. It is significantly difficult for scaling the development of ML models with the exponential growth of a variety of data.

3) Security and Privacy: The security and privacy are critical aspects of communications and networking systems utilizing blockchain and ML solutions. Although blockchain is highly secure, it still suffers attacks in the application layer. For instance, the DAO, which is a program built on Ethereum platform was breached, resulting in $50 million worth of Ether being stolen [241]. As a consequence, It brought serious issues in the DAO community that led to an Ethereum hard fork. ML solutions play a significant role in security systems, and are expected to improve the performance by providing an accurate detection on suspicious activities, such as zero-day malware, new malware, and sophisticated Advanced Persistent Threats (APTs). DL approaches are also well implemented in IDSs for handling significantly large sets of encrypted data. Even though ML algorithms enable more accurate detection of various attacks, there are still a lot of challenges for implementing ML algorithms to combat malicious threats appropriately in the IDS. For a huge dataset containing both benign and malicious data, the security mechanisms for detecting attacks need to deal with high dimensionality of data to pre-process for feature extraction. Then, ML algorithms should perform dimensionality reduction efficiently. Stack auto-encoders, which is a neural network consisting of multiple layers of sparse autoencoders, may be a feasible solution in the IDS. To enable network IDS (NIDS), the authors in [242] propose a non-symmetric deep autoencoder (NDAE) solution to provide non-symmetric data dimensionality reduction and a classifier model that utilizes stacked NDAEs and the random forest (RF) classification algorithm to reduce analytical overheads and training time. However, since ML models are impossible to be trained with large data sources in a real-time fashion,

it is a crucial challenge for online attack detection, especially in large and dynamic networks. Furthermore, different security or privacy problems need different ML solutions. How to design a general training model with proper parameters and structures to address various security problems is still challenging. To solve these problems, the security and privacy approaches need to be carefully studied for ML and blockchain systems.

4) Efficient Consensus Protocols: Existing consensus protocols in middleware layers of blockchain systems, which could be roughly divided into consensus protocols with proof of concept and Byzantine fault-tolerant replication protocols (as discussed in Section III-B). To address the energy and computational consumption issues, several works are going to explore more efficient consensus mechanisms, such as proof of learning [203], proof of DL [204], proof of training quality [181], and proof of useful work [205]. *WekaCoin* is a blockchain-based cryptocurrency that implemented a distributed Proof-of-Learning consensus protocol, which achieves distributed consensus by ranking ML systems for a given task [203]. For instance, a node (supplier) publishes a task, which includes training data, testing data, a submission deadline, a performance metric (accuracy, root mean square error, etc), and a reward. Each validator in validator committee evaluates all the models published in the blockchain by the trainers, derives a candidate ranking for the selected task, and then builds a candidate block from its pool of transactions under the limited time. This consensus ranking and consensus block will be sent to the validators and reach the consensus after a certain number of iterations. In this way, computational and energy resources are used to solve useful ML tasks rather than hashing-based puzzles while creating an open repository of ML models and datasets. However, how to prevent collusion between the three main actors of the network: suppliers, trainers, and validators is still challenging.

Proof of DL (PoDL) protocol is proposed in [204] to maintain blockchain via DL training instead of useless hash calculation. Specifically, PoDL protocol is an improved PoW-like consensus mechanism with which DL power of honest miners provide tamper-proofness. A valid proof for a new block can be generated if and only if a proper DL model is produced. It makes better use of the nodes' energy and computing resources for facilitating and maintaining blockchain. However, the model requester may be generalized to multiple malicious requesters who may collude with miners. The latency of block generation cannot also be ignored since the training DL model will require a long time. Besides, more study needs to be performed with a realistic pattern of block submission and more DL models/datasets.

The *Coin.AI* [205] is also a blockchain-based cryptocurrency which introduces a proof-of-useful-work scheme. In this system, the nodes on blockchain compete to train DL models and mine a new block when the performance of trained model exceeds a threshold. The problem of generating a neural architecture during mining is solved by proposing a method that maps the hash of the previous block, the list of transactions and a nonce into such an architecture. However, due to the difference between the tasks in terms of data volume, expected accuracy and variable dimension, the strategies of nodes to obtain a puzzle solution are different from those in the PoW network.

The authors in [181] propose a proof of training quality protocol, which combines data model training with the consensus process to fully recycle the energy required achieving consensus among nodes on blockchain. To improve the operation efficiency, they introduce consensus committee, where members are selected by retrieving the related nodes for a specific data sharing request in the blockchain. These selected members are responsible for training a data model and performing the consensus process based on the work of collaborative training. In this way, the communication overhead is significantly reduced since the consensus messages are only sent to the committee nodes rather than all the nodes. However, the lack of a solid way of providing global randomness beacon for consensus committee generation may result in the system insecure.

5) Resource Management: To reach the full potential of blockchain and ML, researchers should effectively allocate and schedule resources, including computing resource, storage resource, and communication resource in the underlying networks. Since multiple heterogeneous resources may be shared based on blockchain technology among multiple participants and vendors, efficient coordination and feasible incentive mechanisms should be designed appropriately. Consensus protocols also have to be embedded carefully in these mechanisms. For example, resource allocation solutions should determine how to select mining nodes and allocate limited resources optimally to deal with some difficult issues. The authors in [209], [243], [244] have investigated resource management on blockchains by utilizing ML solutions for the PoW consensus protocol. However, the quantitative analysis of the resources required for other consensus protocol is seldom explored so far.

Another issue in resource allocation is access control, which plays a vital role in communication and networking systems. The goal of access control is to maximize the revenue of the systems and guarantee the QoS of users. Since different users have different service requirements, i.e., best-effort service and delay-sensitive service, resource allocation schemes for communications and networking systems applying ML and blockchain systems have to be dynamically optimized to satisfy the different QoS of users and fulfill specific tasks, such as training models and verify contractions. In the access control schemes which utilize ML and blockchain technologies, they need to conduct accurate estimation for users' requirements and detect the cyber-attacks to guarantee the security and privacy of the systems. Considering the trust features of blockchain nodes and controllers, the authors in [245] have investigated view change, access control, and computational resources allocation with deep Q-learning approach. However, due to the unpredictable properties of insecure and complex network environments, how to measure the trust features of blockchain nodes and controllers is a significant challenge.

B. Broader Perspectives

1) Internet of Things: IoT is the network of physical objects that are able to communicate, sense or interact with their internal states and the external environment. With the rapid development of information technology, the connected devices, including wearables and mobile devices, smart home appliances, sensors, or any other physical "things", are expected to be 500 billion by 2030 [246]. In IoT, sensors or devices can share large amounts of data in real-time. IoT not only provides services for information transfer, analysis, and communications but also allows for independent operation, coordination, and interaction with other systems. IoT has changed the lifestyle of human beings, and our society is moving towards the always-connected systems.

The combination of blockchain, ML, and IoT will bring many potential opportunities and benefits, and plays an important role in numerous ways. On one hand, IoT will be functionally necessary to connect various types of ML application together and will be even more critical in providing large-scale data in real-time streams for training models and performing decision-making. Combining massive global IoT networks and smart ML solutions, it could collect, distribute, and process big data to enable a lot of applications in various aspects, including financial, education, and healthcare service. ML methods with IoT can also analyze human behavior via Bluetooth signals, motion sensors, or facial-recognition technology and make the corresponding prediction and solutions.

On the other hand, blockchain with IoT will disrupt existing processes across a variety of aspects, such as manufacturing, finance, healthcare, trading, and industry. According to the IDC report, up to 10 percent of pilot and production blockchain will incorporate IoT sensors by 2020 [247]. Blockchain can store an immutable record of the behaviors of smart devices and offer ideal, decentralized storage to host all kinds of data generated by these IoT devices. It provides a way for autonomous operation of smart devices without any centralized controller. For example, in supply systems, as the shipment of goods involves different distribution locations, the information of packages could be recorded and maintained on a blockchain. Then all parties involved in the system can share information and status of the goods in a decentralized and distributed manner. By implementing blockchain, IoT mechanisms can track a large number of connected devices, process various microtransactions, and coordinate devices to reduce the cost of operation significantly. Blockchain technology makes IoT systems more secure, private, and resilient. Due to the limited computing power, popular blockchain networks like Bitcoin and Ethereum are often limited by scalability challenges and delays in transaction processing. IOTA [248] is a popular blockchain-based projects to build on the growing ecosystem of the IoT devices. With a vast number of available Internet-connected devices, harnessing these limited power of the connected devices available across the large IoT networks can address such common issues of blockchain systems.

2) Big Data: Big data is the information assets, where traditional data processing solutions are not capable of managing, querying, and analyzing them due to the complexity and energy consumption [249]–[251]. Using any or all three "V" words, the complexity of big data can be characterized as follows,

- Volume: A large amount of data are generated from a wide range of data sources. The volume of stored information may be terabytes, petabytes, or even exabytes.
- Variety: The various kinds of data include unstructured, semi-structured, and structured data which need to be analyzed effectively.
- Velocity: Big data is generated at a very rapid pace over time and needs to derive decision-making more frequently about that data.

With the advent of big data processing, ML-based devices and machines will get much smarter to learn, train, and make decisions. The growing size of big data also leads to consistent improvement, enhancement, and advancement in ML solutions. Moreover, big data analytics which processes large volumes of disparate and complex data comprehensively is widely used to improve the performance of ML applications and enhance ML functions in various aspects, such as wireless networks [252], healthcare systems [253], transportation systems [254], and cyber-physical systems [255].

Meanwhile, since blockchain-based applications are becoming more and more mainstream, the amount of transactional data stored within distributed ledgers becomes enormous. Conventional data storage systems are not capable of storing these large data at a low cost. To address this issue, *Storj* [256] is proposed as an open source end-to-end storage solution that pairs cloud storage with a decentralized encrypted network. In addition, the data within blockchain is valuable as it usually applies in financial services, healthcare applications, and government systems. Big data analysis will be crucial in tracking these transactions and helping organizations make better decisions. Moreover, Big data analysis could also be used to identify fraudulent operations on blockchain. Since data analysis techniques can run pattern recognition tasks from thousands to millions of blockchain interactions to identify vicious behaviors, these solutions are going to be more and more critical to make blockchain and transactions secure and legal. For instance, *DataBlockChain* [257] brings critical information to the world such as government information, industry, voting records, property, and credit bureau data. Considering the distributed feature of blockchain and powerful ML applications, the distributed big data analysis algorithms may bring a significant improvement through where knowledge benefits can be extracted from a large number of blocks. The integration of big data, blockchain, and ML technologies, which provides a suitable trained ML model operating on top of distributed, transparent and immutable blockchain-driven data layers, could bring numerous advantages.

3) Edge Computing: With recent advances of wireless communication technologies, MEC which deploys the computing resources at the edge of the network, is proposed as a promising computing paradigm in communications and networking systems. The MEC models operate at the edge of networks, including end devices, base stations, edge servers, and micro data centers. As compared to traditional cloud computing,

MEC could significantly reduce the resource cost and energy consumption for broad applications [258]. To cope with the high computational demands, as well as to improve the system response time, MEC solutions have applied various ML applications, including DL models [259], DNN models [260], and RL [261]. The exponentially increased computing capabilities of devices are beneficial to the consensus process on blockchain, like solving PoW [106], [262]. Also, by moving complex tasks to the cloud infrastructure, integrated cloud and MEC solutions could not only reduce the network traffic and latency but also improve the resource utilization and users' experience [263]. Edge computing enables parallel processing to train ML models and perform consensus on blockchain in a decentralized manner. It will be a promising research direction for jointly consider blockchain and ML in edge computing environments.

VIII. Conclusion

This paper addresses the integration of blockchain and ML, which is becoming a necessary solution that enables intelligent, secure, and decentralized sharing of data and model as well as the efficient operation of communications and networking systems. We first gave an overview of blockchain and ML, in which the basic concepts, taxonomies, and typical applications were succinctly introduced. Then we presented some key features of blockchain (decentralization, transparency, secure, immutable, etc.) that can benefit ML papers, including data and model sharing, security and privacy, decentralized intelligence, and trustful decision-making. In addition, we showed that ML could benefit various aspects of blockchain, including energy and resource efficiency, scalability, security and privacy, and smart contract implementation. Moreover, we discussed some open issues for future research, such as big data processing, scalability, security and privacy, efficient consensus protocols, and resource management. Finally, we explored some broader perspectives, such as IoT, big data, and edge computing to identify more research opportunities.

In summary, research on the integrated blockchain and ML for communications and networking is quite broad, and a number of challenges lay ahead. This paper attempts to briefly explore the technologies related to the integration of blockchain and ML at a very preliminary level and to discuss future research that may get benefit from the pursuit of this vision.

References

[1] S. Nakamoto. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. [Online]. Available: https://bitcoin.org/bitcoin.pdf

[2] F. M. Benčić and I. P. Žarko, "Distributed ledger technology: Blockchain compared to directed acyclic graph," in *Proc. IEEE Int. Conf. Distrib. Comput. Syst.*, Jul. 2018, pp. 1569–1570.

[3] X. Jiang, F. R. Yu, T. Song, Z. Ma, Y. Song, and D. Zhu, "Blockchain-enabled cross-domain object detection for autonomous driving: A model sharing approach," *IEEE Internet Things J.*, early access, doi: 10.1109/JIOT.2020.2967788.

[4] Q. Xia, E. B. Sifah, K. O. Asamoah, J. Gao, X. Du, and M. Guizani, "MedShare: Trust-less medical data sharing among cloud service providers via blockchain," *IEEE Access*, vol. 5, pp. 14757–14767, 2017.

[5] M. Atzori, "Blockchain technology and decentralized governance: Is the state still necessary?" in *Proc. SSRN*, 2015, Art. no. 2709713.

[6] W. Reijers, F. O'Brolcháin, and P. Haynes, "Governance in blockchain technologies & social contract theories," *Ledger*, vol. 1, pp. 134–151, Dec. 2016.

[7] J. Feng, F. R. Yu, Q. Pei, X. Chu, J. Du, and L. Zhu, "Cooperative computation offloading and resource allocation for blockchain-enabled mobile edge computing: A deep reinforcement learning approach," *IEEE Internet Things J.*, early access, doi: 10.1109/JIOT.2019.2961707.

[8] F. Guo, F. R. Yu, H. Zhang, H. Ji, M. Liu, and V. C. M. Leung, "Adaptive resource allocation in future wireless networks with blockchain and mobile edge computing," *IEEE Trans. Wireless Commun.*, early access, doi: 10.1109/TWC.2019.2956519.

[9] *Blockchain Technology in the Insurance Sector*, FACI, London, U.K., Jan. 2017.

[10] P. Kuo, A. Mourad, and J. Ahn, "Potential applicability of distributed ledger to wireless networking technologies," *IEEE Wireless Commun.*, vol. 25, no. 4, pp. 4–6, Aug. 2018.

[11] M. S. Ali, M. Vecchio, M. Pincheira, K. Dolui, F. Antonelli, and M. H. Rehmani, "Applications of blockchains in the Internet of Things: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1676–1717, 2nd Quart., 2018.

[12] F. Tschorsch and B. Scheuermann, "Bitcoin and beyond: A technical survey on decentralized digital currencies," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 2084–2123, 3rd Quart., 2016.

[13] T. Neudecker and H. Hartenstein, "Network layer aspects of permissionless blockchains," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 838–857, 1st Quart., 2019.

[14] M. Belotti, N. Božić, G. Pujolle, and S. Secci, "A vademecum on blockchain technologies: When, which and how," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3796–3838, 4th Quart., 2019.

[15] Z. Lu, W. Liu, Q. Wang, G. Qu, and Z. Liu, "A privacy-preserving trust model based on blockchain for VANETs," *IEEE Access*, vol. 6, pp. 45655–45664, 2018.

[16] C. Li, Y. Fu, F. R. Yu, T. H. Luan, and Y. Zhang, "Vehicle position correction: A vehicular blockchain networks-based GPS error sharing framework," *IEEE Trans. Intell. Transp. Syst.*, early access, doi: 10.1109/TITS.2019.2961400.

[17] C. Qiu, H. Yao, R. Yu, C. Jiang, and S. Guo, "A service-oriented permissioned blockchain for the Internet of Things," *IEEE Trans. Services Comput.*, early access, doi: 10.1109/TSC.2019.2948870.

[18] T. M. Fernández-Caramés and P. Fraga-Lamas "A review on the use of blockchain for the Internet of Things," *IEEE Access*, vol. 6, pp. 32979–33001, 2018.

[19] C. Esposito, A. D. Santis, G. Tortora, H. Chang, and K.-K. R. Choo, "Blockchain: A panacea for healthcare cloud-based data security and privacy?" *IEEE Cloud Comput.*, vol. 5, no. 1, pp. 31–37, Jan./Feb. 2018.

[20] K. Fan, Y. Ren, Y. Wang, H. Li, and Y. Yang, "Blockchain-based efficient privacy preserving and data sharing scheme of content-centric network in 5G," *IET Commun.*, vol. 12, no. 5, pp. 527–532, Mar. 2018.

[21] Y. Liu, F. R. Yu, X. Li, H. Ji, and V. C. M. Leung, "Decentralized resource allocation for video transcoding and delivery in blockchain-based system with mobile edge computing," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 11169–11185, Nov. 2019.

[22] R. Dennis and G. Owen, "Rep on the block: A next generation reputation system based on the blockchain," in *Proc. Int. Conf. Distrib. Comput. Syst.*, Dec. 2015, pp. 131–138.

[23] Z. Yang, K. Zheng, K. Yang, and V. C. M. Leung, "A blockchain-based reputation system for data credibility assessment in vehicular networks," in *Proc. Annu. IEEE Int. Symp. Pers. Indoor Mobile Radio Commun.*, Oct. 2017, pp. 1–5.

[24] G. Zyskind, O. Nathan, and A. Pentland, "Decentralizing privacy: Using blockchain to protect personal data," in *Proc. IEEE Security Privacy Workshops*, May 2015, pp. 180–184.

[25] A. Dorri, M. Steger, S. S. Kanhere, and R. Jurdak, "Blockchain: A distributed solution to automotive security and privacy," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 119–125, Dec. 2017.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.

[27] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014. [Online]. Available: arXiv:1409.1556.

[28] M. Chen, W. Saad, and C. Yin, "Echo-liquid state deep learning for 360 content transmission and caching in wireless VR networks with cellular-connected UAVs," *IEEE Trans. Commun.*, vol. 67, no. 9, pp. 6386–6400, Sep. 2019.

[29] M. Chen, W. Saad, and C. Yin, "Liquid state based transfer learning for 360 image transmission in wireless VR networks," in *Proc. Int. Conf. Commun.*, May 2019, pp. 1–6.

[30] M. Chen, W. Saad, and C. Yin, "Liquid state machine learning for resource and cache management in LTE-U unmanned aerial vehicle (UAV) networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1504–1517, Mar. 2019.

[31] U. Challita, A. Ferdowsi, M. Chen, and W. Saad, "Machine learning for wireless connectivity and security of cellular-connected UAVs," *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 28–35, Feb. 2019.

[32] D. Zhang, F. R. Yu, and R. Yang, "Blockchain-based distributed software-defined vehicular networks: A dueling deep Q -learning approach," *IEEE Trans. Cogn. Commun. Netw.*, vol. 5, no. 4, pp. 1086–1100, Dec. 2019.

[33] M. Liu, F. R. Yu, Y. Teng, V. C. M. Leung, and M. Song, "Performance optimization for blockchain-enabled Industrial Internet of Things (IIoT) systems: A deep reinforcement learning approach," *IEEE Trans. Ind. Informat.*, vol. 15, no. 6, pp. 3559–3570, Jun. 2019.

[34] T. N. Dinh and M. T. Thai, "AI and blockchain: A disruptive integration," *Computer*, vol. 51, no. 9, pp. 48–53, Sep. 2018.

[35] J. Wang, M. Li, Y. He, H. Li, K. Xiao, and C. Wang, "A blockchain based privacy-preserving incentive mechanism in crowdsensing applications," *IEEE Access*, vol. 6, pp. 17545–17556, 2018.

[36] H. Yin, D. Guo, K. Wang, Z. Jiang, Y. Lyu, and J. Xing, "Hyperconnected network: A decentralized trusted computing and networking paradigm," *IEEE Netw.*, vol. 32, no. 1, pp. 112–117, Jan./Feb. 2018.

[37] R. Graf and R. King, "Neural network and blockchain based technique for cyber threat intelligence and situational awareness," in *Proc. Int. Conf. Cyber Conflict*, May 2018, pp. 409–426.

[38] C. Liu, K. K. Chai, X. Zhang, E. T. Lau, and Y. Chen, "Adaptive blockchain-based electric vehicle participation scheme in smart grid platform," *IEEE Access*, vol. 6, pp. 25657–25665, 2018.

[39] H. Jang and J. Lee, "An empirical study on modeling and prediction of bitcoin prices with Bayesian neural networks based on blockchain information," *IEEE Access*, vol. 6, pp. 5427–5437, 2018.

[40] M. Swan, "Blockchain thinking: The brain as a decentralized autonomous corporation [commentary]," *IEEE Technol. Soc. Mag.*, vol. 34, no. 4, pp. 41–52, Dec. 2015.

[41] Z. Chen, W. Wang, J. Tian, and X. Yan. (2017). *Cortex-AI on Blockchain the Decentralized AI Autonomous System*. Accessed: May 1, 2019. [Online]. Available: https://www.cortexlabs.ai/Cortex_AI_on_Blockchain_EN.pdf

[42] W. Wang *et al.*, "A survey on consensus mechanisms and mining strategy management in blockchain networks," *IEEE Access*, vol. 7, pp. 22328–22370, 2019.

[43] T. Salman, M. Zolanvari, A. Erbad, R. Jain, and M. Samaka, "Security services using blockchains: A state of the art survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 858–880, 1st Quart., 2019.

[44] J. Al-Jaroodi and N. Mohamed, "Blockchain in industries: A survey," *IEEE Access*, vol. 7, pp. 36500–36515, 2019.

[45] P. Siano, G. De Marco, A. Rolan, and V. Loia, "A survey and evaluation of the potentials of distributed ledger technology for peer-to-peer transactive energy exchanges in local energy markets," *IEEE Syst. J.*, vol. 13, no. 3, pp. 3454–3466, Sep. 2019.

[46] J. Xie *et al.*, "A survey of blockchain technology applied to smart cities: Research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2794–2830, 3rd Quart., 2019.

[47] R. Yang, F. R. Yu, P. Si, Z. Yang, and Y. Zhang, "Integrated blockchain and edge computing systems: A survey, some research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1508–1532, 2nd Quart., 2019.

[48] P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza, "A survey of machine learning techniques applied to self-organizing cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2392–2431, 4th Quart., 2017.

[49] M. A. Alsheikh, S. Lin, D. Niyato, and H. Tan, "Machine learning in wireless sensor networks: Algorithms, strategies, and applications," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1996–2018, 4th Quart., 2014.

[50] F. Musumeci *et al.*, "An overview on application of machine learning techniques in optical networks," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1383–1408, 2nd Quart., 2019.

[51] M. Bkassiny, Y. Li, and S. K. Jayaweera, "A survey on machine-learning techniques in cognitive radios," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1136–1159, 3rd Quart., 2013.

[52] J. Xie *et al.*, "A survey of machine learning techniques applied to software defined networking (SDN): Research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 393–430, 1st Quart., 2019.

[53] Q. Mao, F. Hu, and Q. Hao, "Deep learning for intelligent wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2595–2621, 4th Quart., 2018.

[54] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep learning for IoT big data and streaming analytics: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2923–2960, 4th Quart., 2018.

[55] F. Pacheco, E. Exposito, M. Gineste, C. Baudoin, and J. Aguilar, "Towards the deployment of machine learning solutions in network traffic classification: A systematic survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1988–2014, 2nd Quart., 2018.

[56] T. T. T. Nguyen and G. Armitage, "A survey of techniques for Internet traffic classification using machine learning," *IEEE Commun. Surveys Tuts.*, vol. 10, no. 4, pp. 56–76, 4th Quart., 2008.

[57] Z. M. Fadlullah *et al.*, "State-of-the-art deep learning: Evolving machine intelligence toward tomorrow's intelligent network traffic control systems," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2432–2455, 4th Quart., 2017.

[58] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2224–2287, 3rd Quart., 2019.

[59] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1153–1176, 2nd Quart., 2016.

[60] P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli, "A detailed investigation and analysis of using machine learning techniques for intrusion detection," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 686–728, 1st Quart., 2019.

[61] K. Salah, M. H. U. Rehman, N. Nizamuddin, and A. Al-Fuqaha, "Blockchain for AI: Review and open research challenges," *IEEE Access*, vol. 7, pp. 10127–10149, 2019.

[62] V. Buterin. (2017). *On Sharding Blockchains*. Accessed: May 1, 2019. [Online]. Available: https://github.com/ethereum/wiki/wiki/Sharding-FAQ

[63] Q. Jia, L. Guo, Z. Jin, and Y. Fang, "Preserving model privacy for machine learning in distributed systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 29, no. 8, pp. 1808–1822, Aug. 2018.

[64] M. Wu, K. Wang, X. Cai, S. Guo, M. Guo, and C. Rong, "A comprehensive survey of blockchain: From theory to IoT applications and beyond," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8114–8154, Oct. 2019.

[65] C. Yip. (2018). *Will Greater Transparency Through Blockchain Technology Help the Banking Industry?* Accessed: Aug. 2, 2019. [Online]. Available: https://hackernoon.com/will-greater-transparency-through-blockchain-technology-help-the-banking

[66] M. Conti, E. S. Kumar, C. Lal, and S. Ruj, "A survey on security and privacy issues of bitcoin," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3416–3452, 4th Quart., 2018.

[67] J. D. Harris and B. Waggoner, "Decentralized & collaborative ai on blockchain," 2019. [Online]. Available: arXiv:1907.07247.

[68] B. Marr, "Artificial intelligence and blockchain: 3 major benefits of combining these two mega-trends," 2018.

[69] T. T. A. Dinh, R. Liu, M. Zhang, G. Chen, B. C. Ooi, and J. Wang, "Untangling blockchain: A data processing view of blockchain systems," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 7, pp. 1366–1385, Jul. 2018.

[70] R. C. Merkle, "Protocols for public key cryptosystems," in *Proc. IEEE Symp. Security Privacy*, Apr. 1980, p. 122.

[71] M. Vukolic, "The quest for scalable blockchain fabric: Proof-of-work vs. BFT replication," in *Proc. Int. Workshop Open Problems Netw. Security*, 2015, pp. 112–125.

[72] S. King and S. Nadal. (2012). *PPCoin: Peer-to-Peer crypto-Currency With Proof-of-Stake*. Accessed: 2017. [Online]. Available: https://peercoin.net/assets/paper/peercoin-paper.pdf

[73] D. Larimer, "Delegated proof-of-stake (DPoS)," Blacksburg, VA, USA, Bitshare, White Paper, 2014.

[74] *Proof of Authority Chains*. Accessed: Sep. 7, 2018. [Online]. Available: https://github.com/paritytech/parity/wiki/Proof-of-Authority-Chains

[75] M. Castro and B. Liskov, "Practical Byzantine fault tolerance," in *Proc. USENIX Symp. Oper. Syst. Design Implement.*, vol. 99, 1999, pp. 173–186.

[76] *NEO White Paper*. Accessed: Sep. 7, 2018. [Online]. Available: http://docs.neo.org/en-us/whitepaper.html

[77] D. Schwartz, N. Youngs, and A. Britto, "The ripple protocol consensus algorithm," San Francisco, CA, USA, Ripple Labs Inc., White Paper, 2014.

[78] J. R. Douceur, "The sybil attack," in *Proc. Int. Workshop Peer-to-Peer Syst.*, 2002, pp. 251–260.

[79] V. Buterin and V. Griffith, "Casper the friendly finality gadget," 2017. [Online]. Available: arXiv:1710.09437.

[80] C. Cachin, "Architecture of the hyperledger blockchain fabric," in *Proc. Workshop Distrib. Cryptocurrencies Consensus Ledgers*, vol. 310, 2016, pp. 1–4.

[81] L. Lamport, R. Shostak, and M. Pease, "The Byzantine generals problem," *ACM Trans. Program. Lang. Syst. (TOPLAS)*, vol. 4, no. 3, pp. 382–401, 1982.

[82] A. Garoffolo and R. Viglione, "SideChains: Decoupled consensus between chains," 2018. [Online]. Available: arXiv:1812.05441.

[83] V. Buterin. (2017). *Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform*. Accessed: Oct. 7, 2018. [Online]. Available: http://ethereum.org/ethereum.html

[84] S. Wang, L. Ouyang, Y. Yuan, X. Ni, X. Han, and F. Wang, "Blockchain-enabled smart contracts: Architecture, applications, and future trends," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 1, pp. 2266–2277, Nov. 2019.

[85] G. Wood, "Ethereum: A secure decentralised generalised transaction ledger," vol. 151, Zug, Switzerland, Ethereum Project, Yellow Paper, 2014.

[86] *Litecoin Cash LTCH*. Accessed: Sep. 7, 2018. [Online]. Available: https://litecoin-cash.io/Whitepaper.pdf

[87] Z. Zheng, S. Xie, H.-N. Dai, X. Chen, and H. Wang, "Blockchain challenges and opportunities: A survey," *Int. J. Web Grid Services*, vol. 14, no. 4, pp. 352–375, 2018.

[88] O. Dib, K.-L. Brousmiche, A. Durand, E. Thea, and E. B. Hamida, "Consortium blockchains: Overview, applications and challenges," *Int. J. Adv. Telecommun.*, vol. 11, nos. 1–2, pp. 1–6, 2018.

[89] J. Wu, S. Guo, H. Huang, W. Liu, and Y. Xiang, "Information and communications technologies for sustainable development goals: State-of-the-art, needs and perspectives," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2389–2406, 3rd Quart., 2018.

[90] R. Atat, L. Liu, H. Chen, J. Wu, H. Li, and Y. Yi, "Enabling cyber-physical communication in 5G cellular networks: Challenges, spatial spectrum sensing, and cyber-security," *IET Cyber Phys. Syst. Theory Appl.*, vol. 2, no. 1, pp. 49–54, 2017.

[91] K. Christidis and M. Devetsikiotis, "Blockchains and smart contracts for the Internet of Things," *IEEE Access*, vol. 4, pp. 2292–2303, 2016.

[92] M. Salimitari and M. Chatterjee, "An overview of blockchain and consensus protocols for IoT networks," 2018. [Online]. Available: arXiv:1809.05613.

[93] Y. Liu, K. Wang, Y. Lin, and W. Xu, "LightChain: A lightweight blockchain system for Industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 15, no. 6, pp. 3571–3581, Jun. 2019.

[94] E. Langberg and S. Chen. (2017). *Cblockchains in Mobile Networks*. [Online]. Available: https://e.huawei.com/hk/publications/global/ict$$ insights/201703141505/

[95] S. He, C. Xing, and L.-J. Zhang, "A business-oriented schema for blockchain network operation," in *Proc. Int. Conf. Blockchain*, 2018, pp. 277–284.

[96] Z. Li, J. Kang, R. Yu, D. Ye, Q. Deng, and Y. Zhang, "Consortium blockchain for secure energy trading in Industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 14, no. 8, pp. 3690–3700, Aug. 2018.

[97] X. Zhang and X. Chen, "Data security sharing and storage based on a consortium blockchain in a vehicular adhoc network," *IEEE Access*, vol. 7, pp. 58241–58254, 2019.

[98] M. Cebe, E. Erdin, K. Akkaya, H. Aksu, and S. Uluagac, "Block4Forensic: An integrated lightweight blockchain framework for forensics applications of connected vehicles," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 50–57, Oct. 2018.

[99] Y. Yang, L. Chou, C. Tseng, F. Tseng, and C. Liu, "Blockchain-based traffic event validation and trust verification for VANETs," *IEEE Access*, vol. 7, pp. 30868–30877, 2019.

[100] J. Kang, Z. Xiong, D. Niyato, D. Ye, D. I. Kim, and J. Zhao, "Toward secure blockchain-enabled Internet of Vehicles: Optimizing consensus management using reputation and contract theory," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2906–2920, Mar. 2019.

[101] J. Kang *et al.*, "Blockchain for secure and efficient data sharing in vehicular edge computing and networks," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4660–4670, Jun. 2019.

[102] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella, "On multi-access edge computing: A survey of the emerging 5G network edge cloud architecture and orchestration," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1657–1681, 3rd Quart., 2017.

[103] Z. Xiong, Y. Zhang, D. Niyato, P. Wang, and Z. Han, "When mobile blockchain meets edge computing," *IEEE Commun. Mag.*, vol. 56, no. 8, pp. 33–39, Aug. 2018.

[104] S. Seng, X. Li, C. Luo, H. Ji, and H. Zhang, "A D2D-assisted MEC computation offloading in the blockchain-based framework for UDNs," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[105] Y. Jiao, P. Wang, D. Niyato, and Z. Xiong, "Social welfare maximization auction in edge computing resource allocation for mobile blockchain," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2018, pp. 1–6.

[106] M. Liu, F. R. Yu, Y. Teng, V. C. M. Leung, and M. Song, "Computation offloading and content caching in wireless blockchain networks with mobile edge computing," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 11008–11021, Nov. 2018.

[107] V. Sharma, I. You, F. Palmieri, D. N. K. Jayakody, and J. Li, "Secure and energy-efficient handover in fog networks using blockchain-based DMM," *IEEE Commun. Mag.*, vol. 56, no. 5, pp. 22–31, May 2018.

[108] Z. Chen, S. Chen, H. Xu, and B. Hu, "A security authentication scheme of 5G ultra-dense network based on block chain," *IEEE Access*, vol. 6, pp. 55372–55379, 2018.

[109] K. Kotobi and S. G. Bilen, "Secure blockchains for dynamic spectrum access: A decentralized database in moving cognitive radio networks enhances security and user access," *IEEE Veh. Technol. Mag.*, vol. 13, no. 1, pp. 32–39, Mar. 2018.

[110] X. Ling, J. Wang, T. Bouchoucha, B. C. Levy, and Z. Ding, "Blockchain radio access network (B-RAN): Towards decentralized secure radio access paradigm," *IEEE Access*, vol. 7, pp. 9714–9723, 2019.

[111] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM J. Res. Develop.*, vol. 44, nos. 1–2, pp. 210–229, 1959.

[112] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intell. Syst. Appl.*, vol. 13, no. 4, pp. 18–28, Apr. 1998.

[113] S. Fine, Y. Singer, and N. Tishby, "The hierarchical hidden Markov model: Analysis and applications," *Mach. Learn.*, vol. 32, no. 1, pp. 41–62, 1998.

[114] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics Intell. Lab. Syst.*, vol. 2, nos. 1–3, pp. 37–52, 1987.

[115] M. Yilan and M. K. Özdemir, "A simple approach to traffic density estimation by using kernel density estimation," in *Proc. Signal Process. Commun. Appl. Conf.*, May 2015, pp. 1865–1868.

[116] A. Tabibiazar and O. Basir, "Kernel-based optimization for traffic density estimation in ITS," in *Proc. IEEE VTC-Fall*, Sep. 2011, pp. 1–5.

[117] J. Ran, Y. Chen, and S. Li, "Three-dimensional convolutional neural network based traffic classification for wireless communications," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Nov. 2018, pp. 624–627.

[118] L. Xiao and L. Cheng, "State classification algorithm for bus based on hierarchical support vector machine," in *Proc. Int. Symp. Comput. Intell. Design*, vol. 2, Dec. 2015, pp. 649–652.

[119] T. Auld, A. W. Moore, and S. F. Gull, "Bayesian neural networks for Internet traffic classification," *IEEE Trans. Neural Netw.*, vol. 18, no. 1, pp. 223–239, Jan. 2007.

[120] R. Sun, B. Yang, L. Peng, Z. Chen, L. Zhang, and S. Jing, "Traffic classification using probabilistic neural networks," in *Proc. Int. Conf. Nat. Comput.*, vol. 4, Aug. 2010, pp. 1914–1919.

[121] N. Jing, M. Yang, S. Cheng, Q. Dong, and H. Xiong, "An efficient SVM-based method for multi-class network traffic classification," in *Proc. IEEE Int. Perform. Comput. Commun. Conf.*, Nov. 2011, pp. 1–8.

[122] Z. Chen, N. Pears, M. Freeman, and J. Austin, "Road vehicle classification using support vector machines," in *Proc. IEEE Int. Conf. Intell. Comput. Intell. Syst.*, vol. 4, Nov. 2009, pp. 214–218.

[123] F. Noorbehbahani and S. Mansoori, "A new semi-supervised method for network traffic classification based on x-means clustering and label propagation," in *Proc. ICCKE*, Oct. 2018, pp. 120–125.

[124] H. Jie, Y. Yuexiang, Q. Yong, and T. Chuan, "Accurate classification of P2P traffic by clustering flows," *China Commun.*, vol. 10, no. 11, pp. 42–51, Nov. 2013.

[125] K. Shim, Y. Goo, M. Lee, H. Hasanova, and M. Kim, "The method of clustering network traffic classifications for extracting payload signature by function," in *Proc. Inf. Commun. Technol. Converg.*, Oct. 2018, pp. 1335–1337.

[126] L. Bin and T. Hao, "P2P traffic classification using semi-supervised learning," in *Proc. Int. Conf. Artif. Intell. Comput. Intell.*, vol. 1, Oct. 2010, pp. 408–412.

[127] X. Bian, "PSO optimized semi-supervised network traffic classification strategy," in *Proc. Int. Conf. Intell. Transp. Big Data Smart City*, Jan. 2018, pp. 179–182.

[128] Z. Wang, Y. Dong, S. Mao, and X. Wang, "Internet multimedia traffic classification from QoS perspective using semi-supervised dictionary learning models," *China Commun.*, vol. 14, no. 10, pp. 202–218, Oct. 2017.

[129] Y. Zeng, H. Gu, W. Wei, and Y. Guo, "Deep-full-range: A deep learning based network encrypted traffic classification and intrusion detection framework," *IEEE Access*, vol. 7, pp. 45182–45190, 2019.

[130] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: Deep belief networks with multitask learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2191–2201, Oct. 2014.

[131] L. Yanjun, L. Xiaobo, and Y. Osamu, "Traffic engineering framework with machine learning based meta-layer in software-defined networks," in *Proc. IEEE Int. Conf. Netw. Infrastruct. Digit. Content*, Sep. 2014, pp. 121–125.

[132] L. Mekinda and L. Muscariello, "Supervised machine learning-based routing for named data networking," in *Proc. IEEE GLOBECOM*, Dec. 2016, pp. 1–6.

[133] A. Azzouni, R. Boutaba, and G. Pujolle, "Neuroute: Predictive dynamic routing for software-defined networks," in *Proc. Conf. Netw. Service Manag.*, Nov. 2017, pp. 1–6.

[134] Z. Mammeri, "Reinforcement learning based routing in networks: Review and classification of approaches," *IEEE Access*, vol. 7, pp. 55916–55950, 2019.

[135] F. Li, X. Song, H. Chen, X. Li, and Y. Wang, "Hierarchical routing for vehicular ad hoc networks via reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1852–1865, Feb. 2019.

[136] S.-C. Lin, I. F. Akyildiz, P. Wang, and M. Luo, "Qos-aware adaptive routing in multi-layer hierarchical software defined networks: A reinforcement learning approach," in *Proc. IEEE Int. Conf. Services Comput.*, Jun. 2016, pp. 25–33.

[137] C. Yu, J. Lan, Z. Guo, and Y. Hu, "DROM: Optimizing the routing in software-defined networks with deep reinforcement learning," *IEEE Access*, vol. 6, pp. 64533–64539, 2018.

[138] X. Li, Y. Liu, H. Ji, H. Zhang, and V. C. M. Leung, "Optimizing resources allocation for fog computing-based Internet of Things networks," *IEEE Access*, vol. 7, pp. 64907–64922, 2019.

[139] Z. Zhang, R. Wang, F. R. Yu, F. Fu, and Q. Yan, "QoS aware transcoding for live streaming in edge-clouds aided HetNets: An enhanced actor–critic approach," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 11295–11308, Nov. 2019.

[140] Y. Wei, F. R. Yu, M. Song, and Z. Han, "Joint optimization of caching, computing, and radio resources for fog-enabled IoT using natural actor–critic deep reinforcement learning," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2061–2073, Apr. 2019.

[141] J. Chen, S. Chen, Q. Wang, B. Cao, G. Feng, and J. Hu, "iRAF: A deep reinforcement learning approach for collaborative mobile edge computing IoT networks," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 7011–7024, Aug. 2019.

[142] X. Fu, F. R. Yu, J. Wang, Q. Qi, and J. Liao, "Dynamic service function chain embedding for NFV-enabled IoT: A deep reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 507–519, Jan. 2020.

[143] R. Li *et al.*, "Deep reinforcement learning for resource management in network slicing," *IEEE Access*, vol. 6, pp. 74429–74441, 2018.

[144] A. S. Sadiq *et al.*, "An efficient IDS using hybrid magnetic swarm optimization in WANETs," *IEEE Access*, vol. 6, pp. 29041–29053, 2018.

[145] M. H. Ali, B. A. D. A. Mohammed, A. Ismail, and M. F. Zolkipli, "A new intrusion detection system based on fast learning network and particle swarm optimization," *IEEE Access*, vol. 6, pp. 20255–20261, 2018.

[146] W. Hu, J. Gao, Y. Wang, O. Wu, and S. Maybank, "Online Adaboost-based parameterized methods for dynamic distributed network intrusion detection," *IEEE Trans. Cybern.*, vol. 44, no. 1, pp. 66–82, Jan. 2014.

[147] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A deep learning approach for network intrusion detection system," in *Proc. EAI Int. Conf. BIONETICS*, 2016, pp. 21–26.

[148] K. Hamedani, L. Liu, R. Atat, J. Wu, and Y. Yi, "Reservoir computing meets smart grids: Attack detection using delayed feedback networks," *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 734–743, Feb. 2018.

[149] *Ocean Protocol: A Decentralized Substrate for AI Data and Services Technical Whitepaper*. Accessed: Apr. 15, 2019. [Online]. Available: https://oceanprotocol.com/tech-whitepaper.pdf

[150] R. Doku and D. Rawat, "Pledge: A private ledger based decentralized data sharing framework," in *Proc. SpringSi*, Apr. 2019, pp. 1–11.

[151] P. Mamoshina *et al.*, "Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare," *Oncotarget*, vol. 9, no. 5, p. 5665, 2018.

[152] T.-T. Kuo and L. Ohno-Machado, "ModelChain: Decentralized privacy-preserving healthcare predictive modeling framework on private blockchain networks," 2018. [Online]. Available: arXiv:1802.01746.

[153] Y. Freund *et al.*, "Experiments with a new boosting algorithm," in *Proc. Int. Conf. Mach. Learn.*, vol. 96, 1996, pp. 148–156.

[154] E. C. Ferrer, O. Rudovic, T. Hardjono, and A. Pentland, "RoboChain: A secure data-sharing framework for human–robot interaction," 2018. [Online]. Available: arXiv:1802.04480.

[155] K. Rantos, G. Drosatos, K. Demertzis, C. Ilioudis, and A. Papanikolaou, "Blockchain-based consents management for personal data processing in the IoT ecosystem," in *Proc. SECRYPT ICETE*, 2018, pp. 572–577.

[156] C. Research. (Jan. 2018). *Cognitive Scale Briefing Note*. Accessed: Sep. 7, 2018. [Online]. Available: https://www.cognitivescale.com/wp-content/uploads/2018/01/cognilytica-briefing-note-cognitivescale.pdf

[157] J. Zhang, N. Xue, and X. Huang, "A secure system for pervasive social network-based healthcare," *IEEE Access*, vol. 4, pp. 9239–9250, 2016.

[158] N. Fabiano, "Internet of Things and blockchain: Legal issues and privacy. The challenge for a privacy standard," in *Proc. IEEE iThings IEEE GreenCom IEEE CPSCom IEEE SmartData*, Jun. 2017, pp. 727–734.

[159] S. Cha, J. Chen, C. Su, and K. Yeh, "A blockchain connected gateway for BLE-based devices in the Internet of Things," *IEEE Access*, vol. 6, pp. 24639–24649, 2018.

[160] X. Liang, S. Shetty, D. Tosh, C. Kamhoua, K. Kwiat, and L. Njilla, "ProvChain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability," in *Proc. IEEE/ACM CCGRID*, May 2017, pp. 468–477.

[161] Y. Zhang, S. Wu, B. Jin, and J. Du, "A blockchain-based process provenance for cloud forensics," in *Proc. IEEE ICCC*, Dec. 2017, pp. 2470–2473.

[162] L. Li *et al.*, "CreditCoin: A privacy-preserving blockchain-based incentive announcement network for communications of smart vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 7, pp. 2204–2220, Jul. 2018.

[163] P. Dunphy and F. A. P. Petitcolas, "A first look at identity management schemes on the blockchain," *IEEE Security Privacy*, vol. 16, no. 4, pp. 20–29, Jul. 2018.

[164] N. Kaaniche and M. Laurent, "A blockchain-based data usage auditing architecture with enhanced privacy and availability," in *Proc. IEEE Netw. Comput. Appl.*, 2017, pp. 1–5.

[165] E. Androulaki, S. Cocco, and C. Ferris. (2018). *Private and Confidential Transactions With Hyperledger Fabric*. [Online]. Available: https://hyperledger-fabric.readthedocs.io/en/release-2.0/private_data_tutorial.html

[166] M. Shen, X. Tang, L. Zhu, X. Du, and M. Guizani, "Privacy-preserving support vector machine training over blockchain-based encrypted IoT data in smart cities," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7702–7712, Oct. 2019.

[167] L. Lyu, J. Yu, K. Nandakumar, Y. Li, X. Ma, and J. Jin, "Towards fair and decentralized privacy-preserving deep learning with blockchain," 2019. [Online]. Available: arXiv:1906.01167.

[168] J. Weng, J. Weng, J. Zhang, M. Li, Y. Zhang, and W. Luo, "Deepchain: Auditable and privacy-preserving deep learning with blockchain-based incentive," *IEEE Trans. Depend. Secure Comput.*, early access, doi: 10.1109/TDSC.2019.2952332.

[169] X. Chen, J. Ji, C. Luo, W. Liao, and P. Li, "When machine learning meets blockchain: A decentralized, privacy-preserving and secure design," in *Proc. IEEE Big Data*, Dec. 2018, pp. 1178–1187.

[170] K. Singla, J. Bose, and S. Katariya, "Machine learning for secure device personalization using blockchain," in *Proc. Int. Conf. Adv. Comput. Commun. Informat.*, Sep. 2018, pp. 67–73.

[171] A. Outchakoucht, E. Hamza, and J. P. Leroy, "Dynamic access control policy based on blockchain and machine learning for the Internet of Things," *Int. J. Adv. Comput. Sci. Appl*, vol. 8, no. 7, pp. 417–424, 2017.

[172] D. Lin and Y. Tang, "Blockchain consensus based user access strategies in D2D networks for data-intensive applications," *IEEE Access*, vol. 6, pp. 72683–72690, 2018.

[173] *Decentralized Machine Learning White Paper*. Accessed: Oct. 10, 2018. [Online]. Available: https://decentralizedml.com/DML_whitepaper_31Dec_17.pdf

[174] Y. Miiller, "Decentralized artificial intelligence," *Decentralised AI*, pp. 3–13, 1990.

[175] C. Song, T. Ristenpart, and V. Shmatikov, "Artificial intelligence and blockchain: 3 major benefits of combining these two mega-trends," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2017, pp. 587–601.

[176] L. Melis, C. Song, E. De Cristofaro, and V. Shmatikov, "Inference attacks against collaborative learning," 2018. [Online]. Available: arXiv:1805.04049.

[177] B. Hitaj, G. Ateniese, and F. Perez-Cruz, "Deep models under the GAN: Information leakage from collaborative deep learning," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2017, pp. 603–618.

[178] T. Orekondy, S. J. Oh, B. Schiele, and M. Fritz, "Understanding and controlling user linkability in decentralized learning," 2018. [Online]. Available: arXiv:1805.05838.

[179] N. K. Bore *et al.*, "Promoting distributed trust in machine learning and computational simulation," in *Proc. IEEE Int. Conf. Blockchain Cryptocurrency*, 2019, pp. 311–319.

[180] S. Vyas, M. Gupta, and R. Yadav, "Converging blockchain and machine learning for healthcare," in *Proc. AICAI*, Feb. 2019, pp. 709–711.

[181] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Blockchain and federated learning for privacy-preserved data sharing in industrial IoT," *IEEE Trans. Ind. Informat.*, early access, doi: 10.1109/TII.2019.2942190.

[182] T. A. Butt, R. Iqbal, K. Salah, M. Aloqaily, and Y. Jararweh, "Privacy management in social Internet of Vehicles: Review, challenges and blockchain based solutions," *IEEE Access*, vol. 7, pp. 79694–79713, 2019.

[183] B. Goertzel, S. Giacomelli, D. Hanson, C. Pennachin, and M. Argentieri. (2017). *SingularityNet: A Decentralized, Open Market and Inter-Network for AIs*. Accessed: May 27, 2019. [Online]. Available: https://public.singularitynet.io/whitepaper.pdf

[184] G. A. Montes and B. Goertzel, "Distributed, decentralized, and democratized artificial intelligence," *Technol. Forecast. Soc. Change*, vol. 141, pp. 354–358, Apr. 2019.

[185] *Introducing Open Mined: Decentralised AI*. Accessed: Sep. 7, 2018. [Online]. Available: https://github.com/OpenMined

[186] J. Eisses, L. Verspeek, C. Dawe, and S. Dijkstra. (2018). *Effect Network: Decentralized Network for Artificial Intelligence*. Accessed: May 27, 2019. [Online]. Available: https://cryptorating.eu/whitepapers/Effect/effect_whitepaper.pdf

[187] A. B. Kurtulmus and K. Daniel, "Trustless machine learning contracts; evaluating and exchanging machine learning models on the Ethereum blockchain," 2018. [Online]. Available: arXiv:1802.10185.

[188] Y. Lu, Q. Tang, and G. Wang, "On enabling machine learning tasks atop public blockchains: A crowdsourcing approach," in *Proc. IEEE Int. Conf. Data Min. Workshop*, 2018, pp. 81–88.

[189] G. J. Mendis, M. Sabounchi, J. Wei, and R. Roche, "Blockchain as a service: An autonomous, privacy preserving, decentralized architecture for deep learning," 2018. [Online]. Available: arXiv:1807.02515.

[190] F. Corea, *An Introduction to Data: Everything You Need to Know About AI, Big Data and Data Science*, vol. 50. Cham, Switzerland: Springer, 2018.

[191] H. B. McMahan *et al.*, "Communication-efficient learning of deep networks from decentralized data," 2016. [Online]. Available: arXiv:1602.05629.

[192] S. Wang, Y. Yuan, X. Wang, J. Li, R. Qin, and F.-Y. Wang, "An overview of smart contract: Architecture, applications, and future trends," in *Proc. IEEE Intell. Veh. Symp. (IV)*, 2018, pp. 108–113.

[193] N. Alexopoulos, J. Daubert, M. Mühlhäuser, and S. M. Habib, "Beyond the hype: On using blockchains in trust management for authentication," in *Proc. IEEE Trustcom/BigDataSE/ICESS*, Aug. 2017, pp. 546–553.

[194] B. Yu, J. Wright, S. Nepal, L. Zhu, J. Liu, and R. Ranjan, "IoTChain: Establishing trust in the Internet of Things ecosystem using blockchain," *IEEE Cloud Comput.*, vol. 5, no. 4, pp. 12–23, Jul./Aug. 2018.

[195] M. Singh and S. Kim, "Trust bit: Reward-based intelligent vehicle commination using blockchain paper," in *Proc. IEEE WF-IoT*, Feb. 2018, pp. 62–67.

[196] S. Kirkman, "A data movement policy framework for improving trust in the cloud using smart contracts and blockchains," in *Proc. IEEE Int. Conf. Cloud Eng.*, Apr. 2018, pp. 270–273.

[197] T. Wang, "A unified analytical framework for trustable machine learning and automation running with blockchain," in *Proc. IEEE Int. Conf. Big Data*, Dec. 2018, pp. 4974–4983.

[198] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," *ACM SIGMOD Rec.*, vol. 22, no. 2, pp. 207–216, 1993.

[199] N. K. Bore *et al.*, "Promoting distributed trust in machine learning and computational simulation via a blockchain network," 2018. [Online]. Available: arXiv:1810.11126.

[200] J. Yang, S. He, Y. Xu, L. Chen, and J. Ren, "A trusted routing scheme using blockchain and reinforcement learning for wireless sensor networks," *Sensors*, vol. 19, no. 4, p. 970, 2019.

[201] V. Arya, S. Sen, and P. Kodeswaran, "Blockchain enabled trustless API marketplace," 2018. [Online]. Available: arXiv:1812.02154.

[202] J. Kang, Z. Xiong, D. Niyato, S. Xie, and J. Zhang, "Incentive mechanism for reliable federated learning: A joint optimization approach to combining reputation and contract theory," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10700–10714, Dec. 2019.

[203] F. Bravo-Marquez, S. Reeves, and M. Ugarte, "Proof-of-learning: A blockchain consensus mechanism based on machine learning competitions," in *Proc. IEEE DAPPCON*, Apr. 2019, pp. 119–124.

[204] C. Chenli, B. Li, Y. Shi, and T. Jung, "Energy-recycling blockchain with proof-of-deep-learning," 2019. [Online]. Available: arXiv:1902.03912.

[205] A. Baldominos and Y. Saez, "Coin.ai: A proof-of-useful-work scheme for blockchain-based distributed deep learning," 2019. [Online]. Available: arXiv:1903.09800.

[206] Y. Zhang, H. Wang, T. Zheng, and Q. Yang, "Energy-efficient transmission design in non-orthogonal multiple access," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2852–2857, Mar. 2017.

[207] P. T'Serclaes, *Blockchain Could Be the Missing Link in the Renewable Energy Revolution*, vol. 21, World Econ. Forum, Cologny, Switzerland, 2017.

[208] J. Chen, K. Duan, R. Zhang, L. Zeng, and W. Wang, "An AI based super nodes selection algorithm in blockchain networks," 2018. [Online]. Available: arXiv:1808.00216.

[209] N. C. Luong, Z. Xiong, P. Wang, and D. Niyato, "Optimal auction for edge computing resource management in mobile blockchain networks: A deep learning approach," in *Proc. IEEE Int. Conf. Commun.*, May 2018, pp. 1–6.

[210] C. Xu, K. Wang, and M. Guo, "Intelligent resource management in blockchain-based cloud datacenters," *IEEE Cloud Comput.*, vol. 4, no. 6, pp. 50–59, Nov. 2017.

[211] Z. Zhou, B. Wang, Y. Guo, and Y. Zhang, "Blockchain and computational intelligence inspired incentive-compatible demand response in Internet of Electric Vehicles," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 3, no. 3, pp. 205–216, Jun. 2019.

[212] M. A. Ferrag and L. Maglaras, "DeepCoin: A novel deep learning and blockchain-based energy exchange framework for smart grids," *IEEE Trans. Eng. Manag.*, early access, doi: 10.1109/TEM.2019.2922936.

[213] *Verv VLUX Whitepaper—The Evolution of Energy*. Accessed: Sep. 7, 2018. [Online]. Available: https://vlux.io/media/VLUX_Whitepaper.pdf

[214] J. Xie, F. R. Yu, T. Huang, R. Xie, J. Liu, and Y. Liu, "A survey on the scalability of blockchain systems," *IEEE Netw.*, vol. 33, no. 5, pp. 166–173, Sep. 2019.

[215] L. S. Sankar, M. Sindhu, and M. Sethumadhavan, "Survey of consensus protocols on blockchain applications," in *Proc. IEEE Int. Conf. Adv. Comput. Commun. Syst.*, 2017, pp. 1–5.

[216] *Automating Hiring With Blockchain and Artificial Intelligence Technologies*. Accessed: Sep. 28, 2018. [Online]. Available: https://talentsnap.co/whitepaper.pdf

[217] *Fetch: Technical Introduction: A Decentralised Digital World for the Future Economy*. Accessed: Mar. 15, 2019. [Online]. Available: https://fetch.ai/uploads/technical-introduction.pdf

[218] C. Jentzsch. (2016). *Decentralized Autonomous Organization to Automate Governance White Paper*. [Online]. Available: https://download.slock.it/public/DAO/WhitePaper.pdf

[219] *The Inside Story of Mt. Gox, Bitcoin's $460 Million Disaster*. Accessed: Sep. 7, 2018. [Online]. Available: https://www.wired.com/2014/03/bitcoin-exchange

[220] S. Higgins, *The Bitfinex Bitcoin Hack: What We Know (and Don't Know)*, Coindesk, New York, NY, USA, Aug. 2016.
[221] M. U. Wasim, A. A. Z. A. Ibrahim, P. Bouvry, and T. Limba, "Law as a service (LaaS): Enabling legal protection over a blockchain network," in *Proc. HONET-ICT*, Oct. 2017, pp. 110–114.
[222] H. S. Yin and R. Vatrapu, "A first estimation of the proportion of cyber-criminal entities in the Bitcoin ecosystem using supervised machine learning," in *Proc. IEEE Big Data*, Dec. 2017, pp. 3690–3699.
[223] S. Raje, S. Vaderia, N. Wilson, and R. Panigrahi, "Decentralised firewall for malware detection," in *Proc. IEEE Int. Conf. Auton. Comput.*, Dec. 2017, pp. 1–5.
[224] *Secure and Grow Your Blockchain Business*. Accessed: May 1, 2019. [Online]. Available: https://www.anchain.ai/
[225] S. Pradhan, K. Hacioglu, W. Ward, J. H. Martin, and D. Jurafsky, "Semantic role parsing: Adding semantic structure to unstructured text," in *Proc. IEEE Int. Conf. Data Min.*, Nov. 2003, pp. 629–632.
[226] Y. Hsu and H. Kao, "Curatable named-entity recognition using semantic relations," *IEEE/ACM Trans. Comput. Biol. Bioinformat.*, vol. 12, no. 4, pp. 785–792, Jul. 2015.
[227] Z. Shi, "The design and implementation of domain-specific text summarization system based on co-reference resolution algorithm," in *Proc. Int. Conf. Fuzzy Syst. Knowl. Disc.*, vol. 5, Aug. 2010, pp. 2390–2394.
[228] I. Dagan and A. Itai, "Word sense disambiguation using a second language monolingual corpus," *Comput. Linguist.*, vol. 20, no. 4, pp. 563–596, 1994.
[229] M. Mylrea, "AI enabled blockchain smart contracts: Cyber resilient energy infrastructure and IoT," in *Proc. AAAI Spring Symposia*, 2018. [Online]. Available: https://aaai.org/ocs/index.php/SSS/SSS18/paper/view/17593
[230] S. S. Arumugam *et al.*, "IoT enabled smart logistics using smart contracts," in *Proc. IEEE LISS*, 2018, pp. 1–6.
[231] *Matrix Technical Whitepaper*. Accessed: Sep. 7, 2018. [Online]. Available: https://www.matrix.io/html/MATRIXTechnicalWhitePaper.pdf
[232] M. Wong, H. Haapio, S. Deckers, and S. Dhir, "Computational contract collaboration and construction," in *Proc. IRIS*, 2015, pp. 505–512.
[233] P. Bailis, A. Narayanan, A. Miller, and S. Han, "Research for practice: Cryptocurrencies, blockchains, and smart contracts; hardware for deep learning," *Commun. ACM*, vol. 60, no. 5, pp. 48–51, 2017.
[234] F. Idelberger, G. Governatori, R. Riveret, and G. Sartor, "Evaluation of logic-based smart contracts for blockchain systems," in *Proc. Int. Symp. Rules Rule Markup Lang. Semantic Web*, 2016, pp. 167–183.
[235] L. Luu, D.-H. Chu, H. Olickel, P. Saxena, and A. Hobor, "Making smart contracts smarter," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2016, pp. 254–269.
[236] T. H.-D. Huang, "Hunting the Ethereum smart contract: Color-inspired inspection of potential attacks," 2018. [Online]. Available: arXiv:1807.01868.
[237] W. J.-W. Tann, X. J. Han, S. S. Gupta, and Y.-S. Ong, "Towards safer smart contracts: A sequence learning approach to detecting security threats," 2018. [Online]. Available: arXiv:1811.06632.
[238] L. Zhu, H. Dong, M. Shen, and K. Gai, "An incentive mechanism using Shapley value for blockchain-based medical data sharing," in *Proc. IEEE BigDataSecurity HPSC IDS*, May 2019, pp. 113–118.
[239] S. AVRP and P. K. Baruah, "Blended learning-assimilating authentic data into deep learning models," in *Proc. IEEE HiPCW*, Dec. 2018, pp. 75–80.
[240] C. H. Liu, Q. Lin, and S. Wen, "Blockchain-enabled data collection and sharing for industrial IoT with deep reinforcement learning," *IEEE Trans. Ind. Informat.*, vol. 15, no. 6, pp. 3516–3526, Jun. 2019.
[241] D. Siegel. (2016). *Understanding the DAO Attack*. [Online]. Available: http://www.coindesk.com/understanding-dao-hack-journalists
[242] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A deep learning approach to network intrusion detection," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 2, no. 1, pp. 41–50, Feb. 2018.
[243] H. Yao, T. Mai, J. Wang, Z. Ji, C. Jiang, and Y. Qian, "Resource trading in blockchain-based Industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 15, no. 6, pp. 3602–3609, Jun. 2019.
[244] C. Qiu, H. Yao, C. Jiang, S. Guo, and F. Xu, "Cloud computing assisted blockchain-enabled Internet of Things," *IEEE Trans. Cloud Comput.*, early access, doi: 10.1109/TCC.2019.2930259.
[245] C. Qiu, F. R. Yu, H. Yao, C. Jiang, F. Xu, and C. Zhao, "Blockchain-based software-defined Industrial Internet of Things: A dueling deep Q-learning approach," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4627–4639, Jun. 2019.
[246] Cisco. *Internet of Things at a Glance—Cisco*. Accessed: Feb. 18, 2020. [Online]. Available: https://www.cisco.com/c/dam/en/us/products/collateral/se/internet-of-things/at-a-glance-c45-731471.pdf
[247] C. MacGillivray *et al.*, "IDC future scape: Worldwide IoTs 2018 predictions," in *Proc. IDC Web Conf.*, 2017, pp. 29–32.
[248] S. Popov. (2017). *The Tangle*. Accessed: Oct. 7, 2018. [Online]. Available: http://iotatoken.com/IOTA_Whitepaper.pdf
[249] T. Mahmood and U. Afzal, "Security analytics: Big data analytics for cybersecurity: A review of trends, techniques and tools," in *Proc. NCIA*, Dec. 2013, pp. 129–134.
[250] Y. He, F. R. Yu, N. Zhao, V. C. M. Leung, and H. Yin, "Software-defined networks with mobile edge computing and caching for smart cities: A big data deep reinforcement learning approach," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 31–37, Dec. 2017.
[251] J. Wu, S. Guo, J. Li, and D. Zeng, "Big data meet green challenges: Big data toward green applications," *IEEE Syst. J.*, vol. 10, no. 3, pp. 888–900, Sep. 2016.
[252] M. G. Kibria, K. Nguyen, G. P. Villardi, O. Zhao, K. Ishizu, and F. Kojima, "Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks," *IEEE Access*, vol. 6, pp. 32328–32338, 2018.
[253] W.-J. Wang, M.-C. Su, S.-H. Lee, C.-S. Hung, and C.-C. Chen, "A guideline to determine the training sample size when applying big data mining methods in clinical decision making," in *Proc. IEEE Int. Conf. Appl. Syst. Innov.*, Apr. 2018, pp. 678–681.
[254] L. Zhu, F. R. Yu, Y. Wang, B. Ning, and T. Tang, "Big data analytics in intelligent transportation systems: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 1, pp. 383–398, Jan. 2019.
[255] R. Atat, L. Liu, J. Wu, G. Li, C. Ye, and Y. Yang, "Big data meet cyber-physical systems: A panoramic survey," *IEEE Access*, vol. 6, pp. 73603–73636, 2018.
[256] S. Wilkinson, T. Boshevski, J. Brandoff, and V. Buterin. (2014). *STORJ a Peer-to-Peer Cloud Storage Network*. Accessed: May 28, 2019. [Online]. Available: https://storj.io/storjv3.pdf
[257] *Datablockchain*. Accessed: Sep. 7, 2018. [Online]. Available: https://www.datablockchain.io/pdf/whitepaper.pdf
[258] Y. Liu, F. R. Yu, X. Li, H. Ji, and V. C. M. Leung, "Distributed resource allocation and computation offloading in fog and cloud networks with non-orthogonal multiple access," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 12137–12151, Dec. 2018.
[259] P. S. Chandakkar, Y. Li, P. L. K. Ding, and B. Li, "Strategies for re-training a pruned neural network in an edge computing paradigm," in *Proc. IEEE EDGE*, Jun. 2017, pp. 244–247.
[260] Y. Huang, X. Ma, X. Fan, J. Liu, and W. Gong, "When deep learning meets edge computing," in *Proc. IEEE ICNP*, Oct. 2017, pp. 1–2.
[261] A. Bodas, B. Upadhyay, C. Nadiger, and S. Abdelhak, "Reinforcement learning for game personalization on edge devices," in *Proc. ICICT*, Mar. 2018, pp. 119–122.
[262] Z. Xiong, Y. Zhang, D. Niyato, P. Wang, and Z. Han, "When mobile blockchain meets edge computing," *IEEE Comm. Mag.*, vol. 56, no. 8, pp. 33–39, Aug. 2018.
[263] Y. Liu, F. R. Yu, X. Li, H. Ji, and V. C. M. Leung, "Hybrid computation offloading in fog and cloud networks with non-orthogonal multiple access," in *Proc. IEEE INFOCOM Workshops*, Apr. 2018, pp. 154–159.

Yiming Liu received the B.E. degree in communication engineering from Shanghai University in 2014, and the Ph.D. degree in information and communication engineering from the Beijing University of Posts and Telecommunications (BUPT) in 2019.

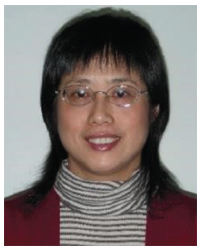
From 2017 to 2018, she was a visiting Ph.D. student with the University of British Columbia in 2018. She is currently a Postdoctoral Researcher with the School of Information and Communication Engineering, BUPT. Her current research interests include ultra-dense networks, non-orthogonal multiple access, resource management, mobile edge computing, blockchain, and distributed ledger technology.

F. Richard Yu (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of British Columbia in 2003.

From 2002 to 2006, he was with Ericsson, Lund, Sweden, and a start-up in California, USA. He joined Carleton University in 2007, where he is currently a Professor. His research interests include wireless cyber-physical systems, connected/autonomous vehicles, security, distributed ledger technology, and deep learning. He received the IEEE Outstanding Service Award in 2016, the IEEE Outstanding Leadership Award in 2013, the Carleton Research Achievement Award in 2012, the Ontario Early Researcher Award (formerly Premiers Research Excellence Award) in 2011, the Excellent Contribution Award at IEEE/IFIP TrustCom 2010, the Leadership Opportunity Fund Award from Canada Foundation of Innovation in 2009, and the Best Paper Awards at IEEE ICNC 2018, VTC 2017 Spring, ICC 2014, Globecom 2012, IEEE/IFIP TrustCom 2009, and Int'l Conference on Networking 2005. He serves on the editorial boards of several journals, including the Co-Editor-in-Chief for *Ad Hoc and Sensor Wireless Networks*, the Lead Series Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, and IEEE COMMUNICATIONS SURVEYS & TUTORIALS. He has served as the Technical Program Committee Co-Chair of numerous conferences. He is a Registered Professional Engineer with the province of Ontario, Canada, a fellow of the Institution of Engineering and Technology. He is a Distinguished Lecturer, the Vice President (Membership), and an Elected Member of the Board of Governors of IEEE Vehicular Technology Society.

Xi Li received the B.E. and Ph.D. degrees in communication and information system from the Beijing University of Posts and Telecommunications (BUPT) in 2005 and 2010, respectively.

From 2017 to 2018, she was a Visiting Scholar with the University of British Columbia, Vancouver, BC, Canada. She is currently an Associate Professor with the School of Information and Communication Engineering, BUPT. She has published more than 100 papers in international journals and conferences. Her current research interests include resource management and intelligent networking in next generation networks, the Internet of Things, and cloud computing. She has also served as a TPC Member of IEEE WCNC 2012/2014/2015/2016/2019, PIMRC 2012/2017/2018/2019, GLOBECOM 2015/2017/2018, ICC 2015/2016/2017/2018/2019, Infocom 2018, and CloudCom 2013/2014/2015, the Chair of Special Track on cognitive testbed in CHINACOM 2011, the Workshop Chair of IEEE GreenCom 2019, and a peer reviewer of many academic journals. She is serving on the editorial boards for the INTERNATIONAL JOURNAL OF COMMUNICATION SYSTEMS (Wiley).



Hong Ji (Senior Member, IEEE) received the B.S. degree in communications engineering and the M.S. and Ph.D. degrees in information and communications engineering from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1989, 1992, and 2002, respectively.

In 2006, she was a Visiting Scholar with the University of British Columbia, Vancouver, BC, Canada. She is currently a Professor with BUPT. She has authored more than 300 journal/conference papers. Several of her papers had been selected for best paper. Her research interests include wireless networks and mobile systems, including cloud computing, machine learning, intelligent networks, green communications, radio access, ICT applications, system architectures, management algorithms, and performance evaluations. She has guest-edited the *International Journal of Communication Systems* (Wiley), special issue on *Mobile Internet: Content, Security and Terminal*. She has served as the Co-Chair for Chinacom'11, and a member of the Technical Program Committee for WCNC'19/15/14/12, Globecom'17/16/15/14/13/12/11/10, ISCIT'17, CITS'16/15/12, WCSP'15, ICC'13/12/11, ICCC'13/12, PIMRC'12/11, IEEE VTC'12S, and MobiWorld'11. She is serving on the editorial boards for the IEEE TRANSACTION ON GREEN COMMUNICATIONS AND NETWORKING and the *International Journal of Communication Systems* (Wiley).

Victor C. M. Leung (Fellow, IEEE) is a Distinguished Professor of computer science and software engineering with Shenzhen University. He is also an Emeritus Professor of electrical and computer engineering and the Director of the Laboratory for Wireless Networks and Mobile Systems with the University of British Columbia (UBC). His research is in the broad areas of wireless networks and mobile systems. He has coauthored more than 1300 journal/conference papers and book chapters. He is serving on the editorial boards for the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, the IEEE TRANSACTIONS ON CLOUD COMPUTING, IEEE ACCESS, IEEE NETWORK, and several other journals. He received the IEEE Vancouver Section Centennial Award, the 2011 UBC Killam Research Prize, the 2017 Canadian Award for Telecommunications Research, and the 2018 IEEE TCGCC Distinguished Technical Achievement Recognition Award. He coauthored papers that won the 2017 IEEE ComSoc Fred W. Ellersick Prize, the 2017 IEEE Systems Journal Best Paper Award, the 2018 IEEE CSIM Best Journal Paper Award, and the 2019 IEEE TCGCC Best Journal Paper Award. He is named in the current Clarivate Analytics list of "Highly Cited Researchers." He is a fellow of the Royal Society of Canada, the Canadian Academy of Engineering, and the Engineering Institute of Canada.