Blockchain and Deep Reinforcement Learning Empowered Intelligent 5G Beyond

Yueyue Dai, Du Xu, Sabita Maharjan, Zhuang Chen, Qian He, and Yan Zhang

Abstract

Blockchain and AI are promising techniques for next-generation wireless networks. Blockchain can establish a secure and decentralized resource sharing environment. AI can be explored to solve problems with uncertain, time-variant, and complex features. Both of these techniques have recently seen a surge in interest. The integration of these two techniques can further enhance the performance of wireless networks. In this article, we first propose a secure and intelligent architecture for next-generation wireless networks by integrating AI and blockchain into wireless networks to enable flexible and secure resource sharing. Then we propose a blockchain empowered content caching problem to maximize system utility, and develop a new caching scheme by utilizing deep reinforcement learning. Numerical results demonstrate the effectiveness of the proposed scheme.

Introduction

The Federal Communications Commission (FCC) proposed the vision of the sixth-generation (6G) wireless network at Mobile World Congress Americas (MWCA) and emphasized that blockchain will play a crucial role toward a more distributed and secure network [1]. Moreover, artificial intelligence (AI) is a promising technique that can be integrated into a 6G network to empower self-aggregating communication and intelligent resource orchestration.

Blockchain is an open database that maintains an immutably distributed ledger typically deployed in a peer-to-peer (P2P) network. Blockchain enables registering and updating transactions securely in a decentralized fashion without relying on a central intermediary. Recently, blockchain technology has received enormous attention due to its features such as decentralization, immutability, anonymity, and security. The authors in [2, 3] proposed to exploit blockchain to develop a secure localized P2P electricity trading system for locally buying and selling electricity among electric vehicles. FCC proposed that blockchain can be utilized in the next-generation wireless network to reduce the administrative expense for dynamic access systems that increase spectral efficiency [1]. The authors in [4] utilized blockchain and smart contract technologies to design a reputation-based data sharing scheme in the vehicular edge network. Since forming blockchain requires a mining task, that is, proof-of-work (PoW), which demands both heavy computation resources and energy, it is hard to widely facilitate blockchain in a wireless network [5]. The authors in [6] proposed to utilize the mobile edge computing (MEC) paradigm to solve this issue.

Heterogeneous networks and device-to-device (D2D) communications are introduced in the 5G network to boost the communication rate and simultaneously guarantee seamless coverage. MEC is a new paradigm that can significantly reduce latency and avoid backhaul congestion via computation offloading [7–9] and distributed content caching [10]. The authors in [7] proposed a two-tier multi-task offloading scheme to prolong battery life of mobile users in heterogeneous networks. The authors in [8] proposed task offloading to minimize task duration in software defined ultra dense networks. The authors in [9] integrated computation offloading with load balancing to maximize system utility in vehicular networks. The authors in [10] proposed a distributed content caching scheme to jointly optimize computing and caching resources on edge servers to alleviate backhaul congestion. However, due to time-variant wireless channels, the diverse and stringent requirements of various emerging applications and unknown traffic systems, designing high-performance algorithms to make full use of the above technologies is quite a challenge that essentially demands novel approaches.

AI has the ability to interact with the wireless environment to facilitate resource management and orchestration. The authors in [11] utilized deep reinforcement learning (DRL) to tackle the joint edge computing and caching resource allocation problem. In addition, with the expansion of the network scale, the amount of data generated by applications and networks will experience an exponential growth. AI exhibits great potential on in-depth feature discovery and events forecasting utilizing such data [12]. The authors in [13] designed a label-less learning-based traffic control algorithm to reduce the amount of offloaded data.

Although some works have introduced blockchain and AI in a wireless network, most of these works have studied them separately. In this article, we integrate blockchain and AI into wireless networks and propose a secure and intelligent architecture for next-generation wireless networks. The proposed architecture enables secure and

Yueyue Dai and Du Xu are with the University of Electronic Science and Technology of China; Sabita Maharjan is with Simula Metropolitan Center for Digital Engineering and University of Oslo; Yan Zhang (corresponding author) is with the University of Oslo; Zhuang Chen and Qian He are with Guilin University of Electronic Technology.

Digital Object Identifier:
10.1109/MNET.2019.1800376

0890-8044/19/$25.00 © 2019 IEEE

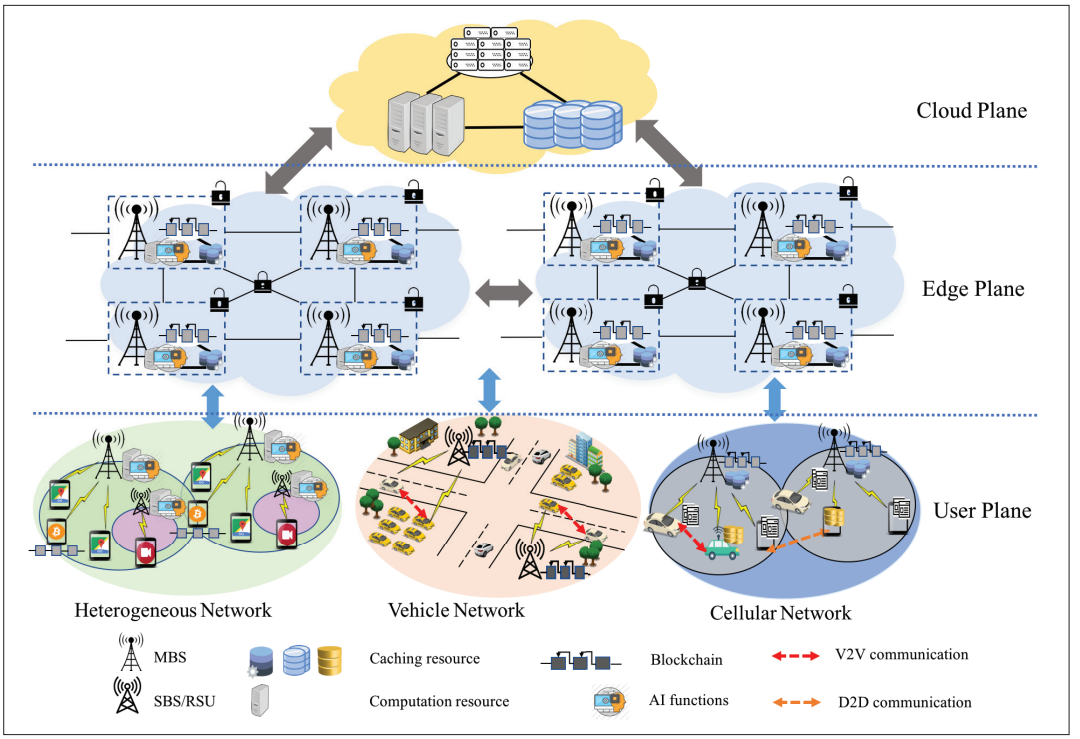

FIGURE 1. The secure and intelligent architecture of next-generation wireless networks.

intelligent resource management and orchestration, that is, spectrum sharing, content caching, energy trading, and computation offloading. Then we exploit consortium blockchain to establish a secure content caching environment and utilize the advanced deep reinforcement learning algorithm to design a caching scheme for maximizing system utility. Numerical results demonstrate that the proposed scheme can efficiently use D2D caching and allocate resources to enhance system utility. The main contributions of this article can be summarized as follows.

We propose a secure and intelligent architecture for the next-generation wireless network. Leveraging AI and blockchain, this architecture enables secure and intelligent resource management and orchestration.

We introduce four typical blockchain empowered wireless resource management schemes, that is, spectrum sharing, content caching, energy trading, and computation offloading.

We formulate a caching-resource sharing problem in a blockchain-enabled framework to maximize system utility, and design a novel caching-resource sharing scheme by utilizing the advanced DRL algorithm.

A Secure and Intelligent Architecture for Next-generation Wireless Networks

Secure and Intelligent Architecture: An Overview

We aim to integrate blockchain and AI into next-generation wireless networks to enable secure network orchestration, flexible networking, and intelligent resource management. Figure 1 illustrates our proposed hierarchical architecture for next-generation wireless networks.

The architecture consists of three planes: the cloud plane, edge plane, and user plane. A number of servers are equipped with strong computation, caching, and processing capabilities at the cloud plane. With a global view, this layer can leverage advanced techniques such as data mining and big data to make a network-level orchestration shifting from reactive network operation to proactive network operation, by predicting some events or pre-allocating some resources. Due to high computing capability and sufficient caching resource, cloud servers can process delay-tolerant applications and store content with large size or less popularity. Further, there is a central authority in the cloud plane. The central authority is equipped with tamper-resistant hardware, and it manages security parameters and keys of all entities: macro base station (MBS), small base station (MBS), roadside units (RSUs), mobile devices, and smart vehicles.

Nearby network infrastructures (e.g., MBSs, SBSs, RSUs) are geo-distributed at the network edge and equipped with MEC servers and blockchain. Network infrastructures can provide radio interfaces for mobile devices and vehicles to achieve seamless coverage and instant wireless communication. MEC servers, with computation resources, caching resources, and AI functions, can provide distributed intelligent wireless computing and caching to implement computation-intensive and delay-sensitive applications and cache the most popular or important content such as the latest news and emergency warning at the network edge. Blockchain can record all transactions generated in the wireless network and maintain a distributed ledger to increase the security and privacy of the wireless ecosystem. The transactions can be spectrum sharing, computation/caching resource allocation, energy trading, and so on.

	Public blockchain	Private blockchain	Consortium blockchain
Access	Anyone	Single organization	Permissioned nodes
Energy cost	High	Low	Low
Delay	Long	Short	Short
Security	High	Low	High
Existing work	MEC [6]	IoT [14]	Energy trading [2, 3], Data sharing [4]

TABLE 1. Comparison of different types of blockchains.

To make a programmable, flexible, and elastic mobile edge plane, network functions virtualization and software-defined network technologies are deployed. Since network functions virtualization is able to abstract physical resources and establish virtual machines, the edge plane can ignore the difference in terms of vendor and protocol, and realize fast function deployment by creating, migrating, and destroying the virtual machines among distributed edge entities. Software-defined networking can decouple network control and management functions from data forwarding such that the edge plane can apply dynamic resource management and intelligent service orchestration.

Heterogeneous networks, vehicle-to-everything (V2X), and cellular networks coexist in the user plane to boost communication rate and simultaneously ensure seamless coverage to support the highly reliable connectivity required for various emerging applications such as automated driving and virtual reality.

In a heterogeneous network, each mobile device has a computation-intensive and delay-sensitive task, such as navigation, video streaming, and bitcoin. MBSs and SBSs are equipped with computation resources and AI functions. Resource-limited mobile devices can offload their tasks to the heterogeneous edge infrastructure, and base stations can utilize fine-grained computation resource allocation policy to process the offloaded tasks. Since the interactions between mobile devices and base stations are trust-based, there is no need to utilize blockchain in this scenario.

In vehicular networks, the scenario supports vehicle-to-RSU and vehicle-to-vehicle (V2V) communication. Based on V2V communication, content or energy can be shared among vehicles. Since vehicles may not trust each other, they need to use pseudonyms when sharing content or energy for better security and privacy protection. To this end, we deploy blockchain on RSUs.

In a cellular network, we consider a more general and complex scenario where D2D and V2V communication are also supported. MBSs often have limited amounts of caching resource. In reality, some state-of-the-art mobile devices and vehicles are deployed with a certain amount of caching resources. Thus, base stations, mobile devices, and vehicles can cooperatively provide distributed edge caching to make full use of the available resources. That is, a specific mobile device or vehicle can select any other mobile device or vehicle with sufficient caching resource as their service provider for content caching. Since D2D and V2V communication are not trusted, BSs also need to utilize blockchain technology to ensure security of transactions in this scenario.

Advantages of Our Proposed Architecture

Leveraging blockchain and AI, the proposed hierarchical architecture is expected to bring a variety of benefits, which are summarized as follows:

Secure and Intelligent Resource Management: In the proposed architecture, spectrum management among mobile users, V2V energy trading, or caching sharing can be recorded into tamper-resistant blocks to achieve secure resource sharing or allocation. AI algorithms can automatically perceive complicated wireless networks, diverse requirements of emerging services, and time-variant states of available resources. Therefore, exploiting blockchain and AI, the proposed architecture can perform optimal resource allocation policies in a secure environment.

Flexible Networking: The coexistence of MBS, SBS, mobile device, and vehicle in the proposed architecture provides an opportunity for flexible networking. AI can accurately analyze the topology, channel assignment, and interference of the current wireless network, and then select the most appropriate wireless access mode (i.e., cellular network, V2V, or D2D) to improve communication rate, reduce energy consumption, or enhance user experience. For example, the architecture can generate user-specific policies to make some mobile users communicate with the MBS while others connect to RSUs to maintain fundamental information exchange.

Reliable and Dynamic Orchestration: Assisted by blockchain, operating reports and network configurations can be replicated and synchronized in a decentralized manner among edge servers, which can facilitate network diagnosis and enable reliable orchestration. AI can generate operating reports to describe and summarize some network features, and can provide a fast and dynamic mechanism to migrate virtual machines. In addition, computation and caching resources at the edge plane can facilitate even more powerful network orchestration.

Blockchain Empowered Secure and Intelligent Resource Management

Blockchain can be categorized into three main types: public blockchain, private blockchain, and consortium blockchain. We show the key features of these blockchains in Table 1. Public blockchain is the traditional blockchain in which anyone can participate in the process of verifying transactions, creating blocks, and getting consensus due to no access limitation, such as Bitcoin and Ethereum. However, since the mining process in public blockchain incurs high energy cost and long delay, the authors in [6] proposed to utilize an MEC system to support physically distributed, low-latency, and quality of service (QoS)-aware applications. Private blockchain is mastered by a specific organization, which is more like the centralized architecture. Due to the advantages of low energy consumption and short delay, private blockchain is suitable for energy-constrained and delay-sensitive networks, such as the Internet of Things (IoT) [14]. However, the security of this type blockchain is low. Consortium blockchain

utilizes permissioned nodes to create new blocks in a mutually untrustworthy environment. Without mining process, consortium blockchain is more efficient compared to public blockchain. Also, the high level of security offered by consortium blockchain can provide a secure resource trading or sharing environment [2–4].

A blockchain consists of three essential components: transactions, blocks about transaction records, and a consensus algorithm. The transaction information includes nodes' pseudonyms used for privacy protection, data type, metadata tags for raw transactional data, complete index history of metadata, an encrypted linked to transaction records, and a timestamp of transaction generation [2]. Each transaction is encrypted and signed with digital signatures to guarantee authenticity. The digitally signed transactions are arbitrarily packed up into a cryptographically tamper-evident data structure named block. The blocks are linked in a linear chronological order by hash pointers to form a blockchain. To maintain the consistency and order of the blockchain, a consensus algorithm is designed to generate an agreement on the order of the blocks and to validate the correctness of the set of transactions that constitute the block. In consortium blockchain, networks often adopt the practical Byzantine fault tolerant (PBFT) consensus algorithm. In PBFT, the leader can create a new block. After receiving the newly created block, a small group of preselected nodes participate in a voting process for reaching the consensus. On the contrary, in a public blockchain, for establishing consensus, every node needs to solve a PoW puzzle and then participates in a voting process.

Security in the wireless network is challenging due to lack of standardization. Moreover, many wireless entities share their resources or contents openly without considering personal privacy. To establish a secure and private wireless communication environment, we integrate blockchain into the wireless network and discuss four potential blockchain-empowered resource management cases, as shown in Fig. 2.

Spectrum Sharing: The need to accommodate diverse types of users, applications with diverse performance requirements, and the need to integrate heterogeneous air interfaces into the next generation wireless networks are likely to make the radio spectrum more congested. Cognitive radio is a spectrum sharing technology that can estimate communication parameters and automatically perform spectrum resource allocation in a time-varying wireless environment. In a cognitive radio system, radio spectrum is owned by the primary user who can lease the spectrum to secondary users according to a specific spectrum sharing scheme. However, secondary users have to share their private information for using the spectrum of the primary users. Blockchain can be utilized in this context to enable a secure spectrum sharing application while integrating privacy protection for the secondary users.

A distributed spectrum sharing system is shown in Fig. 2a, where each SBS is equipped with a blockchain. The SBS with licensed spectrum is the primary user, and the SBS without licensed spectrum is the secondary user. If a primary user successfully leases spectrum to a secondary user, the secondary user pays a reward to the primary user and simultaneously forms a spectrum-leasing transaction. Each spectrum-leasing transaction should be verified and then stored in a block to avoid tampered and faked spectrum-leasing records. Each block is linked to the previous block, forming a chain. The blocks record thousands of spectrum sharing transactions among SBSs in a decentralized manner.

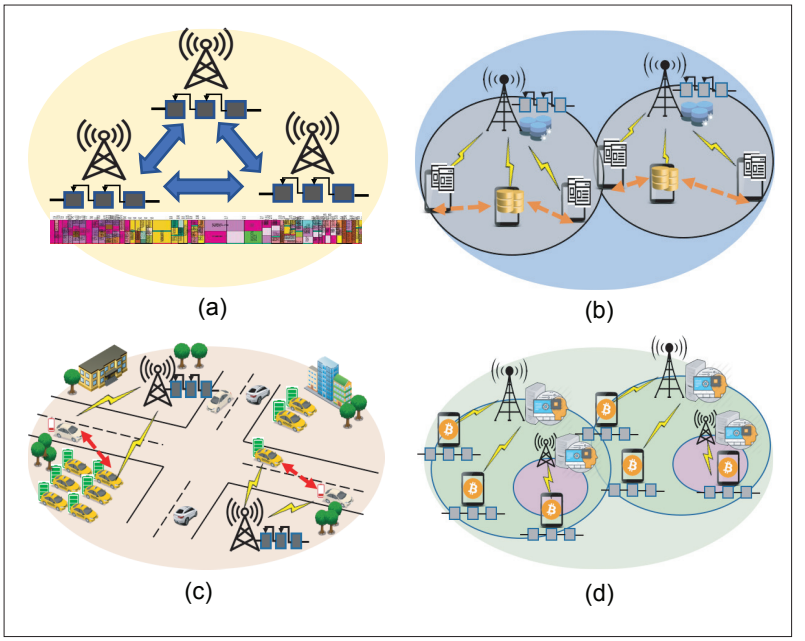

FIGURE 2. Potential blockchain-empowered resource management: a) spectrum sharing; b) D2D caching; c) energy trading; d) computation offloading.

Cognitive radio can utilize AI to design the optimal spectrum sharing police to maximize long-term rewards by interacting with the radio frequency environment.

D2D Caching: Contents generated by sensors or multimedia applications are undergoing exponential growth, which can possibly challenge the capacity of MBSs. Since some state-of-the-art devices (e.g., smartphones) have certain caching resources, large-scale content can be cached in these entities through D2D communications. Caching content at mobile devices is a promising approach for reducing data traffic on backhaul links, as well as for enhancing QoS for end users. However, since a content usually involves much sensitive and critical personal information of its generator, caching requesters are not willing to store their contents with an untrusted caching provider [4]. Since blockchain enables untrusted nodes to interact with each other in a secure manner, it provides a promising way to D2D caching.

Figure 2b shows blockchain empowered D2D caching. In this system, a resource-constrained mobile device with a large-scale content is defined as a caching requester. The device with sufficient caching resources is defined as a caching provider. MBSs are equipped with AI algorithms to predict the D2D communication duration between caching requesters and caching providers, and perform caching pair matching and resource allocation to enhance caching hit ratio or system utility. If a content is successfully stored at one caching provider, the caching requester creates a transaction and sends it to the nearest MBS.

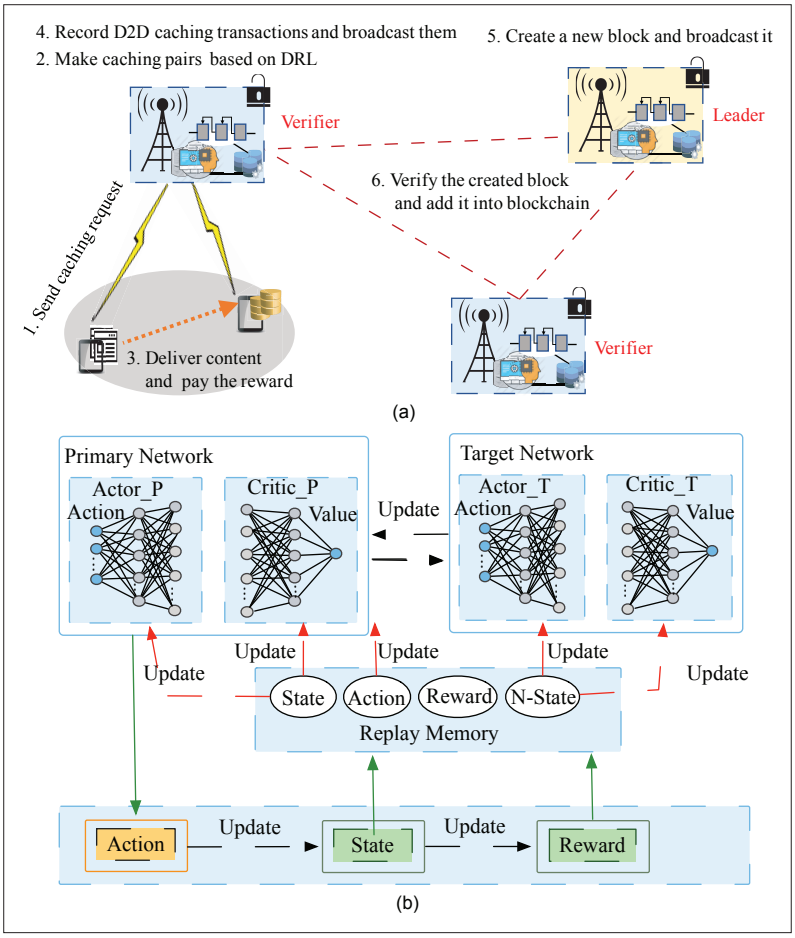

FIGURE 3. DRL and blockchain empowered content caching: a) secure content caching using consortium blockchain; b) DRL-empowered caching scheme.

Each MBS broadcasts received transactions to other MBSs and builds a new block after receiving a certain number of transactions. Note that the caching resources on MBSs are utilized to store the whole transactions about D2D caching.

Energy Trading: Smart vehicles can not only charge electricity from a fixed charging point, but can also get electricity from other smart vehicles with surplus electricity. However, smart vehicles with surplus electricity may not be willing to work as energy suppliers in a localized P2P energy trading market due to privacy concerns [2, 3]. To encourage vehicles with surplus electricity to participate in energy trading, it is necessary to protect the privacy of smart vehicles during the trade.

Figure 2c illustrates a secure V2V energy trading system. There are two types of vehicles: one type needs to charge, and the other type has surplus electricity. The vehicles that need to charge send their charging request to the nearest RSU via V2R channel. The RSUs broadcast the received requests to local vehicles with surplus electricity. Then the vehicles with surplus electricity respond to the RSU with their state of charge. Each RSU utilizes an AI algorithm, such as DRL, to match energy trading pairs. Here, successful energy trading among vehicles is defined as an energy transaction. Leveraging blockchain, all transactions about energy trading are packed as a block and recorded in a blockchain.

Computation Offloading: Blockchain can be considered as an application deployed at mobile devices. For example, as shown in Fig. 2d, each mobile device is equipped with a mining-based blockchain such as bitcoin. To support bitcoin, mobile devices have to solve a PoW puzzle. However, the PoW puzzle is a computation-intensive and energy-consuming task such that resource-constrained devices cannot supply sufficient computation resources and energy to maintain bitcoin. In an MEC framework, mobile devices can offload the PoW task to nearby BSs and utilize the distributed computing enabled by the BSs. Base stations compute and provide results, (i.e., block and hash pointer about the transactions) to mobile devices.

In light of the cases discussed above, it is implied that blockchain and wireless networks are complementary. Blockchain can establish a secure and trusted resource allocation and sharing environment for wireless networks. Conversely, wireless networks provide distributed but accessible computation resources and energy for implementing blockchain.

Deep Reinforcement Learning and Blockchain-Empowered Content Caching

In this section, we utilize content caching as an example case to elaborate on how blockchain and AI can be jointly exploited to achieve secure and intelligent resource management. We first establish a secure content caching environment via consortium blockchain and then utilize DRL to design a content caching scheme for maximizing caching resource utility.

Content Caching Blockchain

Figure 3a illustrates the framework of content caching in D2D networks and shows the detailed caching process based on consortium blockchain. In the content caching system, each MBS maintains a blockchain, and the D2D caching transactions occur among mobile devices. Specifically, if a content is successfully cached at one caching provider, the caching requester should create a transaction record and send it to the nearest MBS. MBSs collect and manage their local transaction records. The transaction records are structured into blocks after finishing the consensus process among the MBSs and then stored in each MBS permanently. In the following, we describe the key operations in a caching blockchain.

System Initialization: To protect privacy, each mobile device needs to register a legitimate identity in the system initialization stage. In D2D caching blockchain, an elliptic curve digital signature algorithm and asymmetric cryptography are used for system initialization. A mobile device d_i can obtain a legitimate identity after passing identity authentication. The identity includes a public key, a private key, and the corresponding certificate (i.e., $\{PK_i, SK_i, Cert_i\}$).

Choosing Roles in D2D Caching: For D2D caching, mobile devices choose their roles (i.e. caching requester and caching provider) according to their current caching resource availability state and future plans. Mobile devices with surplus caching resource can become caching providers to provide caching service for caching requesters.

Caching Transactions: Caching requesters send the amount of caching resource and expected serving time to the nearest MBS. The MBS broadcasts all received caching requests to local caching providers. Caching providers feed back the amount of caching resource to the MBS and their future plans. Then each MBS utilizes a DRL algorithm to match the caching supply and demand pairs among mobile devices, determine the caching resource that each caching provider can provide, allocate bandwidth between the MBS and mobile devices.

Building Blocks in Caching Blockchain: MBSs collect all transaction records in a certain period, and then encrypt and digitally sign these records to guarantee the authenticity and accuracy of these records. The transaction records are structured into blocks, and each block contains a cryptographic hash to the prior block in the consortium blockchain. To verify the correctness of the new block, PBFT is used. According to PBFT, there is a leader that is responsible for creating a new block. Because of broadcast, each MBS has access to the whole transaction record and has the opportunity to be the leader. In consortium blockchain, the leader is chosen before the block building and does not change before finishing the consensus process.

Carrying Out the Consensus Process: The leader broadcasts the created block to other MBSs for verification and audit. All MBSs audit the correctness of the created block and broadcast their audit results. The leader will analyze the audit results and send the block to these MBSs once again for audit if necessary. Moreover, according to audit results and corresponding signatures, compromised MBSs will be found out and held accountable.

DRL-Empowered Caching Scheme

The content caching problem can be formulated as an optimization problem to maximize system utility and solved using a DRL algorithm. We consider a cache network with K BSs, M caching requesters, and N caching providers. MBSs are equipped with blockchain and the DRL algorithm for security and decision making. Each caching request has a large-scale content d_i, such as a multimedia file. If the content of caching requester i is stored at caching provider j, $x_{ij} = 1$; otherwise, $x_{ij} = 0$. Two mobile devices can transmit contents to each other via D2D communications as long as the distance between them is less than a pre-defined communication range. System utility consists of caching utility and energy cost. Caching utility is equal to $x_{ij} \cdot d_i \cdot B_i$, where B_i is the price for storing content. The cost is the energy consumption on communication and caching, respectively.

Exploiting the proposed architecture, the information in terms of caching capabilities, the requirements of caching requesters, and each content size can be collected and sent to the agent. Then the agent designs an action to match caching pairs and allocate resources. There are three key elements in the deep reinforcement learning process, namely state, action, and reward:

State: The state in DRL is a space to reflect the environment. The state consists of three components $S = (D_i, C_j, B_j)$, where D_i denotes the state of content i, C_j is the available caching resource, and B_j is the available bandwidth of caching provider j. In this environment, each MBS assembles the above information as a state and sends it to the agent.

Action: The objective of an agent is to map the space of states to the space of actions. In this system, the action consists of two parts: x_{ij}, and b_{ij}, where x_{ij} is a binary value and b_{ij} is the amount of bandwidth.

Reward: Based on current state and action, the agent obtains a reward from the environment. Since reward function is related to the objective function, in this scenario, system utility can be regarded as the reward function.

The DRL process to design the content caching policy is shown in Fig. 3b, which is based on a deep deterministic policy gradient method [15].

In DRL, the primary network consists of two deep neural networks, namely the actor network and the critic network. The actor network is used to explore the policy, and the critic network estimates the performance and provides the critic value, which helps the actor to learn the gradient of the policy.

The target network can be defined as an old version of the primary network, which is used to generate the target value for training Critic-P. It includes a target actor network and a target critic network. The input of the target network is the next state (i.e., N-State) from replay memory and the output is a critic value for training Critic-P.

The replay memory stores experience tuples that include current state, the selected action, reward, and next state. The stored experience tuples can be randomly sampled for training the primary network and the target network. Randomly sampling experience tuples aim to reduce the effects of data correlation.

The replay memory stores experience tuples which include current state, the selected action, reward, and next state. The stored experience tuples can be randomly sampled for training primary network and target network. Randomly sampling experience tuples aim to reduce the effects of data correlation.

Numerical Results

We evaluate the performance of the proposed DRL-empowered D2D caching scheme through extensive simulations. The proposed architecture implements two parts: the environment and the agent. In the environment, there are 5 caching providers and 20 caching requesters randomly distributed in a 500 m × 500 m area. The caching resources of caching providers are randomly taken from [30, 31, 32, 35, 40] GB. The maximal bandwidth of caching providers is randomly taken from [22, 24, 25, 28, 30] MHz. The data size of each content is randomly taken from [3, 6, 9] GB.

The proposed DRL-empowered D2D caching scheme simultaneously performs caching pairs matching and dynamic bandwidth allocation to maximize the system utility, while the DRL-empowered benchmark caching scheme only performs caching pair matching. We first present the comparison of cumulative average system utility under different schemes in Fig. 4. From Fig. 4, we can see that the cumulative average system utility

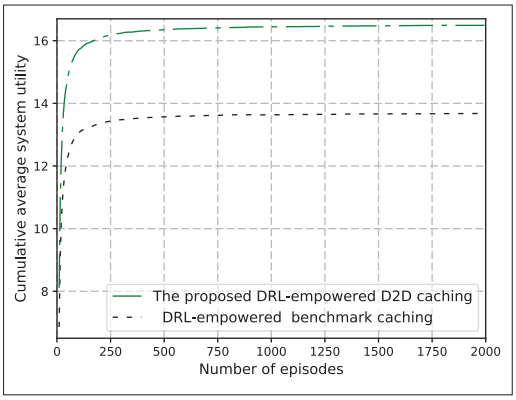

FIGURE 4. Comparison of system utility under different schemes.

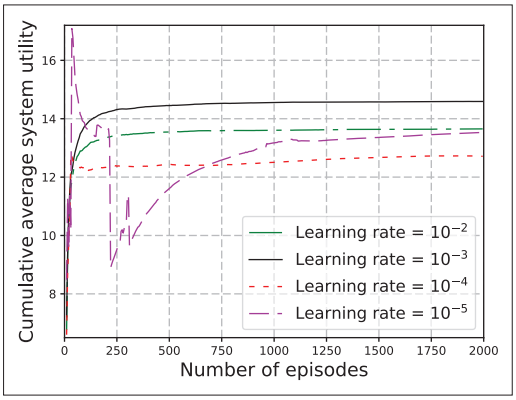

FIGURE 5. Convergence performance under different learning rates.

of the proposed DRL-empowered D2D caching scheme is obviously higher than the benchmark caching scheme. The reason is that the proposed DRL-empowered D2D caching scheme can select the most appropriate caching provider for a specific caching requester and optimize the bandwidth between the caching provider and the caching requester to further improve system utility. However, the benchmark caching scheme performs caching pair matching without bandwidth allocation, which results in higher communication energy cost. Moreover, we observe that each system utility of different schemes is very low at the beginning of the learning process. With an increasing number of episodes, system utilities reach a relatively stable value after running 750 episodes iterations.

Figure 5 shows the convergence performance of the proposed scheme under different learning rates. First, the cumulative average system utilities achieve convergence in all learning rates. Second, when the learning rate is 10^{-3}, the cumulative average system utility is obviously higher than the cases when the learning rate is 10^{-4} and 10^{-5}, which implies that a small learning rate achieves a better performance. However, the performance with learning rate of 10^{-3} is also better than that with learning rate of 10^{-2}. Thus, we can conclude that 10^{-3} is the best learning rate for the proposed DRL-empowered D2D caching scheme. In fact, an appropriate learning rate depends on the architecture of the mode being optimized, as well as the state of the environment in the current optimization process.

Conclusion and Future Work

In this article, we have proposed a secure and intelligent hierarchical architecture for next-generation wireless networks by integrating blockchain and AI into the wireless network. The proposed architecture can enable secure and intelligent resource management, flexible networking, and reliable orchestration. Then we have presented four typical blockchain empowered wireless resource management schemes, that is, spectrum sharing, D2D caching, V2V energy trading, and computation offloading. Furthermore, we have exploited consortium blockchain to establish a secure content caching environment and utilized the advanced deep reinforcement learning to design a caching scheme for maximizing caching resource utility. Numerical results have validated the effectiveness of the proposed scheme.

Acknowledgment

This research has partially received funding from the National Key Research and Development Program of China-No. 2016YFB0800105; the European Union's Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No. 824019; projects 240079/F20 funded by the Research Council of Norway; project 18H86301ZT00100303; the 111 project (B14039); the National Natural Science Foundation of China (61661015); and the study abroad program for graduate student of Guilin University of Electronic Technology.

References

[1] "FCC's Rosenworcel Talks Up 6G?"; https://www.multichannel.com/news/fccs-rosenworcel-talks-up-6g, 2018.
[2] J. Kang *et al.*, "Enabling Localized Peer-to-Peer Electricity Trading Among Plug-In Hybrid Electric Vehicles Using Consortium Blockchains," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, 2017, pp. 3154–64.
[3] Z. Li *et al.*, "Consortium Blockchain for Secure Energy Trading in Industrial Internet of Things," *IEEE Trans. Ind. Informat.*, 2017.
[4] J. Kang *et al.*, "Blockchain for Secure and Efficient Data Sharing in Vehicular Edge Computing and Networks," *IEEE Internet of Things J.*, 2018.
[5] A. Panarello *et al.*, "Blockchain and IoT Integration: A Systematic Survey," *Sensors*, vol. 18, no. 8, 2018, p. 2575.
[6] M. A. Rahman *et al.*, "Blockchain-Based Mobile Edge Computing Framework for Secure Therapy Applications," *IEEE Access*, 2018.
[7] Y. Dai *et al.*, "Joint Computation Offloading and User Association in Multi-Task Mobile Edge Computing," *IEEE Trans. Vehic. Tech.*, vol. 67, no. 12, Dec. 2018, pp. 12,313–25.
[8] M. Chen and Y. Hao, "Task Offloading for Mobile Edge Computing in Software Defined Ultra-Dense Network," *IEEE JSAC*, vol. 36, no. 3, 2018, pp. 587–97.
[9] Y. Dai *et al.*, "Joint Load Balancing and Offloading in Vehicular Edge Computing and Networks," *IEEE Internet of Things J.*, 2018, pp. 1–1.
[10] M. Chen *et al.*, "Edge-Cocaco: Toward Joint Optimization of Computation, Caching, and Communication on Edge Cloud," *IEEE Wireless Commun.*, 2018, p. 2.
[11] Y. Dai *et al.*, "Artificial Intelligence Empowered Edge Computing and Caching for Internet of Vehicles," *IEEE Wireless Commun.*, accepted, 2018.
[12] M. Mohammadi and A. Al-Fuqaha, "Enabling Cognitive Smart Cities Using Big Data and Machine Learning: Approaches and Challenges," *IEEE Commun. Mag.*, vol. 56, no. 2, Feb. 2018, pp. 94–101.
[13] M. Chen *et al.*, "Label-Less Learning for Traffic Control in an Edge Network," *IEEE Network*, vol. 32, no. 6, Dec. 2018, pp. 8–14.
[14] T.-T. Kuo and L. Ohno-Machado, "Modelchain: Decentralized Privacy-Preserving Healthcare Predictive Modeling Framework on Private Blockchain Networks"; https://arxiv.org/abs/1802.01746, 2018.
[15] T. P. Lillicrap *et al.*, "Continuous Control with Deep Reinforcement Learning," *Proc. Int'l. Conf. Learning Representations*, 2016.

Biographies

Yueyue Dai [S'17] (yueyuedai@ieee.org) received her B.Sc. degree in communication and information engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, in 2014, where she is currently pursuing a Ph.D. degree. Since October 2017, she has been a visiting student with the Department of Informatics, University of Oslo, Norway. Her current research interests include wireless networks, mobile edge computing, the Internet of Vehicles, blockchain, and deep reinforcement learning.

Du Xu [M'17] (xudu.uestc@gmail.com) is a professor at UESTC. He received a B.S., an M.S., and a Ph.D. from South-East University and UESTC in 1990, 1995, and 1998, respectively. His research interests include network modeling and performance analysis, switching and routing, network virtualization, and security. He has presided over many advanced research projects, including NSFC, National 863 Plans, and the National Key Research and Development Program of China.

Sabita Maharjan [M'09] (sabita@simula.no) received her Ph.D. degree in networks and distributed systems from the University of Oslo and Simula Research Laboratory, Norway, in 2013. She is currently a senior research scientist at the Simula Metropolitan Center for Digital Engineering, Norway, and an associate professor at the University of Oslo. Her current research interests include wireless networks, network security and resilience, smart grid communications, the Internet of Things, machine-to-machine communications, software defined wireless networking, and the Internet of Vehicles.

Zhuang Chen (zhuangchenuio@gmail.com) received his B.S. degree in Internet of Things engineering from Qingdao University of Science and Technology, China, in 2017. He is pursuing an M.S. degree with the School of Computer and Information Security, Guilin University of Electronic Technology, China. He is also currently a visiting student with the Department of Informatics, University of Oslo. His current research interests include mobile edge computing, multimedia cache, wireless networks, deep reinforcement learning, and blockchain.

Qian He [M] (heqian@guet.edu.cn) is a full professor at Guilin University of Electronic Technology, China. He received his B.Sc. from Hunan University, Changsha, China, in 2001, his M.S. from Guilin University of Electronic Technology in 2004, and his Ph.D. from Beijing University of Posts and Telecommunications, China, in 2011. His research interests include network security and distributed computing. He is a member of ACM and a Senior Member of CCF.

Yan Zhang [SM'10] (yanzhang@ieee.org) is a full professor at the University of Oslo. He is an Editor of several IEEE publications, including *IEEE Communications Magazine*, *IEEE Network*, *IEEE Transactions on Green Communications and Networking*, *IEEE Communications Surveys & Tutorials*, and *IEEE Internet of Things Magazine*. He received the Highly Cited Researcher award (Web of Science top 1 percent most cited) according to Clarivate Analytics. His current research interests include next generation wireless networks leading to 5G and cyber physical systems. He is an IEEE VTS Distinguished Lecturer and a Fellow of IET.