IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON SECURITY AND PRIVACY IN EMERGING DECENTRALIZED COMMUNICATION ENVIRONMENTS

Received October 30, 2019, accepted November 20, 2019, date of publication December 2, 2019,
date of current version December 18, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2956955

Proof-of-Reputation Based-Consortium Blockchain for Trust Resource Sharing in Internet of Vehicles

HAOYE CHAI, SUPENG LENG, KE ZHANG, AND SUN MAO
School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

Corresponding author: Supeng Leng (spleng@uestc.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFE0117500, in part by the Science and Technology Program of Sichuan Province, China, under Grant 2019YFH0007, in part by the Fundamental Research Funds for the Central Universities, China, under Grant 2672018ZYG X2018J001, and in part by the EU H2020 Project COSAFE under Grant MSCA-RISE-2018-824019.

ABSTRACT Resource sharing among vehicles can highly improve the capability and efficiency of Internet of Vehicles (IoV). However, it is challenging to establish trust and preserve privacy during the resource sharing process because of the high mobility and topological variability in IoV. Emerging blockchain technology expresses the excellent performance in handling distributed trust due to its verifiable and immutable ledger. In this paper, we first propose a consortium blockchain-based resource sharing paradigm in IoV, in which the resource sharing interactions are encapsulated as transactions and recorded by Road Side Units (RSUs). Moreover, a lightweight consensus mechanism named as Proof-of-Reputation is proposed to reduce computational power consumption and motivate vehicles involved in resource sharing. Finally a differentiated resource pricing scheme is proposed based on the dynamic match game of resource demand and supply. The reputation value is designed to indicate the trustworthy degree of vehicles, and the trust is established via the consensus procedure. We couple the resource sharing process and consensus together by utilizing the reputation value of each vehicle. The security and privacy analysis as well as simulation experiments on communication performance can verify the efficiency of the proposed blockchain system.

INDEX TERMS Internet of Vehicles, lightweight blockchain, trust management.

I. INTRODUCTION

With the emergence of internet of Vehicles (IoV), various smart devices are linked to enable the fast and efficient applications [1]. IoV is an emerging paradigm to support the evolution of Intelligent Transportation Systems (ITS), which is highly characterized by gathering, sharing, processing, computing, and secure release of data services onto information platforms. IoV shows advantages on providing various services such as autonomous driving, accident alarming and mobile advertising [2], [3]. However, due to the high mobility and short interaction time of vehicles, it is difficult to directly utilize traditional cellular-based communication technologies in IoV to implement latency-sensitive applications. To enable the latency-sensitive applications, Vehicle-to-Vehicle (V2V) communication has emerged as a promising technology for IoV [4]. In this case, the connected vehicles enable a cooperative computation pattern to implement the intensive tasks in traffic scenarios. Vehicles tend to share their spare resources over proximate peer-to-peer links to implement those low latency services [5]. In this situation, IoV enables a flexible resource sharing market for vehicles to trade their spare computation and spectrum resources.

Nevertheless, because of the high mobility and variability of vehicular networks, it is difficult to maintain long-term stable connections among vehicles, and the trading partners are often strange vehicles [6]. The provision of security and privacy is very challenging during the resource sharing process [7]. On the one hand, resource payers may repudiate the transaction and refuse to pay the reward. On the other hand, resource payees may be dishonest during the trading process and do not provide the resource as their promise. In this case, how to effectively realize the trustworthiness of vehicles during resource sharing process is a big challenge in vehicular networks.

The associate editor coordinating the review of this manuscript and approving it for publication was Shui Yu.

The trust management of vehicles is usually implemented by a centralized entity using rating mechanisms in the existing research work. The entity calculates the rating scores and provides network operators with a reward and punishment benchmark for vehicles. According to the evaluations of vehicles, researchers proposed several rating methods for vehicular systems [8]–[10]. However, there exists significant security and privacy challenges for the centralized methods. The centralized entity is vulnerable to the single point of failure. In addition, the rating mechanisms relying on the sensitive information of vehicles, such as identity information and trading performance, which would result in the disclosure of privacy.

Recently, emerging Blockchain technology is considered as a natural tool to tackle the trust and privacy issues due to its encrypted and tamper-proof ledger. Blockchain technology is initially leveraged to enable trade peers without a third party [11]. Regarding its decentralization and highly security, blockchain is migrated to the application of IoV scenarios, which implement the reliable and trustworthy data management during information sharing and resource trading processes [12]–[14]. Nevertheless, most of the blockchain systems rely on exorbitant computing power (*e.g.* Proof-of-Work) or broadcasting of verifying results (*e.g.* Practical Byzantine Fault Tolerance). These blockchain systems are inapplicable for practical IoV scenarios, since the computing resource of vehicles is limited and the communication between vehicles is unstable.

In order to address the aforementioned problems, we propose a lightweight blockchain-based resource sharing scheme. Considering the dynamic characteristic of resource availability and vehicles location, we further propose a Deep Reinforcement Learning (DRL)-based resource pricing scheme through smart contract technology. The proposed blockchain system provides an efficient platform for the trust management, and a reputation-based consensus mechanism is proposed to replace the mining process. The main contributions of this paper are threefold as follows,

- Unlike existing public chain architecture, we present a trust management paradigm for resource sharing in IoV based on consortium blockchian. The proposed architecture separates trading vehicles and block publishers, so that it can effectively reduce the communication cost of publishing blocks while maintaining the privacy of trading vehicles.
- We develop a DRL-based smart contract scheme to match the supply and demand during resource sharing process. The contract is included in the blockchain system as a script to realize trading flexibility and programmable society, which is beneficial for the vehicle-centric trading pattern.
- We originally propose a Proof-of-Reputation (PoR) consensus mechanism to construct a lightweight blockchain, in which the RSU can publish the block only if it has collected the highest sum reputation of transactions. This is inspired by the Ethereum *gas* setting [15], which can spur vehicles to perform good behaviours.

The remainder of this paper is organized as follows. We present related works of trust management in Section II. The system model is proposed in Section III. In Section IV, a DRL-based smart contract is designed for the sharing process. Section V analyses the communication and security performance of proposed PoR protocol, followed by the simulation in Section VI. Finally, we conclude this paper in Section VII.

II. RELATED WORKS

Recent research within this field has focused on how to establish trust among vehicles during the cooperation process. In general, the methods of trust management of vehicles can be broadly classified into two categories, namely non-blockchain trust management and blockchain-based trust management.

A. NON-BLOCKCHAIN TRUST MANAGEMENT

Non-blockchain trust management in IoV has been widely studied. Existing works leverage the reputation value to evaluate the trust degree of vehicles. The rating methods can be further divided into two types: infrastructure-based and self-organized [16]. In the Infrastructure-based approach, the trust rating process is executed at trusted authorities by managing the certificates of vehicles. However, due to the high mobility of vehicles, they are continuously confronted with other unknown vehicles, which are registered with other trusted authorities in other regions [17]. Therefore, the infrastructure-based approach may be inapplicable in IoV scenarios. In the self-organized approach, the rating process is executed among vehicles through their interactions. Three-Valued Subjective Logic is utilized to calculate the trust value according to the quality and quantity of the vehicular interactions [8]. The authors of [9], [10] conduct the trust assessment through direct and indirect trust evaluations. The direct trust evaluation is based on the own vehicle and the indirect trust evaluation is based on environment. The proposed evaluation methods make sure the reliable of message receivers. In [18], a reputation threshold is designed, in which vehicles cannot participate in the message sharing process if their reputation scores are less than a certain threshold. Thus, the trustiness is guaranteed from the message sources. However, non-blockchain trust management cannot maintain the privacy when it encounters the resource trading scenarios in IoV.

B. BLOCKCHAIN-BASED TRUST MANAGEMENT

In order to cope with the problems of non-blockchain methods in resource trading scenarios, blockchain technology is introduced for trust management. The transactions among the trading process are recorded and consented by all the peers in the networks. Any one that has fraudulent behaviours will be recorded in the ledger and cannot be modified. The transactions are packed into blocks and chained with

TABLE 1. Notions.

variable	The definition
X_i	The amount of task need to be offloaded from the TO_i
Loc_i,Loc_j	The location of TO_i and RP_j vehicle
Rep_j^i	The Reputation value of TO_i against RP_j
CR_j	The available computing resource of RP_j
SR_j	The available spectrum resource of the RP_j
ω_j^i	The amount of allocated spectrum resource from TO_i to RP_j
η_j^i , ρ_j^i	The choice variables indicating if TO_i purchase computing and spectrum resources from RP_j
$\triangle t$	The length of one time slot
K	The total number of BNs
N_I	The number of interfering BNs
N_z^t	The number of collected transactions of BN_z at t moment
Ω	The blocksize threshold
λ_t	The resource trading frequency
λ_b	The BN density
p_j^i	The differentiated resource price of TO_i against RP_j
sr_j^i	The purchased spectrum resource by TO_i from RP_j
cr_j^i	The purchased computing resource by TO_i from RP_j
x_j^i	The amount of allocated task from TO_i to RP_j

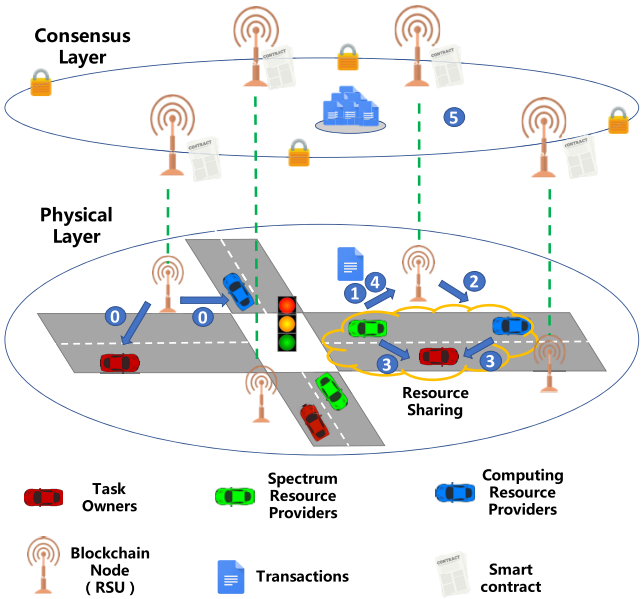

FIGURE 1. System model.

tampered-resistant Hash values, in order to maintain the privacy of trading peers [19], [20]. Due to its decentralization and highly security, blockchain has been widely studied in IoV scenarios. By integrating both Proof-of-Work (PoW) and Proof-of-Stake consensus mechanisms, trust value offset schemes are leveraged to reduce the computing cost [12], [13]. The authors in [14] propose a reputation-based delegated Byzantine fault tolerance consensus algorithm to maintain the trustiness of electric vehicles. In [21], the authors propose a CreditCoin vehicular announcement network under blockchain architecture. The CreditCoin can not only motivate users to share traffic information, but also maintain a trustworthy interaction environment in IoV scenarios.

Nevertheless, the consensus mechanisms of the above works rely on computing power and multiple interaction time, which is inappropriate for the high topology-variability networks of IoV. Accordingly, we propose a lightweight consensus mechanism for resource sharing in this paper, a sum reputation method is discussed as the extension of our previous work [22]. The trust is built and the effectiveness is proved by security analysis.

III. SYSTEM MODEL

In this section, we present the system model which includes the blockchain-enabled resource sharing model and the dynamic resource pricing model. Table 1 presents the main notions adopted in the following part of this work.

A. BLOCKCHAIN-ENABLED RESOURCE SHARING MODEL

As shown in Fig.1, we consider a blockchain-enabled resource sharing scenario in IoV network, of which the vehicles can be divided into two categories, Task Owners (TOs) and Resource Providers (RPs). TOs intend to offload their computation tasks to an adjacent RP to implement cooperative computation. A resource trading market is investigated for potential sellers and buyers, *i.e.*, RPs and TOs. To compensate RPs whenever they provide resources for an offloadable task, a digital cryptographic currency named Resource Coins (RCs) is introduced. In addition, RPs have two further divisions, one is computation resource providers (CRPs) which provide Computing Resource (CR) to execute the task of TOs, and the other is spectrum resource providers (SRPs) which provide spectrum resource (SR) to assist the task transmission.

Different from traditional resource sharing systems, where TOs often treat RPs as trustworthy parties [23], our proposed blockchain-enabled resource sharing system assumes that TOs and RPs cannot fully trust with each other. Thus, it is crucial to tackle the trust issue during the resource sharing process. We propose the Consortium Blockchain architecture to overcome the problem. The reason why we choose the consortium blockchain rather than the public chain is that totally decentralized public chain cannot supply the large-scale IoV network, considering the high mobility of vehicles. When every vehicle needs to maintain one global ledger, the maintenance cost is huge. Therefore, we adopt consortium blockchain and assign a type of specific authorized nodes to reach a fast consensus with moderate cost.

The interactions among vehicles are described as *transaction*s and recorded by every Blockchain Nodes (BNs) in networks. BNs can be RoadSide Units (RSUs), which are responsible for collecting transactions from the vehicles and generating blocks. Diversified communication models make it difficult to directly utilize conventional public blockchain architecture on the vehicular system. Therefore, we adopt the consortium blockchain architecture in the proposed resource sharing system, where a specific type of node BN is responsible for publishing blocks. Vehicles are only responsible for trading resources with other vehicles. Note that in consortium

blockchain architecture, as the consensus procedure is executed among BNs, vehicles are only able to generate transactions and have no right to participant in the consensus procedure.

The general format of transaction can be illustrated as TX : $\{from\|value\|data\|to\}$, the terms *from* and *to* represent the wallet address of blockchain entities, which include senders, receivers and smart contract accounts. The term *value* denotes the transferred RCs and the term *data* denotes the certain parameters during the resource sharing procedure. The basic workflow of Blockchain-Enabled resource sharing in IoV networks consists the following six steps:

Step 0: System initialization: Each vehicle becomes a legitimate entity after it is registered on a trusted party such as macro-base station or RSU. A pair of keys including public key PL and private key PK are sent to the vehicle, together with an initialized reputation value (*Rep*). Using key PK, the vehicle generates several wallet accounts WAs to realize transaction with others.

Step 1: Uploading demands and supplies: Each vehicle determines its role according to its service requirement and resource availability. Specifically, TOs send the amount of computation task to the BNs and RPs send the amount of available resources to the BNs. The sending messages are encapsulated as transactions TX_{req} and TX_{pro} to trigger the smart contracts (*STs*) embedded in the blockchain system. After activating the contract accounts, the dynamic match game of resource demand and supply will be automatically implemented within a predefined time. Differentiated prices and resources allocation schemes will be obtained. The general format of TX_{req} and TX_{pro} are shown as,

$$\begin{aligned} TX_{req} &: \{WA_{TO}\|0\|X, Loc, Rep\|WA_{ST}\} \\ TX_{pro} &: \{WA_{RP}\|0\|CR, SR, Loc, Rep\|WA_{ST}\}, \end{aligned} \tag{1}$$

where *Loc* indicates current location of the vehicles to help the contracts measure channel condition, and WA_{ST} is the address of contract.

Step 2: Triggering smart contract: Upon receiving TX_{req} and TX_{pro}, STs decide the prices vector $\mathbf{p}$ and the trading amount vectors of spectrum and computing resources $\{\mathbf{sr}, \mathbf{cr}\}$ according to their reputation value, task amount and supplement amount. Meanwhile, the specific assignment scheme of offloading $\{\mathbf{x}, \boldsymbol{\omega}\}$ would also be settled, where $\mathbf{x}$ denotes the task assignment vector, and $\boldsymbol{\omega}$ is the spectrum assignment vector. The contract aims to maximize the utility of both TOs and RPs while maintaining the latency requirement. The joint optimization problem can be well solved by means of existing heuristic algorithm [24] with perfect priori-knowledge of environment information. However, the information of task arrival and network state is difficult to be acquired in advance. In addition, the mobility of vehicles makes the resource trading a dynamic process. In order to address the problem, we leverage a learning framework to assist STs to obtain the optimal policies of resource pricing and task assignment, which would be elaborated in Section IV.

Step 3: Resource trading and computation offloading: After triggering smart contract, both TOs and RPs get the price and offloading assignment $\{\mathbf{p}, \mathbf{cr}, \mathbf{sr}, \mathbf{x}, \boldsymbol{\omega}\}$. TOs would buy resources from both CRPs and SRPs according to the resource price $\mathbf{p}$ and trading amount $\{\mathbf{cr}, \mathbf{sr}\}$. Then TOs will utilize purchased resources to implement latency-driven task offloading process. During the task offloading process, TOs transmit $\mathbf{x}$ tasks by using $\boldsymbol{\omega}$ spectrum resource to specific CRPs.

Step 4: Transaction production: Once completing offloaded tasks, RPs return computation results back to TOs for the inspecting of integrity and correctness. Afterwards, TOs create trading transaction TX_c with their signatures including $\mathbf{p}$*($\mathbf{cr}$+$\mathbf{sr}$) RCs, reputation **Rep** and judgement of computation results **jud**. For a certain TO, the format of generated TX_c is shown as,

$$TX_c : \{WA_{TO}\|p*(cr+sr)\|jud, Rep\|WA_{RP}\}. \tag{2}$$

Step 5: PoR Consensus Process: As the core section of blockchain system, conventional consensus mechanisms rely on the computing power to solve hash puzzles and waste huge computing resource. In this article, we propose a lightweight consensus mechanism named Proof-of-Reputation (PoR) based on the sum reputation of collected transactions. We refer to a fix block size model [25] with threshold Ω. When one BN collects Ω transactions, it will broadcast the block to the network for consensus. After consensus procedure, the block is recorded in the ledger and the reputation value of the trading vehicles are updated according to their performance during the resource trading process. The detail designs and performance evaluation of PoR will be discussed in section V.

B. DYNAMIC RESOURCE PRICING MODEL

To illustrate the dynamic feature of computation task generation and provision of spare resources, the system operation time is divided into several discrete time intervals composed of equal length $\triangle t$. Therefore, the t-th slot can be described as $(t, t+\triangle t)$.

At time slot t, suppose that there are N TOs and M RPs, where TO_i $(i \in N)$ has $X_i(t)$ computation tasks to be completed, RP_j $(j \in M)$ has $CR_j(t)$ and $SR_j(t)$ spare resources. According to the computation task requirement $X_i(t)$, TO_i purchases spare resources from ambient vehicles, the V2V resource sharing process can be implemented with respect to their reputation value $Rep_j^i(t)$ and resource prices $p_j^i(t)$. We set $\eta_j^i(t) \in \{0,1\}$ and $\rho_j^i(t) \in \{0,1\}$ as the choice variables indicating that if TO_i purchases computing resource and spectrum resource from RP_j, respectively. Therefore, the buying cost of TO_i at t-th slot can be described as

$$Cost_i(t) = -\sum_{j=1}^{M} \frac{p_j^i(t)}{Rep_j^i(t)} [\eta_j^i(t) cr_j^i(t) + \rho_j^i(t) sr_j^i(t)], \tag{3}$$

where $p_j^i(t)$ denotes the differentiated price of resources and $Rep_j^i(t)$ is the reputation value of TO_i towards RP_j at slot t.

We define the reputation value to indicate that an entity with a high reputation value would acquire minor cost when it purchases resources and vice versa. $sr_j^i(t)$ and $cr_k^i(t)$ are the amount of spectrum and computing resources purchased by TO_i from RP_j which satisfies,

$$\sum_{i=1}^{N} sr_j^i(t) \le SR_j(t), \quad \sum_{i=1}^{N} cr_j^i(t) \le CR_j(t), \tag{4}$$

where $SR_j(t)$ and $CR_j(t)$ are the available spectrum and computing resources of RP_j at slot t. Similarly, the selling revenue of RP_j at slot t can be described as,

$$Rev_j(t) = \sum_{i=1}^{N} [p_j^i(t) Rep_j^i(t) - \epsilon][\eta_j^i(t) cr_j^i(t) + \rho_j^i(t) sr_j^i(t)], \tag{5}$$

where ϵ denotes the cost parameters of resource provision. Regarding that the cost and revenue functions are related to different states of SR_j, CR_j and Rep, the pricing scheme for resource is a dynamic procedure in the long term.

IV. PROBLEM FORMULATION FOR DRL-BASED SMART CONTRACT

A. LATENCY-DRIVEN TASK ASSIGNMENT

During the resource pricing process, it can be figured that if TO_i chooses RP_j as a computing resource seller, it must assign the computation tasks to RP_j. Therefore, the task assignment process is coupled with dynamic resource pricing procedure. Regarding the selection of resource providers, the total task execution latency includes the task transmission latency and the task computation latency.

1) TASK TRANSMISSION LATENCY

In the proposed system, *TO*s utilize purchased spectrum resource to transfer the computation tasks to the selected *CRP*s. The V2V communication link between *TO*s and *CRP*s is achieved through Frequency Division Multiple Access (FDMA) technology (as in LTE-A systems), in which spectrum transmission relies on the transfer of Physical Resource Blocks (PRBs). In this paper, we assume that the spectrum trading process is the occupation of PRBs. A vehicle with more spectrum resource means a higher transmission rate. Thus, the task transmission latency from TO_i to CRP_j is shown as,

$$\begin{cases} l_{i,j}^{v2v}(t) = \dfrac{x_j^i(t)}{\omega_j^i(t) \log(1 + SNR_j^i(t))} \\ SNR_j^i(t) = \dfrac{h_j^i(t) P}{\sigma^2}, \end{cases} \tag{6}$$

where $h_j^i(t)$ is channel power gain depending on the distance between TO_i and CRP_j, *i.e.*, $|Loc_i(t) - Loc_j(t)|$. P is the transmission power and $x_j^i(t)$ denotes the allocated computing task to the CPRs and $\omega_j^i(t)$ denotes the allocated spectrum resource from TO_i to CRP_j which satisfies,

$$\sum_{j=1}^{M} \omega_j^i(t) \le \sum_{j=1}^{M} \rho_j^i(t) sr_j^i(t), \quad \forall i \in N. \tag{7}$$

2) TASK COMPUTING LATENCY

The computation latency at distributed *CRP*s depends on the amount of computation tasks and the clock-rate of CPU. We assume that one bit input data requires U CPU cycles [26]. Under the assignment of computation tasks $x_j^i(t)$, the computing latency of task offloading will be

$$l_{i,j}^{com}(t) = \begin{cases} \dfrac{x_j^i(t) U}{\eta_j^i(t) cr_j^i(t)}, & if\ \eta_j^i(t) = 1 \\ 0, & if\ \eta_j^i(t) = 0. \end{cases} \tag{8}$$

In the distributed V2V architecture, the total service latency is determined by the longest execution time of assigned computation tasks. Comparing to the input size of assigned tasks, the output size is far less than that of input. Thus, the return time of computation results can be omitted [27]. The total latency of task assignment is $l_i^{tot} = max(l_{i,j}^{v2v} + l_{i,j}^{com}), \forall j \in M$.

We consider a latency-constraint scenario where the computation results of tasks must be returned within $\triangle t$ epoch. If the original tasks $X_i(t)$ are too large to be executed within the current time slot, TO_i would choose to split the tasks and send partial tasks to ambient *CRP*s, which can be described as,

$$\begin{cases} \sum_{j=1}^{M} x_j^i(t) \eta_j^i(t) \le X_i(t) \\ l_i^{tot} \le \triangle t, \end{cases} \quad \forall i \in N. \tag{9}$$

Though TO_i has no priori-knowledge about the arrival amount of computation tasks, it is reasonable to consider that the amount of tasks can be observed at the beginning of each time slot. Consequently, the total computation tasks to be offloaded include the leftovers in the last time epoch and the newly generated tasks X_i^{ari}, as shown below,

$$X_i(t+1) = X_i^{ari}(t+1) + [X_i(t) - \sum_{j=1}^{M} x_j^i(t) \eta_j^i(t)]. \tag{10}$$

B. REPUTATION VALUE

In this paper, the reputation value (*Rep*) is the indicator of trust degree of vehicles during the resource sharing process. Moreover, *Rep* is the key element in the proposed consensus mechanism. In our previous work [22], a linear model is proposed to depict *Rep*. However, in the linear model, the peer with the highest *Rep* tends to have the highest *Rep* in the next moment. Thus, this will result in sustained growing of *Rep* and monopoly of resources. Furthermore, the punishment to dishonest devices is not reflected on the model. In this case, we refer to Gompertz function to calculate *Rep* to model the

concept of trust in vehicles interactions. The Gompertz function is often used to reflect the regular pattern of population growth, defined as $Rep(t) = a \times e^{-b \times e^{-c \times r(t)}}$. The variable t denotes the time slot, a, b and c are the function parameters, and r_t is the input of the function.

The *Rep* function consists three phases, firstly slopes gently then increases and finally stay stably. The characteristic can well fit reputation feature in reality that reputation begins with doubting phase, then go through growing phase and can be widely recognized at last. In this paper, we design *Rep* as the accumulation of historical reputation records, which can be described as $r(t) = \tau r(t-1) + r^*(t)$, where $r^*(t)$ is the current reputation value, $\tau \in (0, 1)$ accounts for the fact that recent reputation value is more relevant than the past. Regarding the existence of dishonest *RP*s that do not dedicate to the task as their promise, the current reputation value $r^*(t)$ can be expressed as,

$$r^*_{i,j}(t) = \begin{cases} \dfrac{\eta^i_j(t) cr^i_j(t) + \rho^i_j(t) sr^i_j(t)}{\sum_{j=1}^{M} [\eta^i_j(t) cr^i_j(t) + \rho^i_j(t) sr^i_j(t)]} & honest \\ -1 & dishonest. \end{cases} \tag{11}$$

The honesty of *RP*s relies on the *jud* term of TX_c. The judgement includes the verification of computation correctness and integrity [28], [29]. These are applications dependent and out of the scope of this article, so it will not be described in detailed.

C. JOINT OPTIMIZATION PROBLEM IN SMART CONTRACT

After receiving TX_{req} and TX_{pro}, the joint optimization problem is automatically executed in *ST*s. Our objective aims to maximize the total revenue U^{tot} of both *RP*s and *TO*s under delay constraints. We use a five tuple $v(t)$ to represent the allocation variables during the resource sharing process for the simplicity of analysis, *i.e.*, $v(t) = \{p(t), cr(t), sr(t), x(t), \boldsymbol{\omega}(t)\}$. Therefore, the latency-driven resource trading and task offloading problem can be described as,

$$\begin{aligned} \max_{\boldsymbol{\eta}(t), \boldsymbol{\rho}(t), v(t)} \quad & U^{tot}(t) = \alpha \sum_{i=1}^{N} Cost_i(t) + \kappa \sum_{j=1}^{M} Rev_j(t) \\ s.t.\ C1:\ & \eta^i_j, \rho^i_j \in \{0, 1\} \quad \forall i \in N, j \in M \\ C2:\ & p_{min} \le p_j \le p_{max} \quad \forall j \in M \\ C3:\ & 0 \le R_j \le R_{max} \quad \forall j \in M \\ C4:\ & (4), (7), (9) \quad \forall i \in N,\ j \in M \end{aligned} \tag{12}$$

where α and κ represent the control parameters, and the control parameters represent the weight relationship between buying cost and selling revenue. $v(t)$ can be derived via the perfect system information. In practical, we have no priori-knowledge of arrival distribution of computation tasks, channel state and vehicles locations. In addition, considering that the resource trading process is a game over long period of time, our goal of the proposed system is to optimize the cumulative and expected utility from a long-term perspective.

We formulate the joint resource trading and task offloading process as an infinite discounted continuous state Markov Decision Process (MDP) problem. The system components of MDP contain:

1) STATE SPACE S

The system state at the tth slot is denoted as $s(t) \in S$. The state $s(t)$ is a joint state space, which includes the coordinates of vehicles $Loc(t)$, the computation tasks $X(t)$, the spare spectrum resource $SR(t)$, the spare computing resource $CR(t)$, and the reputation value $Rep(t)$. Therefore, the system state can be described as a five tuple $s(t) = [Loc(t), X(t), SR(t), CR(t), Rep(t)]$.

2) ACTION SPACE A

At the tth slot, the action $a(t) \in A$ is a series decisions of the system state $s(t)$. The decisions include resource selection indicators $\boldsymbol{\eta}(t)$, $\boldsymbol{\rho}(t)$, the differentiated resources prices $p(t)$, the task assignment $x(t)$, the purchase amount of resources $cr(t)$, $sr(t)$ and the bandwidth allocation $\boldsymbol{\omega}(t)$. Thus, the action in each decision epoch can be described as $a(t) = [\boldsymbol{\eta}(t), \boldsymbol{\rho}(t), v(t)]$.

3) STATE TRANSITION

Given the current state $s(t)$ and the action $a(t)$, the system state transition probability to the next state $s(t+1)$ is denoted as $p(s(t+1)|s(t), a(t))$.

4) REWARD FUNCTION

To perform effective computation offloading, the reward function will aim to maximize the utility of both task owners and resource providers. Accordingly, the system reward is denoted as $R(s(t), a(t))$, which is shown as

$$R(s(t), a(t)) = U^{tot}(t). \tag{13}$$

5) POLICY

We define the policy as $\pi : S \to A$, realizing the mapping from current state to the actions, *e.g.*, $a = \pi(s)$. We denote Π as the set of all policies. Given the initial state $s(0)$ and corresponding policy $\pi \in \Pi$, the expected discounted long-term system profit is defined as

$$V^{\pi}(s(0)) = \mathbb{E}[\sum_{t=0}^{\infty} \gamma^t R(s(t), a(t))|s(0)], \tag{14}$$

where $\gamma \in (0, 1)$ denotes the discount parameter. Consequently, the objective of the paper is to find the optimal policy π^* to maximize the long-term system profit, *i.e.*

$$\pi^* = arg \max_{\Pi} V^{\pi}(s(0)). \quad s(0) \in S \tag{15}$$

D. DEEP REINFORCEMENT LEARNING APPROACH FOR DECISION MAKING

The optimal pricing scheme and resource allocation policy could be obtained according to Bellman equation, we omit index (t) and set s' as the next state for the sake of simplicity.

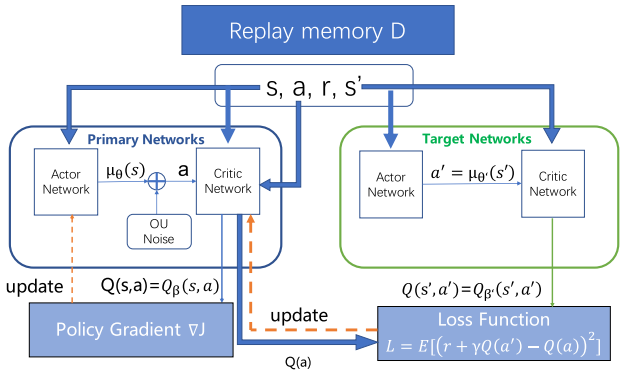

FIGURE 2. DDPG framework.

Therefore, the expected optimal reward $V^*(s)$ and optimal policy $\pi^*(s)$ at s are shown as

$$\begin{cases} V^*(s) = \max\limits_{a\in A}\{R(s,a) + \gamma \sum\limits_{s'\in S} p(s'|s,a)V^*(s')\} \\ \pi^*(s = arg \max\limits_{a\in A}\{R(s,a) + \gamma \sum\limits_{s'\in S} p(s'|s,a)V^*(s')\}. \end{cases} \tag{16}$$

As the environment state cannot be acquired in a prior manner, several works focus on the model-free reinforcement learning methods to obtain the optimal policy π^*, such as Q-learning and Sarsa, etc. [30]. Q-learning adopts a state-action function $Q(s,a) = q(s,a) + \gamma \sum_{s'\in S} p(s'|s,a)V^*(s')$. The Q-value is considered as the expected reward for taking action a at state s. Furthermore, equation (16) can be equivalent to the following

$$\begin{cases} V^*(s) = \max\limits_{a\in A} Q^*(s,a) \\ \pi^*(s) = arg \max\limits_{a\in A} Q^*(s,a), \end{cases} \tag{17}$$

where $Q^*(s,a)$ is the optimal Q-value of the state-action pair (s, a). Through constructing a two-dimension Q-table, the Q-learning algorithm can learn the optimal policy at any observed system state. The updated Q-function can be expressed as

$$Q(s,a) \leftarrow Q(s,a) + \Upsilon[R(s,a) + \gamma \max_{a'\in A} Q(s',a') - Q(s,a)], \tag{18}$$

where Υ is the learning rate. The Q-learning algorithm is subject to the curse of dimensionality, as the computation complexity will increase largely with the growth of the number of state-action pairs. Furthermore, the algorithm typically assumes the discrete action space to simplify the action search. In our scenario, however, the pricing and resource allocation are continuous numbers. The continuous action space will lead to a large number of iterations. For this case, deterministic policy gradient algorithm is developed recently that allows to directly learn the policy $\mu(s)$ instead of the policy distribution $\pi(s|a)$ [31]. Therefore, we implement Deep Deterministic Policy Gradient (DDPG) algorithm to obtain the optimal pricing scheme and resource allocation policy.

DDPG explicits the deterministic actor-critic (DAC) model, which learns a parametrized deterministic policy function $\theta_\mu : S \rightarrow A$ to perform the resource sharing scheme.

Algorithm 1 DDPG-based Joint Decision Algorithm

Input: $\mathbb{D}$, K,,T, l, γ, σ, s_0
initialize replay memory $\mathbb{D}$ to capacity D^*, $\mathbb{D}(K) = \emptyset$;
initialize networks θ_μ and θ_Q with random weight, target networks $\theta_{\mu'} = \theta_\mu$ and $\theta_{Q'} = \theta_Q$;
for *episode* $= 0 \rightarrow T - 1$ **do**
  initialize V2V resource sharing scenario, ST receives TX_{req} and TX_{pro} as the state $s_0 = \{X, CR, SR, Loc, Rep\}$;
  for $t = 0 \rightarrow l - 1$ **do**
    perform action $a = \theta_\mu + N_t$;
    execute action a, *observe* reward r and s' based on (16), (12) and (13);
    Store transition (s, a, r, s') in $\mathbb{D}$;
    if *replay memory* $\mathbb{D}$ *is full* (D^*) **then**
      sample a random batch of K transitions (s, a, r, s') from $\mathbb{D}$;
      Set $\triangle_t = r + \upsilon Q(s', a'|\theta_Q)$;
      Update the parameter of critic network θ_Q by minimizing the loss: $L = E^2[\triangle_t - Q(s, a|\theta_{Q'})]$;
      Update the parameter of actor network θ_μ by using the sampled policy gradient in (24) ;
      Update target network $\theta_{\mu'}$ and $\theta_{Q'}$ using (25);
return θ_μ and θ_Q;

DDPG also learns another critic Q-function $\theta_Q : S \times A \rightarrow R$ to evaluate the action policy. A Neural Network (NN) is embedded in actor and critic network to fit the model of θ_μ and θ_Q. The gradient descent method is used to train θ_μ and θ_Q over minibatch data that randomly sampled from the replay buffer $\mathbb{D}$. The training of minibatch is called memory replay procedure.

As shown in Figure 2, the learning framework of DDPG allows the actor and critic networks to learn according to their respective objective function, contributing to a more convergent learning process. The rule of updating of critic network refers to the *Loss Function*, which is shown in the figure. The gradient of actor network is computed as follows

$$\begin{aligned} \nabla_{\theta_\mu} J &\approx \frac{1}{|\mathbb{D}|} \sum_t [\nabla_{\theta_\mu} Q(s, a|\theta_Q)|_{s,a=\mu(s|\theta_\mu)}] \\ &= \frac{1}{|\mathbb{D}|} \sum_t [\nabla_a Q(s, a|\theta_Q)|_{s,a=\mu(s)} \nabla_{\theta_\mu} \mu(s|\theta_\mu)|_s]. \end{aligned} \tag{19}$$

To maintain the stable updating of $Q_{\theta_Q}(s,a)$, a dual network with evaluation of target network is further proposed in DDPG. Similar with the structure in [32], the target networks are updated using the soft updates as follows

$$\begin{cases} \theta_{\mu'} \leftarrow \sigma\theta_\mu + (1-\sigma)\theta_{\mu'} \\ \theta_{Q'} \leftarrow \sigma\theta_Q + (1-\sigma)\theta_{Q'}, \end{cases} \tag{20}$$

where σ controls the updating amplitude. Higher σ means faster updating speed and less stability. The detail of

the DDPG-based joint decision algorithm is shown in Algorithm 1. The computational complexity of the proposed algorithm is mainly decided from Line 3 to Line 14 in Algorithm 1. The algorithm starts up the loop in Line 3 and Line 5, which leads to the $O(lT)$ time complexity. According to [33], the time complexity of sampling and updating from Line 10 to Line 14 is $O(logN_{tree})$, where N_{tree} is the number of nodes in the sum-tree used to implement the priority probability. Therefore, the total time complexity of Algorithm 1 is $O(lTlogN_{tree})$.

V. SYSTEM ANALYSIS OF PROOF-OF-REPUTATION

In this section, the details of the proposed PoR consensus mechanism will be discussed. In order to prove the effectiveness of the proposed mechanism, we further analyse the communication and security performance.

A. OVERVIEW OF POR DESIGN

In conventional blockchain systems, peers reach the consensus through contributing a great number of computing power, which will cause the waste of energy. Our previous work proposed a reputation-based consensus mechanism of which the blocks publishing do not rely on mining procedure. Thus, it can effectively reduce the power expenses. However, we adopt a linear model for reputation value in [22]. This will cause that a peer can always publish a block if it has the highest reputation value, which violates system fairness. Regarding of this, we propose Proof-of-Reputation mechanism to maintain the fast and lightweight blockchain meanwhile ensuring the fairness among vehicles in IoV networks. The basic procedure of the PoR mechanism includes following three steps:

1) COLLECTING TRANSACTIONS

The BNs continuously collect transactions, including TX_{req}, TX_{Pro} and TX_c. These transactions are sent by vehicles during their resource sharing process. As the transactions are signed with vehicular signatures, BNs cannot forge the transactions by their own, and the fake station issue can be avoided.

2) GENERATING BLOCKS AND CALCULATING REPUTATION

In the PoR mechanism, there is no need for vehicles to calculate the hash puzzles. Therefore, how to generate a new block is a crucial issue in the proposed consensus mechanism. In this case, we design the block with a fixed size Ω. The BNs can generate new blocks only if they have collected Ω transactions. Regarding that the proposed trading system is divided into several time slots, we further stipulate that the generating time of blocks is always at the end of each slot.

Considering that each transaction is corresponding to a reputation value of certain vehicle, it is apparently that each block has a sum reputation (SR) consisting of Ω transactions. The format of new block is shown as Figure 3. We set that the collected transactions in one block must obey a chronological order. This can be realized through the timestamp technology.

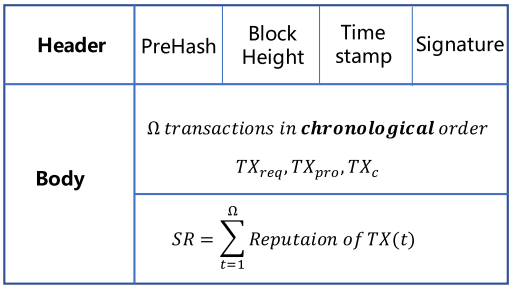

FIGURE 3. The format of block.

The chronological design can guards against the malicious BNs, which will be elaborated in security performance.

3) VERIFYING AND REACHING CONSENSUS

After receiving a new block, the BNs would stop collecting transactions and inspect the SR value of the received block to check if the block has the highest SR. Meanwhile, the integrity of computation result and signature can also be checked. After the verifying procedure, the block with the highest SR value would be added into the blockchain ledger. At this point, the consensus procedure is completed and BNs continue to collect the transactions in the network and prepare for a new round.

The sum of reputation in PoR is inspired by the Ethereum *gas* setting where miners in Ethereum would always collect transactions with high *gas* reward. In this paper, BNs would collect those transactions with high reputation value. Consequently, it can spur vehicles to behave in good manners. On the contrary, the vehicles with a bad reputation would suffer a long on-chain latency and reduce the trading experience.

B. SYSTEM ANALYSIS

1) COMMUNICATION PERFORMANCE

In this subsection, we will discuss the communication performance of the proposed PoR scheme in terms of SINR and Transaction Successful Rate (TSR), which refers to the excellent work [34]. As the blocksize in our paper is fixed, the TSR behaves in a different manner regarding to the blocksize Ω and trading frequency λ_t. We will first calculate the PDF of SINR, then the TSR can be derived according to SINR.

The communication model is illustrated as Figure.4. Let all the BNs: $\{BN_0, BN_1, BN_2...BN_K\}$ follow uniform distribution in a circle area with a radius R_0. It is reasonable to assume that BN_c is the center of the circle. Suppose that BN_0 is the block publisher with the highest sum reputation value and broadcasts the block to other BNs. Meanwhile, other BNs can also broadcast their own generated blocks for consensus, which leads to a interference. Consequently, the SINR can be expressed from the perspective of receiver BN_c as follows

$$SINR(D, N_I) = \frac{P * h(d_0)}{\sum_{l=1}^{N_I} P * h(d_l) + \sigma^2}, \quad N_I \leq K, \quad (21)$$

where $D = [d_0, d_1, \ldots, d_{N_I}]$ represents the distance vector from the receiver BN_c to block publishers $[BN_0, BN_2, \ldots, BN_{N_I}]$, N_I is the number of interfering

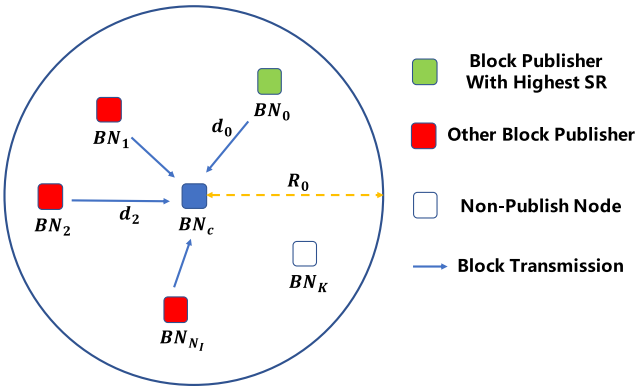

FIGURE 4. Wireless Communication model.

block publishers. The transmission power is denoted as P and channel power gain is $h(d)$.

For convenience, we use capital letters to denote random variables, and the corresponding lowercases to the value of random variables. As SINR is related to distance and interfering numbers, the PDF of SINR can be described as

$$f_{SINR}(D = d, N_I = n_I) = f_{D_0}(d_0) f_{D1,N_I}(d_1, n_I), \quad (22)$$

where d_1 represents the distance vector of interfering BNs. We denote a random variable RX indicating the required time for BNs to generate a block. For our proposed system, the generating time of blocks is at the end of each time slot, BN could publish one block only if it has collected Ω transactions within $\triangle t$. Therefore, the PDF of distance D_0 between BN_0 and BN_c can be shown as,

$$f_{D_0}(d_0) = P\{D_0 = d_0\} * P\{RX \leq \triangle t\}. \quad (23)$$

To obtain the probability of publishing blocks, we further denote a random variable N_t^z, where $z = \{0, 1, \ldots, K\}$ represents the number of collected transactions of BN^z at t moment. In this paper, we suppose that the transaction arrival rate is modelled as Poisson Point Process (PPP) with the density λ_t. Consequently we have following proposition:

Proposition 1: The PDF of X follows Gamma Distribution, *i.e.*

$$f_{RX}(\triangle t) \sim \Gamma(\triangle t; \lambda_t, \Omega). \quad (24)$$

Proof: Denoting $F_{RX}(t)$ as the CDF of RX, considering the arrival number of transaction obeys PPP with parameter λ_t we have

$$\begin{aligned} F_{RX}(\triangle t) &= P\{RX \leq \triangle t\} = P\{N_t \geq \Omega\} \\ &= \sum_{y=\Omega}^{\infty} \frac{(\lambda_t \triangle t)^y}{y!} e^{-\lambda_t} = 1 - \sum_{y=0}^{\Omega-1} \frac{(\lambda_t \triangle t)^y}{y!} e^{-\lambda_t}. \end{aligned} \quad (25)$$

Hence, based on (30), the PDF of X can be described

$$\begin{aligned} f_{RX}(\triangle t) &= \frac{dF_{RX}(\triangle t)}{d\triangle t} \\ &= -[-\lambda_t e^{-\lambda_t \triangle t} \sum_{y=0}^{\Omega-1} \frac{(\lambda_t \triangle t)^y}{y!} \\ &\quad + \lambda_t e^{-\lambda_t \triangle t} \sum_{y=1}^{\Omega-1} \frac{(\lambda_t \triangle t)^{y-1}}{(y-1)!}] \\ &= \lambda_t e^{-\lambda_t \triangle t} \frac{(\lambda_t \triangle t)^{\Omega-1}}{(\Omega-1)!} = \frac{\lambda_t^{\Omega} e^{-\lambda_t \triangle t}}{\Gamma(\Omega)} \triangle t^{\Omega-1} \\ &\sim \Gamma(\triangle t; \lambda_t, \Omega), \end{aligned} \quad (26)$$

where $\Gamma(\Omega) = (\Omega - 1)!$ is the Gamma function. ■

As the distribution of BNs follows a uniform distribution, the CDF of the distance between BN_0 and BN_c can be denoted as $F_{D_0}(d_0) = P\{D_0 \leq d_0\} = \pi d_0^2 / \pi R_0^2$. Consequently, the PDF of D_0 is given by

$$P\{D_0 = d_0\} = F'_{D_0}(d_0) = \frac{2d_0}{R_0^2}. \quad (27)$$

For the interfering BNs, we can figure that only those BNs that have collected Ω transactions are regarded as the interfering nodes. Thus, the joint PDF $f_{D_1,N_I}(d_1, n_I)$ can be divided into two parts according to Bayes' theorem

$$f_{D_1,N_I}(d_1, n_I) = f_{N_I}(n_I) * f_{D_1|N_I}(d_1|n_I). \quad (28)$$

The PDF of the interfering number of N_I can be derived according to Total Probability Theorem:

$$f_{N_I}(n_I) = \sum_{a=n_I}^{K} P\{N_I = n_I | A = a\} P\{A = a\}, \quad (29)$$

where A is the variable indicating the total number of BNs in the network. As the transactions are transmitted by vehicles and broadcasting in the network, it is practical to assume that the transaction arrival rate is same for all BNs. Consequently, the transaction collecting process can be regarded as a Bernoulli process and the probability of $N_I = n_I$ at t moment can be calculated for a given BNs number a:

$$\begin{aligned} &P\{N_I = n_I | A = a\} \\ &\quad = C_a^{n_I} [\Gamma(\triangle t; \lambda_t, \Omega)^{n_I}][1 - \Gamma(\triangle t; \lambda_t, \Omega)]^{a-n_I}, \end{aligned} \quad (30)$$

where $C_a^{n_I}$ is the combination number. The number of BNs can also regarded as a Poisson Point Process with the density of λ_b. Thus, the probability of BNs number at t moment can be calculated as

$$P\{A = a\} = \frac{(\pi R_0^2 \lambda_b)^a}{a!} e^{-\pi R_0^2 \lambda_b}. \quad (31)$$

Due to the uniform distribution of BNs, the PDF of distance between interfering BN and BN_c can be shown as $f_{D_0}(d_0) = f_{D_1}(d_1) = \ldots = f_{D_{N_I}}(d_{n_I})$. Given BNs number A and interfering number n_I, the conditional probability of interfering distance is $f_{D_1|N_I}(d_1|n_I) = \prod_{l=1}^{N_I} (2d^l / R_0^2)$.

From (22) and (31), the PDF of SINR can be obtained. According to the definition of Transaction Successful Rate, the new block is regarded as valid only if SINR is greater than a certain threshold, *i.e.* SINR $\geq \beta$. By referring to the close-form expression of the probability of Transmission Successful Rate (TSR) [34], we will directly give the expression of

TSR with respect to our scenario:

$$P(SINR \geq \beta) = \int_0^{R_0} f_{D_0}(d_0)\Phi(\xi(d_0))d(d_0), \quad (32)$$

where Φ is the CDF of standard normal distribution, and $\xi(d_0)$ is related μ and δ which can be expressed as,

$$\begin{aligned}\xi(d_0) &= [\frac{ph(d_0)}{\beta} - (\mu + \sigma)]/\sqrt{\delta} \\ \mu &= 2p\pi\lambda_b\Gamma(\triangle t; \lambda_t, \Omega)\mathbb{L}(d_1) \\ \delta &= \frac{2p^2\sqrt{\pi\lambda_b\Gamma(\triangle t; \lambda_t, \Omega)}}{R_0}[\mathbb{L}(d) - 2(\frac{\mathbb{L}(d_1)}{R_0})^2], \end{aligned} \quad (33)$$

where $\mathbb{L}(d_1) = \int_0^{R_0} d_1 h(d_1)d(d_1)$ represents the integral value. Hence, the probability of SINR can be calculated analytically when the related parameters are known. It is worth noting that the density of BN number λ_b and the threshold of SINR β are usually predefined, which depend on the communication environment. Thus, the TSR is only related to the length of time slot $\triangle t$ and blocksize threshold Ω. For a given trading frequency λ_t, the system can adaptively adjust threshold Ω and $\triangle t$ to maximize TSR. Numerical results will be detailed in Section VI.

2) SECURITY AND PRIVACY PERFORMANCE

Traditional blockchain system such as Bitcoin utilizes mining process to maintain the integrity of block information, which consumes huge computation power and sustain long computation latency. However, for the scenario of vehicular resource sharing, each node has limited computation resource, and the transaction of resource sharing needs to be fast identified as the high mobility of vehicles. It is impractical to utilize the computation process to maintain the blockchain system. Therefore, we proposed a new consensus mechanism, which cancels the computation process mechanism. The proposed mechanism can achieve the integrity of the global ledger and the defensibility of double-spending attack. In addition, the proposed architecture can prevent several malicious attacks in the resource sharing process. Next, we will discuss the security and privacy performance and show the effectiveness of the proposed PoR-based blockchain system.

a: DOUBLE-SPENDING ATTACKS

This is the most common attack of all the digital currency systems [35]. Double-spending is a potential flaw in the digital cash systems in which the same single digital token can be spent more than once. Bitcoin system resolves the attack under the assumption that there are no peers holding more than 50% computing power of the networks. In this paper, vehicles are only responsible for producing transactions, having no rights to publish blocks. Thus, vehicles are unable to implement Double-Spending attack. BNs collect transactions in the network and generate blocks according to the sum reputation. It is almost impossible for two blocks to have the same sum reputation at the same time. Therefore, BNs cannot control or predict the publishing of blocks and Double-Spending attack will be well resolved.

TABLE 2. Simulation parameters.

Parameters	Value
Number of TOs	50
Number of RPs	50
channel power gain $h(d)$	$35.2 + 35lod(d)$
channel nosie power σ^2	1
Initial Reputation value Rep	uniform distributed within $(0, 1)$
The length of timeslot $\triangle t$	10 s
The radius of interfering area R_0	30 m
The density λ_b of BN	$0.8\ /m^2$

b: INTERGITY

To achieve the integrity of blockchain, it must guarantee that no adversary could tamper the transactions of blocks recorded in the ledger. For the historical blocks, the attack can be well resolved based on *preHash* design of block. For newly chained block, Bitcoin adopts the target difficulty mechanism to resolve the attack. The Hash value of newly chained block must be less than the target, and any one wants to tamper the block will violate the condition. In this paper, we leverage the chronological order to resolve the attack. BNs encapsulate the transactions into one block in a chronological order. If other BNs want to tamper the newly chained block, they must replace the tampered transactions with the same timestamps. This is regarded as impossible in practice, thus the integrity of blockchain can be guaranteed.

c: BAD BEHAVIOURS

The attack is that vehicles provide fraudulent amount of resources or provide dishonest recommendation to defame good peers [36]. This can be well prevented by granting reward to good peers and punishing the bad. In this paper, we design the sum reputation mechanism. The BNs would always collect those transactions with higher reputation value to compete with other BNs to publish a new block. If a vehicle has a low reputation value, the transactions of the vehicle cannot be timely collected by BNs. In this case, other vehicles are not willing to trade with this vehicle. Consequently, PoR can inspire vehicles to behave in good manners and prevent the bad behaviours.

d: PRIVACY

Frequent resource sharing among vehicles may raise privacy concerns of sensitive information leakage, such as identities, driving routes, and trading preferences. In this paper, the resource sharing process is achieved by means of wallet addresses *WA*s. One vehicle could create multiple *WA*s according to its private key *PK* during different sharing processes. Therefore, it is difficult for attackers to invade the privacy of sharing vehicles and the preservation of privacy is achieved.

VI. PERFORMANCE EVALUATION

In order to validate the effectiveness and feasibility of proposed resource sharing system, the performance evaluation is conducted in this section. We evaluate the proposed

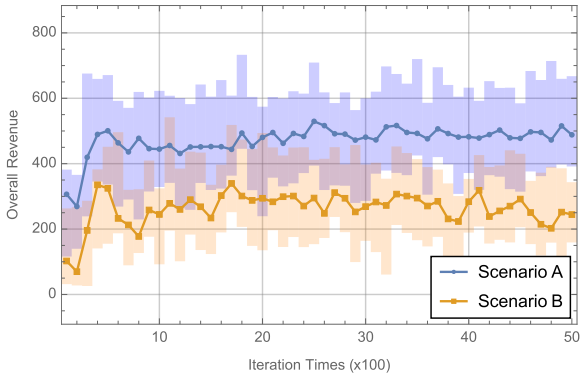

FIGURE 5. Overall Revenue under two scenarios.

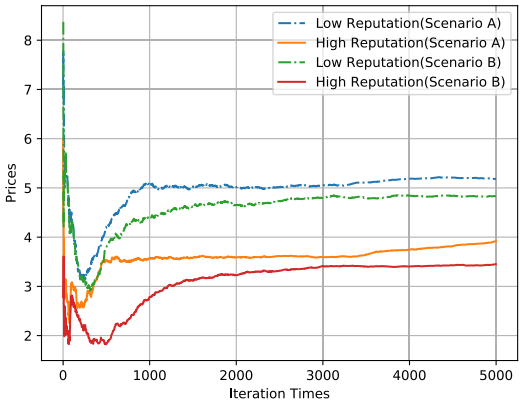

FIGURE 6. Prices for vehicles with different reputation.

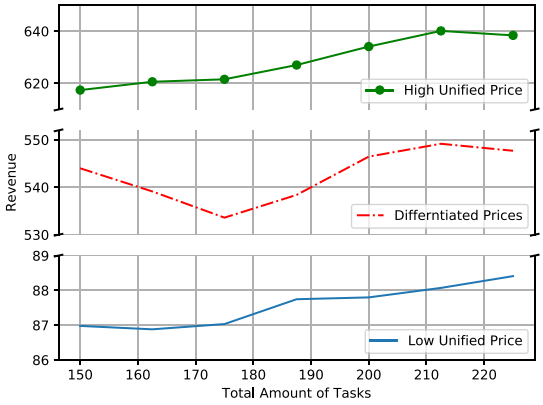

FIGURE 7. The revenue of RPs for different amount of Tasks.

DRL-based pricing scheme during the resource sharing process, and analyse the impact of reputation value on the collection rate. The configurations of key parameters are listed in Table 1.

A. ENERGY COST ANALYSIS

Since the proposed consensus mechanism does not rely on mining process, we define the system energy cost E_{total} as

$$E_{total} = E_{Sig} + E_{Aud}, \tag{34}$$

where E_{Sig} denotes the energy cost of signaling overhead for broadcasting blocks during the consensus procedure and E_{Aud} is the energy cost of audit procedure for validating blocks. The signaling overhead mainly depends on the execution of smart contract and consensus procedure. In our simulation scenario, the signaling overhead of executing smart contract is about $2\times(N+M)$. In the proposed reputation-based consensus procedure, the BNs collecting Ω transactions will broadcast a new block to other BNs. Then, each BN needs to validate the reputation value and transactions of the received blocks. Finally, all BNs send the validated results to the other BNs to finish the consensus procedure. Hence, the total amount of signaling overhead is about $K\times(K-1)+2\times(N+M)$. For the block audit procedure, the total amount of auditing is about $K\times(K-1)$. Assume that the size of one single block is 1 MBytes, the energy consumption of broadcasting and auditing one block is 1.0 J and 1.3 J, respectively [22]. Therefore, the system energy cost is about $2\times(N+M)+2.3\times K(K-1)$ J/block. Comparing to the huge energy consumption of mining process in Bitcoin and Ethereum, which consumes nearly 3.6GJ/block [37], our proposed blockchain system can effectively reduce the system energy cost.

B. NUMERICAL RESULTS OF PROPOSED RESOURCE SHARING MECHANISM

In the first experiment, we simulate the DDPG approach under two different task scenarios A and B, where A has more tasks than B. Thus, vehicles in scenario A need more resources to be traded. The pricing scheme and resource allocation scheme of each scenario are trained separately. The actor and critic networks of DDPG are trained, and the training history of the revenue of vehicles for the two scenarios are shown in Figure 5. It shows that DDPG approach achieves the converge under both scenarios, and the convergence rate is around 1000 iterations. This quick convergence attribution is crucial for the smart-contract based system, where the contract implementing cost is based on the executing times of blockchain virtual machine. Furthermore, we can figure that the overall revenue of scenario A is higher than that of B. This can be explained as with higher tasks, more spare resources of RPs are traded to satisfy the demands of TOs, contributing to the increase of overall revenue.

Figure 6 shows the influence of reputation value on the pricing scheme under the scenarios A and B. It can be observed that the prices in scenario A are higher than the prices in B. This is because that with the increase of task amount, the limited spare resources of RPs cannot satisfy the increasing demand. Therefore, the price will accordingly increase. For a specific scenario, we can figure that vehicles with a higher reputation could purchase the resources at relatively low prices. This proves the incentive effect of the reputation value. In this case, vehicles prefer to acquire high reputation values to implement the resource trading at lower prices. Figs 5-6 prove the feasibility and effectiveness of proposed utility function of vehicles.

As a differentiated pricing scheme is exploited for the resources, we choose two unified pricing schemes as the benchmark methods which includes high unified pricing

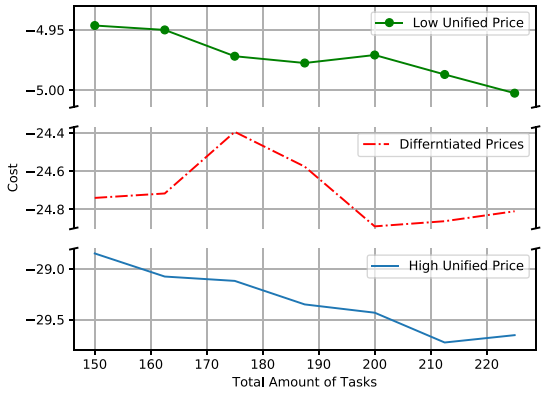

FIGURE 8. The cost of TOs for different amount of Tasks.

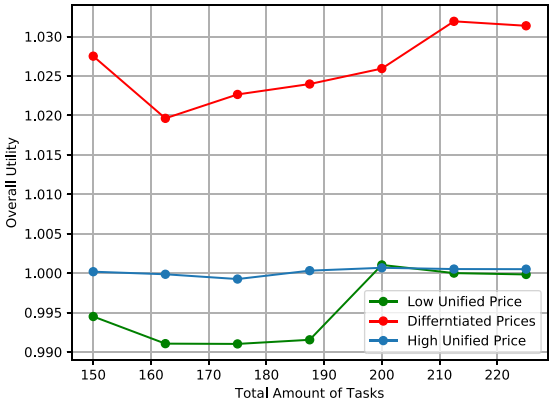

FIGURE 9. The overall utility for different amount of Tasks.

and low unified pricing schemes. Figure 7 investigates the relationship between the amount of computation tasks and the RPs' revenue under different price schemes. It can be observed that the RPs' revenue increases with the size of computation tasks under the two unified schemes. This is due to the fact that more tasks mean larger computing requirements, and the revenue of resource providers is accordingly increasing. Nevertheless, the curve of the proposed differentiated pricing scheme first falls and then rises. The falling part reveals the relationship between demand and supply. This can be explained that when demand is less than the supply (*i.e.*, tasks $<$ 175), the RPs tend to lower their resource prices to compete with others. The revenue reaches the minimum valley when demand equals supply. Afterwards TOs would raise their bid prices to purchase the limited resources, which results in the increases of the RPs' revenue consequently.

Figure 8 displays the TOs' cost with respect to the size of tasks under the different pricing schemes. The curve of unified schemes are gradually increasing. This is because that with the growing of the resource demands, TOs have to purchase resources from those RPs which have low reputation value. The cost of proposed scheme first decreases and then raises up. This also indicates the relationship of supply and demand. Figure 9 shows the overall utilities of the three pricing schemes. According to Figs. 7-9, we can deduce that

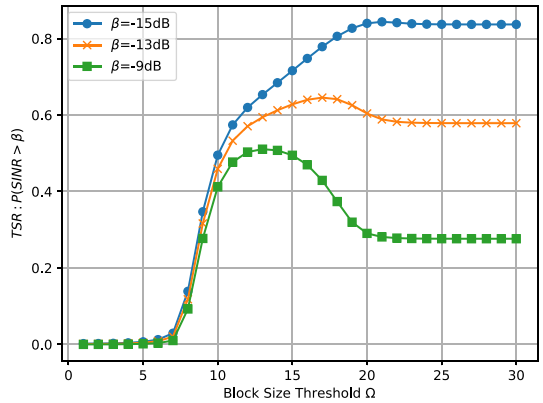

FIGURE 10. TSR under different block size.

comparing to both high and low unified pricing schemes, our proposed differentiated scheme can not only satisfy the utility of TOs, but also take into account of the revenue of RPs. For the normalized overall utility U^{tot}, the proposed differentiated pricing scheme achieves at most 30% increase compared with unified pricing schemes. Therefore, it can well realize the matching of resource supply and demand according to their individual requirement.

We then discuss the communication performance of the proposed PoR mechanism in Figure 10 and 11. Figure 10 shows the impact of the blocksize Ω on the TSR. It is worth noticing that the unit for block size threshold Ω is 10 transactions, $\Omega = 20$ means that the block size is 200 transactions. Considering the blocksize of conventional blockchain system Bitcoin is 2000 to 3000, the publishing interval of blocks is about 6 min. However, the publishing interval of the proposed blockchain is 10 s, which is far shorter than Bitcoin. Therefore, it is reasonable to design a smaller blocksize for the proposed system. As shown in Figure 10, with the increasing of Ω, the probability of TSR first rises and then goes down to a certain value. In addition, we also observe that different threshold values β contribute to different peak values of TSR. It can be explained that a larger blocksize will reduce the amount of candidate blocks while too large block will aggravate the difficulty for all publishers to generate the blocks. This feature is very important in practical scenarios, where the blocksize Ω should be carefully designed according to the communication condition.

In Figure 11, we discuss the influence of trading frequency on TSR. For one certain blocksize, the probability of TSR decreases with the increase of trading frequency λ_t. The TSR shows similar trend as the raise of $\triangle t$. This is because a larger λ_t and longer $\triangle t$ will increase the number of collected transactions, which further leads to more interfering blocks from other BNs. Observed from Figs. 10-11, given the SINR threshold β, the optimal blocksize Ω can be obtained. Furthermore, the timeslot $\triangle t$ can also be well designed according to the trading frequency λ_t to maximize TSR.

We finally examine the defensive ability against *Bad Behaviours* attack of the proposed PoR mechanism

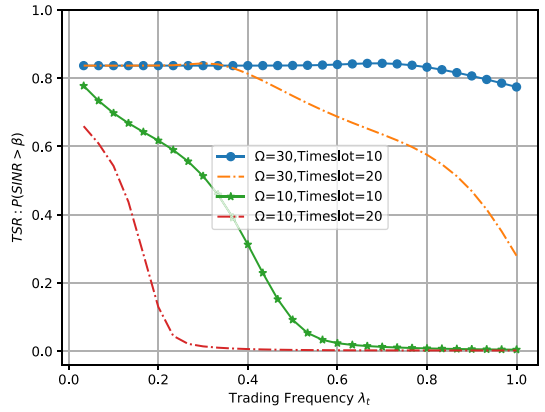

FIGURE 11. TSR under different trading frequency.

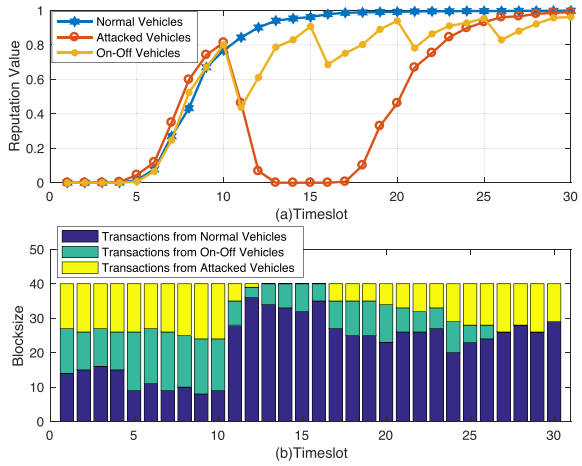

FIGURE 12. Defensive ability against bad behaviours attack.

in Figure 12. Here we define three types of vehicles: *a)Normal* refers to the honest vehicles during entire simulation period. *b)On − Off* refers to the vehicles with irregular behaviours performing well or badly alternatively to remain undetected [25]. *c)Attacked* refers to the vehicles being attacked by malicious party in a certain time and come back to *Normal* type after repairing the attack. We simulate 30 trading slots and the blocksize Ω is set as 40. Assuming that all vehicles behave *Normal* in the first 10 slots. As shown in Figure 11, the reputation value of all three types vehicles increase first. Before the 10-*th* trading slot, the BNs collect similar amount of transactions from the three types of vehicles. While at the beginning of timeslot 11, BNs rarely collect the transactions sent by *On−off* or *Attacked* vehicles. The reason is due to that BNs will prefer to collect those transactions with higher reputation value. Furthermore, after timeslot 25, the transaction collecting rate of *On − off* vehicles decreases rapidly due to their irregular behaviours. The transaction collecting rate of *Attacked* vehicles slowly raises up because that the vehicles constantly gain the reputation values after repairing the attack. This shows the effectiveness of our proposed *PoR* mechanism and the resistance against malicious attacks.

VII. CONLUSION

In this paper, a proof-of-Reputation based blockchain architecture is proposed for the resource sharing in IoV networks. The lightweight blockchain takes charge of the trust management as well as the privacy preservation during the resource sharing process and the reduction of computing expenditure. The Proposed PoR consensus mechanism guarantees the security of blockchain and the incentive of vehicles. We also analyse the communication performance of proposed PoR protocol with respect to the block size and trading frequency. Simulation results demonstrate the effectiveness of the proposed mechanism, our proposed differentiated pricing scheme achieves at most 30% increase compared to the unified pricing schemes and the proposed PoR based blockchain system effectively reduce the energy consumption compared to traditional blockchain systems. As future work, we will focus on studying the overhead and transaction throughput of PoR based on actual distribution application platforms.

REFERENCES

[1] R. Morello, S. C. Mukhopadhyay, Z. Liu, D. Slomovitz, and S. R. Samantaray, "Advances on sensing technologies for smart cities and power grids: A review," *IEEE Sensors J.*, vol. 17, no. 23, pp. 7596–7610, Dec. 2017.

[2] S. Yu, G. Wang, X. Liu, and J. Niu, "Security and privacy in the age of the smart Internet of Things: An overview from a networking perspective," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 14–18, Sep. 2018.

[3] C. Shao, S. Leng, Y. Zhang, A. Vinel, and M. Jonsson, "Performance analysis of connectivity probability and connectivity-aware MAC protocol design for platoon-based VANETs," *IEEE Trans. Veh. Technol.*, vol. 64, no. 12, pp. 5596–5609, Dec. 2015.

[4] F. Lyu, N. Cheng, H. Zhu, H. Zhou, W. Xu, M. Li, and X. S. Shen, "Intelligent context-aware communication paradigm design for IoVs based on data analytics," *IEEE Netw.*, vol. 32, no. 6, pp. 74–82, Nov./Dec. 2018.

[5] X. Wang, S. Leng, and K. Yang, "Social-aware edge caching in fog radio access networks," *IEEE Access*, vol. 5, pp. 8492–8501, 2017.

[6] L. Liang, S. Xie, G. Y. Li, Z. Ding, and X. Yu, "Graph-based resource sharing in vehicular communication," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4579–4592, Jul. 2018.

[7] L. Wu, Q. Sun, X. Wang, J. Wang, S. Yu, Y. Zou, B. Liu, and Z. Zhu, "An efficient privacy-preserving mutual authentication scheme for secure V2V communication in vehicular ad hoc network," *IEEE Access*, vol. 7, pp. 55050–55063, 2019.

[8] T. Cheng, G. Liu, Q. Yang, and J. Sun, "Trust assessment in vehicular social network based on three-valued subjective logic," *IEEE Trans. Multimedia*, vol. 21, no. 3, pp. 652–663, Mar. 2019.

[9] S. Guleng, C. Wu, X. Chen, X. Wang, T. Yoshinaga, and Y. Ji, "Decentralized trust evaluation in vehicular Internet of Things," *IEEE Access*, vol. 7, pp. 15980–15988, 2019.

[10] T. Rosenstatter and C. Englund, "Modelling the level of trust in a cooperative automated vehicle control system," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 4, pp. 1237–1247, Apr. 2018.

[11] L. Huang, G. Zhang, S. Yu, A. Fu, and J. Yearwood, "Customized data sharing scheme based on blockchain and weighted attribute," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, United Arab Emirates, Dec. 2018, pp. 206–212.

[12] Z. Yang, K. Yang, L. Lei, K. Zheng, and V. C. M. Leung, "Blockchain-based decentralized trust management in vehicular networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1495–1505, Apr. 2019.

[13] 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.

[14] Z. Su, Y. Wang, Q. Xu, M. Fei, Y.-C. Tian, and N. Zhang, "A secure charging scheme for electric vehicles with smart communities in energy blockchain," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4601–4613, Jun. 2019.

[15] V. A. Siris, D. Dimopoulos, N. Fotiou, S. Voulgaris, and G. C. Polyzos, "Interledger smart contracts for decentralized authorization to constrained things," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Paris, France, Apr./May 2019, pp. 336–341.
[16] K. Govindan and P. Mohapatra, "Trust computations and trust dynamics in mobile adhoc networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 2, pp. 279–298, 2nd Quart., 2012.
[17] T. Gazdar, A. Belghith, and H. Abutair, "An enhanced distributed trust computing protocol for VANETs," *IEEE Access*, vol. 6, pp. 380–392, 2018.
[18] J. Cui, X. Zhang, H. Zhong, Z. Ying, and L. Liu, "RSMA: Reputation system-based lightweight message authentication framework and protocol for 5G-enabled vehicular networks," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6417–6428, Aug. 2019.
[19] M. Li, L. Zhu, and X. Lin, "Efficient and privacy-preserving carpooling using blockchain-assisted vehicular fog computing," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4573–4584, Jun. 2019.
[20] H. Li, K. Wang, T. Miyazaki, C. Xu, S. Guo, and Y. Sun, "Trust-enhanced content delivery in blockchain-based information-centric networking," *IEEE Netw.*, vol. 33, no. 5, pp. 183–189, Sep./Oct. 2019.
[21] L. Li, J. Liu, L. Cheng, S. Qiu, W. Wang, X. Zhang, and Z. Zhan, "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.
[22] G. Qiao, S. Leng, H. Chai, A. Asadi, and Y. Zhang, "Blockchain empowered resource trading in mobile edge computing and networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, May 2019, pp. 1–6.
[23] F. Guo, H. Zhang, H. Ji, X. Li, and V. C. M. Leung, "An efficient computation offloading management scheme in the densely deployed small cell networks with mobile edge computing," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2651–2664, Dec. 2018.
[24] B. Du, Q. Wei, and R. Liu, "An improved quantum-behaved particle swarm optimization for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 6003–6017, Aug. 2019.
[25] F. Gai, B. Wang, W. Deng, and W. Peng, "Proof of reputation: A reputation-based consensus protocol for peer-to-peer network," in *Database Systems for Advanced Applications* (Lecture Notes in Computer Science), vol. 10828, J. Pei, Y. Manolopoulos, S. Sadiq, and J. Li, Eds. Cham, Switzerland: Springer, 2018.
[26] A. K. Datta and R. Patel, "CPU scheduling for power/energy management on multicore processors using cache miss and context switch data," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 5, pp. 1190–1199, May 2014.
[27] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg, "Real-time detection and tracking for augmented reality on mobile phones," *IEEE Trans. Vis. Comput. Graphics*, vol. 16, no. 3, pp. 355–368, May/Jun. 2010.
[28] B. Dong, R. Liu, and H. W. Wang, "Trust-but-verify: Verifying result correctness of outsourced frequent itemset mining in data-mining-as-a-service paradigm," *IEEE Trans. Services Comput.*, vol. 9, no. 1, pp. 18–32, Jan./Feb. 2016.
[29] W. Zhang, Y. Lin, and G. Qi, "Catch you if you misbehave: Ranked keyword search results verification in cloud computing," *IEEE Trans. Cloud Comput.*, vol. 6, no. 1, pp. 74–86, Mar. 2015.
[30] Y. Wei, F. R. Yu, M. Song, and Z. Han, "User scheduling and resource allocation in HetNets with hybrid energy supply: An actor-critic reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 680–692, Jan. 2018.
[31] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," in *Proc. 31st Int. Conf. Mach. Learn. (PMLR)*, 2014, 32, no. 1, pp. 387–395.
[32] C. H. Liu, Z. Chen, and Y. Zhan, "Energy-efficient distributed mobile crowd sensing: A deep learning approach," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1262–1276, Jun. 2019.
[33] Z. Xu, J. Tang, J. Meng, W. Zhang, Y. Wang, C. H. Liu, and D. Yang, "Experience-driven networking: A deep reinforcement learning based approach," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Honolulu, HI, USA, Apr. 2018, pp. 1871–1879.
[34] Y. Sun, L. Zhang, G. Feng, B. Yang, B. Cao, and M. A. Imran, "Blockchain-enabled wireless Internet of Things: Performance analysis and optimal communication node deployment," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5791–5802, Jun. 2019.
[35] 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.
[36] Y. Zhao, Y. Li, Q. Mu, B. Yang, and Y. Yu, "Secure pub-sub: Blockchain-based fair payment with reputation for reliable cyber physical systems," *IEEE Access*, vol. 6, pp. 12295–12303, 2018.
[37] D. Rusovs, S. Jaundālders, and P. Stanka, "Blockchain mining of cryptocurrencies as challenge and opportunity for renewable energy," in *Proc. IEEE 59th Int. Sci. Conf. Power Elect. Eng. Riga Tech. Univ. (RTUCON)*, Riga, Latvia, Nov. 2018, pp. 1–5.

HAOYE CHAI received the B.Sc. degree in information and communication engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2016, where he is currently pursuing the Ph.D. degree. His research interests include mobile edge computing, the Internet of Vehicles, and blockchain in wireless networks.

SUPENG LENG received the Ph.D. degree from Nanyang Technological University (NTU), Singapore. He is currently a Full Professor and the Vice Dean of the School of Information and Communication Engineering, University of Electronic Science and Technology of China (UESTC). He is also the Leader of the Research Group of Ubiquitous Wireless Networks. He has been working as a Research Fellow of the Network Technology Research Center, NTU. His research focuses on resource, spectrum, energy, routing and networking in the Internet of Things, vehicular networks, broadband wireless access networks, smart grid, and the next generation mobile networks. He has published over 180 research articles in recent years. He serves as an Organizing Committee Chair and a TPC member for many international conferences, as well as a Reviewer for over ten international research journals.

KE ZHANG received the Ph.D. degree from the University of Electronic Science and Technology of China, in 2017. He is currently a Lecturer with the School of Information and Communication Engineering, University of Electronic Science and Technology of China. His research interests include scheduling of mobile edge computing, design and optimization of next-generation wireless networks, and the Internet of Things.

SUN MAO received the B.Sc. degree in information and communication engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2016, where he is currently pursuing the Ph.D. degree. His research interests include mobile edge computing, the Internet of Vehicles, and blockchain in wireless networks.