When Deep Reinforcement Learning Meets 5G-Enabled Vehicular Networks: A Distributed Offloading Framework for Traffic Big Data

Zhaolong Ning, *Senior Member, IEEE*, Peiran Dong, Xiaojie Wang, Mohammad S. Obaidat, *Fellow, IEEE*, Xiping Hu, Lei Guo, Yi Guo, Jun Huang, Bin Hu, and Ye Li, *Senior Member, IEEE*

***Abstract*—The emerging 5G-enabled vehicular networks can satisfy various requirements of vehicles by traffic offloading. However, limited cellular spectrum and energy supplies restrict the development of 5G-enabled applications in vehicular networks. In this article, we construct an intelligent offloading framework for 5G-enabled vehicular networks, by jointly utilizing licensed cellular spectrum and unlicensed channels. A cost minimization problem is formulated by considering the latency constraint of users and is further decomposed into two subproblems due to its complexity. For the first subproblem, a two-sided matching algorithm is proposed to schedule the unlicensed spectrum. Then, a deep-reinforcement-learning-based method is investigated for the second one, where the system state is simplified to realize distributed traffic offloading. Real-world traces of taxies are leveraged to illustrate the effectiveness of our solution.**

***Index Terms*—Deep reinforcement learning (DRL), distributed offloading, traffic big data, 5G-enabled vehicular networks.**

Manuscript received May 29, 2019; revised August 1, 2019; accepted August 14, 2019. Date of publication August 22, 2019; date of current version January 14, 2020. This work was supported in part by the Ministry of Education of China under Distinguished Possessor Grant MS2017BJKJ003 given to Prof. M. S. Obaidat, in part by the National Nature Science Foundation of China under Grant 61971084, Grant 61802159, Grant 61771120, Grant 61632014, and Grant 61671092, in part by China Postdoctoral Science Foundation under Grant 2018T110210, in part by the Fundamental Research Funds for the Central Universities under Grant DUT19JC18, in part by the State Key Laboratory of Integrated Services Networks, Xidian University under Grant ISN20-01, and in part by the Science and Technology Innovation Program of National Defense, Science, and Technology Planning Project of Shenzhen under Grant JCYJ20170818111012390. Paper no. TII-19-2109. *(Corresponding authors: Xiping Hu; Lei Guo; Yi Guo; Jun Huang; Bin Hu; Ye Li.)*

Z. Ning is with the School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China, with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710126, China, with the School of Software, Dalian University of Technology, Dalian 116024, China, and also with the Chongqing Key Laboratory of Mobile Communication Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: zhaolongning@dlut.edu.cn).

P. Dong and X. Wang are with the School of Software, Dalian University of Technology, Dalian 116024, China (e-mail: peiran_dong@outlook.com; wangxj1988@mail.dlut.edu.cn).

M. S. Obaidat is with the College of Computing and Informatics, University of Sharjah, Sharjah 27272, UAE, with Nazarbayev University, Astana 010000, Kazakhstan, with the King Abdullah II School of Information Technology, The University of Jordan, Amman 11180, Jordan, and also with the University of Science and Technology Beijing, Beijing 100083, China (e-mail: msobaidat@gmail.com).

X. Hu is with the School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China, and also with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (e-mail: huxp@lzu.edu.cn).

L. Guo and J. Huang are with the Chongqing Key Laboratory of Mobile Communication Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: guolei@cqupt.edu.cn; jhuang@cqupt.edu.cn).

Y. Guo is with Second Clinical Medical College, Jinan University, Guangzhou 510632, China (e-mail: xuanyi_guo@163.com).

B. Hu is with the School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China (e-mail: bh@lzu.edu.cn).

Y. Li is with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (e-mail: ye.li@siat.ac.cn).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2019.2937079

I. Introduction

With the recent advancements in 5G-enabled wireless communication technologies, electronic devices are expected to be connected to provide smart services for humans and facilitate our daily lives. The Internet of vehicles (IoV) is an emerging paradigm, which enables vehicles to perform various applications and provide comfortable traveling experiences for drivers and passengers [1], [2]. However, these applications and services generally have strict processing deadline constraints and limited energy supplies. Meanwhile, the explosive growth of vehicular data and complicated interactions between vehicles and servers pose great challenges for mobile network operators to guarantee the quality of service (QoS) of users and decrease the energy cost of network communications [3].

According to the recent report by Cisco [4], global mobile data traffic will remain rapid growth in the near future due to the increasing number of end devices and a wide variety of relevant applications. Specifically, the total number of end devices around the world will be 75 billion, and the amount of data will surpass 30.6 exabytes per month by 2020. Compared with 4G long-term evolution (LTE) networks, the 5G-enabled wireless communication system can effectively enhance the performance of vehicular networks [5]. Since the era of 5G-enabled wireless communication is coming, current access strategies may face great challenges in serving massive connected devices.

The nonorthogonal multiple access (NOMA) technology is proposed to enable concurrent transmissions of multiple users, by taking full advantages of wireless spectrum resources. By exploiting successive interference cancellation, receivers can divide superimposed messages from users and serve users simultaneously at different power levels [6].

Mobile edge computing (MEC) is promising for 5G-enabled communication networks. Since edge devices are proximate to users, MEC systems can meet low-latency requirements of users for various applications. Considering the time variety of vehicular networks, edge caching and computation offloading at road-side units (RSUs) are promising to deal with the delay-sensitive applications, which efficiently enhance the QoS of network users. Along with the advanced multiaccess technologies in 5G-enabled communication networks, cooperative and distributive computing architectures gain much attentions in both academia and industry [7], [8]. Since the spectrum and computing resources are constrained in edge servers, it is imperative to design an efficient scheme by jointly considering computation offloading and radio frequency spectrum allocation.

There are some existing studies focusing on computational task offloading and spectrum resource allocation. In [9], a collaborative quality-aware service access system is built for reliability assurance and service quality improvement in IoVs. The authors of [10] construct a multiuser offloading system and prove the NP-hardness of the formulated optimization problem. Then, they utilize a game-theoretic approach to achieve a Nash equilibrium. The authors of [11] consider a partial computation offloading architecture, where users offload partial tasks to MEC servers while executing others locally. Although partial offloading can utilize computational resources placed at servers, interference exists among subchannels in practice when servers allocate spectrum resources for multiple users. The authors of [12] design a resource allocation policy to minimize the energy cost for an MEC offloading framework. In [13], the authors jointly optimize the allocation of spectrum and computing resources to minimize the weighted sum energy cost in a multiple-input multiple-output system. However, these research studies do not consider real traffic data and may not be feasible for IoV systems to tackle the big data generated from vehicles and servers in smart cities.

Artificial intelligence (AI) is an advanced tool to tackle the challenge of big data when addressing the radio and computational resource allocation problems in 5G-enabled communication networks [14]. Deep reinforcement learning (DRL) plays an important part in AI, which gains a great success in many fields, such as game playing, robot controlling, and dialogue generation [15]–[17]. According to its strong cognitive ability for time-varying environments, DRL has widespread applications in vehicular networks [18], [19]. In large-scale vehicular networks, DRL can significantly improve the decision-making speed to control task offloading, spectrum access, and allocate resources dynamically for a massive number of vehicles. With the explosion of traffic data, traditional centralized methods based on DRL cannot learn knowledge from environments rapidly. Specially, in MEC-enabled systems, merely selecting the base station as the agent will have great pressures to make decisions for all vehicles. Hence, it is critical to develop a distributed DRL method to accelerate the speed of data processing and learning.

In order to minimize the offloading cost (i.e., transmission and computation cost) of vehicles while satisfying the latency constraint, this article constructs an intelligent offloading framework for 5G-enabled vehicular networks. Since the number of reused subchannels is limited, it is irrational to offload tasks merely through the reused subchannels by NOMA. In this article, cellular channels and reused subchannels are cooperative to undertake task transmissions, by which the transmission load can be balanced to cope with the large number of vehicles. Latency sensitive tasks can be processed on the macrocell to ensure the QoS of vehicle-to-infrastructure (V2I) users. Correspondingly, computation-intensive tasks can be offloaded to RSUs to decrease the offloading overhead of vehicle-to-RSU (V2R) users. However, long-term cognition of users and environment is challenging for traditional offloading scheduling. Therefore, a distributed DRL-based solution is further developed for the optimization problem. The main contributions are as follows.

1) We construct an intelligent offloading framework for 5G-enabled vehicular networks and formulate an optimization problem to minimize the offloading cost while satisfying the latency constraint of users.
2) The formulated problem is divided into two subproblems, i.e., the V2R scheduling and V2I allocation, due to its complexity. For the first subproblem, a two-sided matching algorithm is proposed to schedule the unlicensed spectrum, while a distributed DRL-based resource allocation method is developed for the second one.
3) Different from most DRL-based algorithms performing in a centralized manner, we simplify the system state and develop a distributed DRL-based method, which can greatly decrease the communication overhead between vehicles and the macrocell.
4) Traces of taxies in Hongkou district Shanghai (China) are utilized to illustrate the effectiveness of our solutions from the aspect of offloading cost reduction.

The rest of this article is organized as follows. The system model is illustrated in Section II. Section III formulates the optimization problem. Two-sided matching and the DRL-based schemes are specified in Section IV. Performance evaluation results are illustrated in Section V, and Section VI concludes this article.

II. System Model

A. System Overview

As illustrated in Fig. 1, various kinds of network elements coexist in the 5G-enabled traffic management system, such as macrocells, RSUs, and vehicles [20]. Both the computing servers equipped by the macrocell and MEC servers are capable to process offloading tasks. The primary difference between them is that the former has sufficient resources and powerful computing capabilities, while resources and computing capabilities of the latter are usually constrained. The considered 5G-enabled vehicular network includes N V2I and M V2R users.

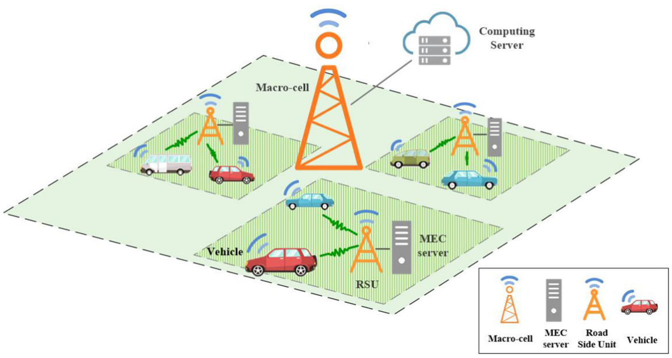


Fig. 1. Illustration of 5G-enabled vehicular networks.

We consider one macrocell server and J RSUs coexist in a local region. Denote V2I transmitters by $n \in \mathcal{N} = \{1, \ldots, N\}$, V2R users by $m \in \mathcal{M} = \{N+1, \ldots, N+M\}$, and RSUs by $j \in \mathcal{J} = \{1, \ldots, J\}$. The main components in the 5G-enabled vehicular network are illustrated as follows: the macrocell station is generally deployed and managed by network operators to provide a seamless communication coverage for vehicles. For traffic offloading, computing servers are also integrated into macrocells. However, the bandwidth of the macrocell is increasingly saturated, and additional costs will be charged for traffic offloading through macrocells.

RSUs are deployed along roads to fulfill communication and computing tasks of vehicles, and they are generally equipped with MEC servers to store contents and offload network traffic. Since deploying a large number of RSUs is expensive, we consider that the number of RSUs is constrained, and the location information of nearby RSUs can be acquired through information synchronization.

On-board units equipped by vehicles can be utilized to fulfill the vehicle-to-vehicle communications through short-distance communication technologies. With the increasingly powerful sensors and communication technologies, information can be exchanged among vehicles following a device-to-device manner. Vehicles can also connect to either the macrocell or RSUs for traffic offloading.

B. V2I Model

Cellular V2I users can access the dedicated cellular spectrum by the NOMA technique. The bandwidth of the dedicated cellular channels for V2I is divided into K_I subchannels, i.e., $\mathcal{K}_I = \{1, 2, \ldots, K_I\}$. For each vehicle accessing to the macrocell, its task to be processed is $\tau_n = \{d_n, c_n, T_n^{\max}, n \in \mathcal{N}\}$. Herein, the task size to be offloaded by user n is d_n, the number of CPU cycles for processing the task is c_n, and the tolerable latency is $T_n^{\max}$. Since the transmission of V2I users follows LTE standards in the dedicated spectrum, the length of each task frame is regarded to be constant. As long as one frame has been received, the transmission process can be finished. We assume that one task contains T_n data frames, and the length of each subframe is $d_n(t), t \in \mathcal{T}_n = \{1, \ldots, T_n\}$. Similar to [21], a quasi-static network scenario is considered in this article, i.e., during the processes of transmission and offloading, the channel condition is stable.

The transmission power of V2I users is denoted as $p_{n,0}$. Let $U_n \subseteq \mathcal{N}$ be the set of V2I users, occupying the same channel with user n. The received signal of the macrocell from vehicle n can be computed by

$$y_{n,0} = \sqrt{p_{n,0}} h_{n,0} x_{n,0} + \sum_{i \neq n, i \in U_n} \sqrt{p_{i,0}} h_{i,0} x_{i,0} + \sigma \tag{1}$$

where $h_{n,0}$ denotes the channel gain between user n and the macrocell station, and $x_{n,0}$ is the corresponding transmitted signal. The first part on the right of (1) is the desired signal from user n, and the second part is the summation of the intracell interference from others occupying the same communication channel. Variable σ is the Gaussian noise.

The path-loss model in [22] is selected for the channel gain modeling of cellular V2I users. Let p_0 and p denote the received signal power at d_0 and d away from the user, respectively. Received power $p = p_0(d/d_0)^{-\alpha}|h_0^2|$ holds, where α is the exponent of the path loss, and the complex Gaussian variable h_0 represents the Rayleigh fading. For simplicity, we set $d_0 = 1$. Channel gain $h_{n,0}$ from user n to the macrocell can be expressed as

$$|h_{n,0}|^2 = G|d_{n,0} + v_{n,0} t_w|^{-\alpha} |h_0|^2 \tag{2}$$

where G is the fixed power gain, $d_{n,0}$ and $v_{n,0}$ are the distance and the relative velocity from user n to the macrocell, respectively, and t_w is the interval of the waiting queue.

There are different ways for V2I and V2R users to occupy channels. For V2I users, they share the channel by separating it into distinct subchannels, while the whole channel is occupied by one V2R user. We define a binary variable $\theta_{n,0,k}(t)$ to indicate subchannel allocation, where

$$\theta_{n,0,k}(t) = \begin{cases} 1, & \text{if subchannel } k \text{ is allocated to V2I} \\ & \text{user } n \text{ in subframe } t \\ 0, & \text{otherwise} \end{cases}. \tag{3}$$

For one subframe transmission, we consider that one subchannel can be reused by at most q V2I users simultaneously, i.e., $|U_n| \leqslant q$. The signal-to-interference-plus-noise ratio (SINR) is leveraged to measure channel interference. In addition, since V2R and V2I users do not share the same channel, their interference can be ignored. Therefore, the SINR of V2I user can be obtained by

$$\Gamma_{n,0,k}(t) = \frac{p_{n,0}(t) h_{n,0}^2 \theta_{n,0,k}(t)}{\sum_{i \in U_n, i \neq n} \theta_{i,0,k}(t) p_{i,0}(t) h_{i,0}^2 + \sigma^2}. \tag{4}$$

Both the macrocell and RSUs are equipped with computing servers, providing offloading services for V2I and V2R users, respectively. For the macrocell, we assume that its computation resources are sufficient, where the offloading latency (i.e., queue waiting latency and computation latency) can be neglected. However, users are charged for renting bandwidths and offloading messages. Following the charging policy in [20], the cost function can be defined as follows:

$$\mathcal{U}_n(t) = \mathcal{Y}(t)_i(d_n, c_n) \cdot \left(1 - e^{-d_n \eta \Delta t} \cdot \theta_{n,0,k}(t)\right) \tag{5}$$

where $\mathcal{Y}(t)_i(d_n, c_n)$ is a utility function that reflects the QoS of users, which has been investigated in [23]. Variable η denotes the unit cost for offloading messages to the macrocell, and Δt is the gap of execution latency caused by offloading to RSUs and that by the macrocell. In our model, since the execution latency of the macrocell can be neglected, Δt equals the execution latency caused by offloading to RSUs.

C. V2R Model

Contrary to the macrocell, RSUs offer free and preemptive services with extra delay because of their constraint computation capabilities. In order to save limited spectrum resources, V2R users access RSUs over the unlicensed spectrum rather than the dedicated cellular spectrum based on orthogonal frequency-division multiplexing (OFDM) technology. Correspondingly, the unlicensed spectrum can be divided into K_R subchannels, denoted by $\mathcal{K}_R = \{K_I+1, K_I+2, \ldots, K_I+K_R\}$, to support the transmission of multiple V2R users. There are M V2R users within the coverage of RSUs. Similar to V2I users, their task can be expressed by $\tau_m = \{d_m, c_m, T_m^{\max}, m \in \mathcal{M}\}$.

If V2R user m transmits data to RSU j ($j = 0$ for the macrocell), the corresponding transmission power is denoted by $p_{i,j}$. The received signal of RSU j can be calculated by

$$y_{m,j} = \sqrt{p_{m,j}} h_{m,j} x_{m,j} + \sigma. \tag{6}$$

Different from V2I, V2R adopts OFDM technology, where no interference affects the transmission performance. Indicator $\theta_{m,j,k}(t)$ can be represented by

$$\theta_{m,j,k}(t) = \begin{cases} 1, & \text{if subchannel } k \text{ is allocated to V2R} \\ & \text{user } m \text{ in subframe } t \\ 0, & \text{otherwise} \end{cases}. \tag{7}$$

Note that the V2R connection follows 802.11 standards, where the transmission cannot be interrupted until the whole task has been received. Thus, for any subframe t and t', $(t, t' \in T_m)$, $\theta_{m,j,k}(t) = \theta_{m,j,k}^{(t')}$ holds. The SINR of V2R users is

$$\Gamma_{m,j,k}(t) = \frac{p_{m,j}(t) h_{m,j}^2 \theta_{m,j,k}(t)}{\sigma^2}. \tag{8}$$

The bandwidth of subchannel k is denoted by b_k. When subframe t is transmitted, the transmission rate of V2R user m over subchannel k to RSU j can be worked out by

$$r_{m,j,k}(t) = \theta_{m,j,k}(t) b_k \log_2 \left(1 + \Gamma_{m,j,k}(t)\right). \tag{9}$$

If user m selects RSU j to offload tasks, the transmission latency for subframe t can be computed by

$$T_{m,j,k}^{T(t)} = \frac{d_m(t)}{r_{m,j,k}(t)}. \tag{10}$$

The processing time of task τ_m through unlicensed channel k is denoted by $T_{m,j,k}^{P(t)}$. Given the computation capability of RSU j (denoted by f_j), the delay of user m to process task τ_m by RSU j can be calculated by

$$T_{m,j,k}^{P(t)} = \frac{f_j}{c_m(t)}. \tag{11}$$

Therefore, the total execution latency for user m to offload task to RSU j is

$$T_{m,j,k}(t) = T_{m,j,k}^{T(t)} + T_{m,j,k}^{P(t)}. \tag{12}$$

III. Problem Formulation

An overhead minimization problem is formulated by focusing on the dedicated cellular and unlicensed spectrum allocation in this section. The total offloading cost is defined as the summation cost of V2I users. Due to its complexity, we then reformulate this problem as a Markov-based reinforcement learning problem in vehicular networks.

A. Spectrum Allocation Problem

Various constraints should be considered in the formulated problem, including latency, power allocation, and channel assignment. The optimization problem is formulated as follows:

$$\min_{\theta_{i,j,k}(t)} \sum_{i \in \mathcal{N} \cup \mathcal{M}} \sum_{t \in \mathcal{T}_i} \mathcal{U}_i(t) \tag{13}$$

s.t.

$$\sum_{j=0}^{J} \theta_{i,j,k}(t) = 1 \quad \forall i \in \mathcal{N} \cup \mathcal{M}, k \in \mathcal{K}_I \cup \mathcal{K}_R, t \in \mathcal{T}_i \tag{13a}$$

$$\sum_{i=1}^{N} \theta_{i,0,k}(t) \leqslant q \quad \forall k \in \mathcal{K}_I, t \in \mathcal{T}_i \tag{13b}$$

$$\sum_{i=N+1}^{N+M} \sum_{j=1}^{J} \theta_{i,j,k}(t) = 1 \quad \forall k \in \mathcal{K}_R, t \in \mathcal{T}_i \tag{13c}$$

$$\sum_{i=N+1}^{N+M} \sum_{j=1}^{J} \sum_{k=K_I+1}^{K_I+K_R} \theta_{i,j,k}(t) \leqslant J \quad \forall t \in \mathcal{T}_i \tag{13d}$$

$$\sum_{i=1}^{N} \sum_{k=1}^{K_I} \theta_{i,0,k}(t) b_k \leqslant B_I \quad \forall t \in \mathcal{T}_i \tag{13e}$$

$$\sum_{i=N+1}^{N+M} \sum_{j=1}^{J} \sum_{k=K_I+1}^{K_I+K_R} \theta_{i,j,k}(t) b_k \leqslant B_R \quad \forall t \in \mathcal{T}_i \tag{13f}$$

$$\sum_{t=1}^{T_i} \theta_{i,j,k}(t) T_{i,j,k}(t) \leqslant T_i^{\max} \quad \forall i \in \mathcal{M} \tag{13g}$$

$$\Gamma_{i,0,k}(t) \geqslant \Gamma^{\min} \quad \forall i \in \mathcal{N} \tag{13h}$$

$$\theta_{i,j,k}(t) \in \{0, 1\} \quad \forall i \in \mathcal{N} \cup \mathcal{M}, k \in \mathcal{K}_I \cup \mathcal{K}_R, t \in \mathcal{T}_i. \tag{13i}$$

Constraint (13a) indicates that one task should be offloaded through either RSUs or the macrocell station. Constraints (13b) and (13c) show that each dedicated cellular and unlicensed subchannel can merely be allocated to at most q V2I and one V2R users, respectively. Constraint (13d) indicates that the number of occupied RSUs cannot surpass the number of available RSUs. The constraints for bandwidth allocation of both the macrocell and RSUs are shown in (13e) and (13f), respectively. The execution latency of task offloading by RSUs is constrained in (13g). Constraint (13h) requires that the interference generated by subchannel sharing cannot exceed threshold $\Gamma^{\min}$, indicating the minimum SINR to decode packets successfully.

B. Subproblem Formulation

Since RSUs provide free offloading services based on the unlicensed spectrum, users can compete for RSUs to reduce their cost. However, the extra latency caused by MEC may affect the user's QoE when the latency constraint is violated. Therefore, the procedure of spectrum allocation can be divided into two subproblems: centralized unlicensed spectrum allocation (CUSA) and distributed cellular spectrum allocation (DCSA).

For CUSA, the main constraint is the latency tolerance [i.e., constraint (13g)]. In order to minimize the offloading cost, we first allocate unlicensed spectrum resources for V2R users, while satisfying their latency constraints. The optimization target can be represented by

$$\min_{\theta_{i,j,k}(t)} \sum_{i\in\mathcal{M}}\sum_{t\in\mathcal{T}_i}\mathcal{U}_i(t) \tag{14}$$

s.t.

$$\sum_{j=0}^{J}\theta_{i,j,k}(t)=1 \quad \forall i\in\mathcal{M}, k\in\mathcal{K}_R, t\in\mathcal{T}_i \tag{14a}$$

$$\sum_{i=N+1}^{N+M}\sum_{j=1}^{J}\theta_{i,j,k}(t)=1 \quad \forall k\in\mathcal{K}_R, t\in\mathcal{T}_i \tag{14b}$$

$$\sum_{i=N+1}^{N+M}\sum_{j=1}^{J}\sum_{k=K_I+1}^{K_I+K_R}\theta_{i,j,k}(t)\leqslant J \quad \forall t\in\mathcal{T}_i \tag{14c}$$

$$\sum_{i=N+1}^{N+M}\sum_{j=1}^{J}\sum_{k=K_I+1}^{K_I+K_R}\theta_{i,j,k}(t)b_k\leqslant B_R \quad \forall t\in\mathcal{T}_i \tag{14d}$$

$$\sum_{t=1}^{T_i}\theta_{i,j,k}(t)T_{i,j,k}(t)\leqslant T_i^{\max} \quad \forall i\in\mathcal{M} \tag{14e}$$

$$\theta_{i,j,k}(t)\in\{0,1\} \quad \forall i\in\mathcal{M}, k\in\mathcal{K}_R, t\in\mathcal{T}_i. \tag{14f}$$

A two-sided matching algorithm is developed to solve the CUSA subproblem in Section IV. After scheduling MEC resources, remaining users (i.e., V2I users) offload tasks to the macrocell station through the cellular spectrum based on NOMA. Herein, the main focus is how to guarantee that the SINR of every V2I user is above threshold $\Gamma^{\min}$ [i.e., constraint (13h)]. In order to tackle this obstacle, we reformulate the DCSA problem as a Markov decision problem as follows.

In DCSA, each V2I user is considered as an agent, and cellular channels are selected based on the observed instantaneous system states. Generally, the channel condition of the next subframe is only relevant to the current situation and the selected action. The obtained rewards depend only on the system state and the action taken by the V2I user. It can be observed that the procedure of DCSA follows the Markov decision process. We define the long-term reward of V2I user i as

$$\mathcal{R}_i(t)=-\sum_{t=1}^{T_i}(\gamma^t)\mathcal{U}_i(t) \tag{15}$$

where discount factor $\gamma\in[0,1)$, making a tradeoff between the immediate reward and the future reward. It is a constant hyperparameter, which can be viewed as the weight of the reward at time slot t. Generally, recent rewards are more valuable than long-term rewards. Thus, the weight of the reward at time slot t is set as γ^t, which decreases with the increasing of time t. Since the optimization object is to maximize the total rewards of all agents, we define the long-term reward as the weighted summation of the opposite number of the offloading cost.

Algorithm 1: The Pseudo-code of DJOA.

Input: τ_i, $h_{i,j}$, $p_{i,j}$, $\mathcal{K}_I\cup\mathcal{K}_R$, $i\in\mathcal{N}\cup\mathcal{M}, j\in\mathcal{J}$
Output: Spectrum allocation result
Allocate unlicensed spectrum by dividing all users into V2R and V2I users with **Algorithm 2**.
foreach *V2I user* **do**
  Allocate cellular spectrum with **Algorithm 3**.
end

IV. DRL-Enabled Distributed Offloading

In this section, a two-sided matching algorithm and a distributed DRL method are proposed to solve CUSA and DCSA, respectively. We first present the whole offloading procedure; then, details of two subproblems are illustrated.

A. Offloading Procedure

Given offloading task τ_i, channel gain $h_{i,j}$, transmission power $p_{i,j}$, and communication channels $\mathcal{K}_I\cup\mathcal{K}_R$, the distributed joint offloading algorithm (DJOA) is illustrated in Algorithm 1. Unlicensed spectrum is first allocated by dividing all users into V2R and V2I users following a two-sided matching algorithm. Then, a distributed DRL algorithm is developed to schedule cellular channels.

B. Two-Sided Matching Algorithm

We initialize the offloading cost of all users by assuming that they occupy one cellular channel based on OFDM (i.e., there is no interference among users) and upload the cost information to the macrocell to allocate unlicensed channels in a centralized approach, which can also be viewed as choosing $|\mathcal{M}|$ V2R users from all $\mathcal{N}\cup\mathcal{M}$ users. Note $|\mathcal{M}|$ can be any positive integer less than J.

We define two kinds of preference lists to schedule the priority of users. Since the optimization object is to minimize the cost of users, users with higher costs ought to have higher priorities to be accepted. Thus, the cost preference list is initialized in a descending order of the offloading cost. If RSU k receives multiple proposals, it always provides services to the individual with the highest cost priority. In addition, users construct the

Algorithm 2: The Pseudo-code of Two-Sided Matching Algorithm.

```
Initialize the offloading cost.
Initialize the cost preference list C.
Initialize the V2R users set M = ∅.
foreach user i do
    Compute the execution latency L_{i,j} if it is accessed
    to RSU j.
end
Construct latency preference list L_i.
while |M| < J do
    foreach user i ∈ C do
        user i proposes itself to L_i[1].
        Remove L_i[1] from L_i
    end
    foreach RSU k has receieved any proposal do
        if RSU k has not been allocated then
            Accept user i' with the highest cost priority.
            Reject other users.
        end
        else
            if current option < user i' then
                Accept user i'.
                Reject other users.
            end
            else
                Reject all requests.
            end
        end
    end
    Until no more users send offloading requests.
end
```

latency preference list in a descending order of latency $L_{i,j}$, reflecting users' preferences for RSUs. If latency $L_{i,j}$ exceeds $T_i^{\max}$, RSU j is removed from the preference list. The match loops until no more users send offloading requests or all RSUs have been allocated. Details are presented in Algorithm 2.

C. Distributed DRL

Different from traditional centralized DRL, where the macrocell is selected as the agent to make offloading decisions, we develop a distributed DRL method by considering multiple agents (i.e., V2I users). The cumulative reward of one V2I user is inevitably influenced by various channel states and other users' actions. Note that the behaviors of other users cannot be observed in the distributed manner. Instead of sending complete messages (e.g., task information, channel states, and selected policies) to the central macrocell traditionally, each user iteratively sends only one binary bit indicator to the central macrocell. Then, the macrocell broadcasts the received messages so that global state information can be obtained by all individuals. The communication cost incurred by information exchange is negligible. In the hierarchical cellular network, the macrocell server is responsible for global information acquisition and each user acts as an agent. Next, we formulate the system state, action, and reward in detail.

1) *System state:* We define the interference degree of all users as state $s(t)$, represented by

$$s(t) = \{s_1(t), s_2(t), \ldots, s_N(t)\} \tag{16}$$

where $s_i(t) \in \{0, 1\}$. Variable $s_i(t) = 0$ indicates that the interference of users exceeds the interference threshold (i.e., $\Gamma_{i,0,k} < \Gamma^{\min}$), and $s_i(t) = 1$ means that the interference degree is satisfied with constraint (13h). Even though we simplify the system state to a binary variable, the space of possible states can still be very large with such a large number of V2I users (i.e., 2^N for N users).

2) *System action:* It is considered that all users select cellular channels for data transmission, and the system action of each user can be defined as

$$a_i(t) = \{\theta_{i,0,1}(t), \theta_{i,0,2}(t), \ldots, \theta_{i,0,K_I}(t)\} \tag{17}$$

where $\theta_{i,0,k}(t) \in \{0, 1\}$. The searching space of possible actions is 2^{K_I} for K_I cellular channels.

3) *Reward:* User i obtains an immediate reward when it executes action $a_i(t)$ by observing system state $s(t)$. We define the long-term reward as (15). The immediate reward is reformulated based on the observed states as follows:

$$R_i(t) = \begin{cases} -\mathcal{U}_i(t), & \text{if } \Gamma_{i,0,k} \geqslant \Gamma^{\min} \\ -\infty, & \text{otherwise} \end{cases}. \tag{18}$$

This reward function guarantees the maximum long-term reward (i.e., the minimum offloading cost), while satisfying the interference constraint.

The long-term reward obtained by each user is the expectation of the cumulative discounted immediate reward, represented by

$$\mathcal{R}_i^{\pi}(s) = \mathbb{E}\left[\sum_{t=0}^{+\infty} \gamma^t R_i\left(s(t), a_i(t)\right) | s(0) = s, a(0) = a_i\right]. \tag{19}$$

The agent selects actions based on policy π, which maps system states to actions.

The policy is evaluated by Q-learning, where the Bellman equation is utilized to obtain the optimal Q-value function

$$Q_i^*(s, a_i) = \mathbb{E}(R_i(s, a_i)) + \gamma \sum_{s' \in \mathcal{S}} P_{ss'}(a_i) \max_{a_i' \in \mathcal{A}_i} Q_i^*(s', a_i') \tag{20}$$

where $\mathcal{S}$ and $\mathcal{A}$ are feasible sets of system states and actions, respectively. Variable $P_{ss'}(a_i)$ is the transition probability from the current state s to the next state s'. The optimal policy intends to maximize the Q-value, i.e.,

$$\pi_i^*(s) = \max_{a_i' \in \mathcal{A}_i} Q_i^*(s, a_i'). \tag{21}$$

In a realistic scenario, it is difficult to obtain the precise transition probability. Thus, the optimal Q-value is obtained recursively based on available system states, actions, and rewards.

The iterative formula is as follows

$$lQ_i\left(s, a_i\right) = (1-\delta)Q_i\left(s, a_i\right) + \delta\left[u_i\left(s, a_i\right) + \gamma \max_{a_i' \in A_i} Q_i\left(s', a_i'\right)\right] \tag{22}$$

where δ is the learning rate that affects the convergence rate of the iteration process.

Based on the formulated Markov decision process, the distributed reinforcement learning method can be investigated. In traditional Q-learning, it is challenging to obtain the optimal Q-value through searching the large Q-value table. To deal with the storage issue, a deep neural network (DNN) is introduced to approximate the Q-value function in an artificial intelligent manner (i.e., $Q_i\left(s, a_i; \theta\right) \approx Q_i^*\left(s, a_i\right)$, where θ indicates the network parameter of the DNN).

In a deep Q-learning network (DQN), DNN θ is leveraged for both action selection [i.e., (21)] and Q-value approximation [i.e., (22)], resulting in overestimation. A deep double Q-learning network (DDQN) is proposed to mitigate this problem, where evaluated and target DNNs are initialized with θ and θ^-, respectively. The former is in charge of action selection and evaluation. The latter approximates the Q-value function, which can be obtained by

$$y_i^{\mathrm{DDQN}} = R_i\left(s, a_i\right) + \gamma Q_i\left(s', \arg\max_{a_i' \in \mathcal{A}_i} Q_i\left(s', a_i'; \theta\right); \theta^-\right). \tag{23}$$

For parameter updating, the loss function is defined as follows [24]:

$$L_i(\theta) = \mathbb{E}\left[\left(y_i^{\mathrm{DDQN}} - Q_i\left(s, a_i; \theta\right)\right)^2\right] \tag{24}$$

where the parameter of target DNN θ^- is actually a backup of θ before the evaluated DNN updates.

The procedure of DDQN is illustrated as follows. For each episode (i.e., subframe), V2I users select actions following the ε-greedy policy. Then, the immediate reward of the chosen action and the next system state can be obtained through global information synchronization. The DQN introduces experience replay buffer to store the transition, which consists of the current state, action, reward, and next state. In order to eliminate coupling among training data, each user updates DNN parameters by sampling minibatch transitions from the replay buffer. Finally, the target DNN and random probability are updated periodically.

V. Performance Evaluation

The performance of the DJOA is evaluated in this section. The simulation setup is first illustrated. Then, we analyze the simulation results.

A. Simulation Setup

With the objective of validating the performance gained by the DJOA, traces of taxies in Hongkou district, Shanghai (China), are leveraged for evaluation. As illustrated in Fig. 2, a practical situation is considered by utilizing the map of Hongkou district, Shanghai. There are one macrocell station and five RSUs deployed in the region to provide offloading services. GPS locations of the macrocell and RSUs are illustrated in Table I. There are two reasons for the selection of GPS locations. First, in order to avoid data errors in local areas, RSUs need to be scattered within Hongkou district. Second, to ensure that the amount of data are large enough, RSUs need to be deployed near streets with high traffic density. The data size of task offloading is between 20 and 120 MB. The value of q is set as 3, i.e., more than three users occupying one subchannel can cause serious interference. To implement the proposed DJOA algorithm, Tensorflow 0.12.1 is employed with python anaconda 4.3 on Ubuntu 16.04 LTS.

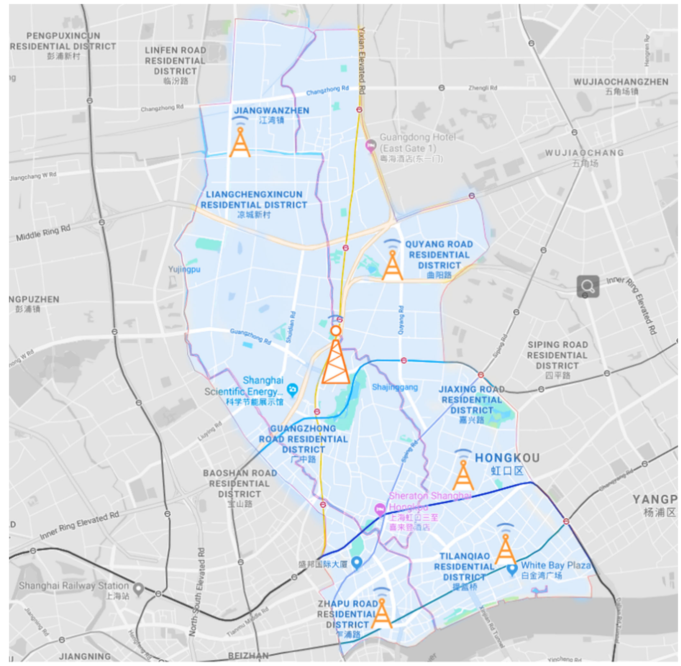


Fig. 2. Map of Hongkou District in Shanghai (China) for evaluation.

TABLE I
GPS Locations of the Macrocell and RSUs in Hongkou District

Server ID	GPS locations
Macro-cell	31.280, 121.490
RSU 1	31.250, 121.490
RSU 2	31.258, 124.515
RSU 3	31.268, 121.503
RSU 4	31.291, 121.497
RSU 5	31.304, 121.467

To demonstrate the effectiveness of our proposed DJOA, the following schemes are leveraged for comparison.

1) *DQN:* The traditional DQN method constructs one DNN to select and evaluate actions, which results in overestimation inevitably.
2) *Q-learning:* It is a classic value-based reinforcement learning method. The defined Q-value function always chooses the action that obtains the largest reward based on the recorded Q table.

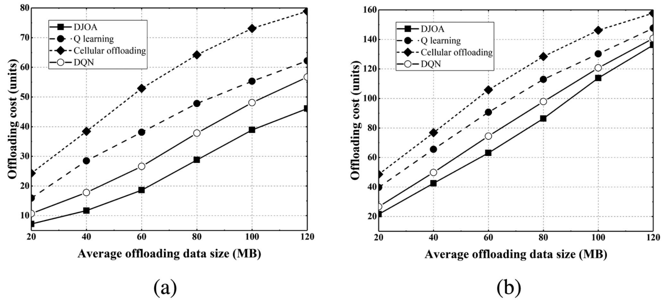


Fig. 3. Offloading cost with different data sizes. (a) $|\mathcal{N} \cup \mathcal{M}| = 10$. (b) $|\mathcal{N} \cup \mathcal{M}| = 20$.

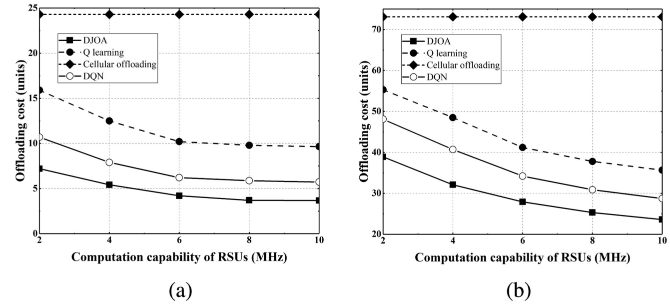


Fig. 4. Offloading cost with different computation capabilities of RSUs. (a) d_m = 20 MB. (b) d_m = 100 MB.

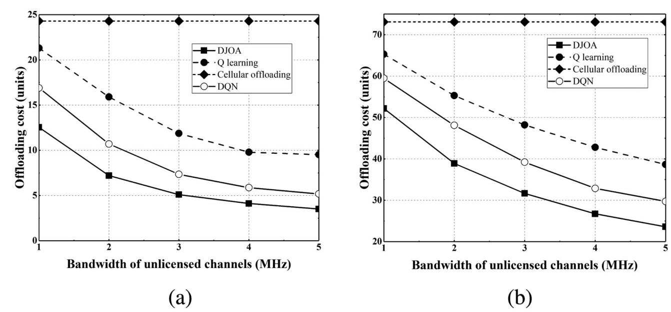


Fig. 5. Offloading cost with different bandwidth of unlicensed channels. (a) d_m = 20 MB. (b) d_m = 100 MB.

3) *Cellular offloading:* A baseline algorithm, where all users offload their messages to the macrocell through cellular spectrum.

B. Performance Evaluation

The following performance comparison is provided based on two scenarios.

Based on different user densities (i.e., ten and 20 users within the wireless coverage of the macrocell), Fig. 3 illustrates the performance of offloading costs varying with different data sizes, where the data size increases from 20 to 120 MB. Since large amounts of data require much time and computing resources, the latency constraint may be violated when tasks are assigned to RSUs. Therefore, more users choose the macrocell when their task sizes are large. Correspondingly, the overall offloading cost rises gradually with the increase of the data size. When the user density is relatively low (i.e., $|\mathcal{N} \cup \mathcal{M}| = 10$), the DJOA can reduce the offloading cost by 60% on average compared with the cellular offloading scheme. As the average offloading data size increases, the ratio decreases from 70% to 40%. It is because that most users cannot access RSUs to guarantee their QoS. When the density of users is high [i.e., Fig. 3(b)], the performance of our proposed DJOA is close to that of the cellular offloading scheme, especially when the offloading burden is very heavy. This is because only a few users can offload through RSUs in this scenario, and available offloading schedules are limited, declining the performance of the DJOA.

Performance comparisons with different computation capabilities of RSUs are provided in Fig. 4. Since the computation capability and the offloading data size jointly influence the total execution time, situations for light and heavy offloading tasks are illustrated in Fig. 4(a) and (b), respectively. With the increase of computation capabilities, RSUs are able to process tasks with large data sizes, while satisfying the latency constraint. Thus, the total offloading cost declines. Note that in Fig. 4(a), when the computation capability exceeds 6 MHz, the offloading cost becomes stable. The reason is that when the offloading task size is relatively small (e.g., d_m = 20 MB), 6-MHz computation capability can consume less time than the latency constraint to finish processing all tasks. Extra computing resources no longer have gains. Correspondingly, the heavy load requires much computing resources, and the rising of computation capability exactly meets this demand. Therefore, the offloading cost decreases steadily when the computation capability increases in Fig. 4(b). In summary, when the offloading burden is relatively light, it does not need to deploy more MEC servers to decrease the offloading cost. In contrast, it is encouraged to guarantee users' QoS by deploying more computing resources when the data size of task offloading is relatively large for processing.

The effect of different bandwidths of unlicensed channels on performance is illustrated in Fig. 5. The bandwidth of unlicensed channels influences the transmission delay of V2R users. The transmission rate between V2R users and RSUs can be very low when the bandwidth is small, resulting in the validation of latency constraint. Thus, the offloading cost is large when the bandwidth is merely 1 MHz in Fig. 5(a) and (b). With the increase of bandwidths, RSUs can easily provide offloading services for V2R users with the transmission time lower than the constraint. It is striking that the decrease ratio of the offloading cost drops when the bandwidth grows larger than 3 MHz in Fig. 5(a). In contrast, the offloading cost continues to decline in Fig. 5(b). This is because 3-MHz bandwidth is sufficient for a light offloading burden (i.e., d_m = 20 MB), while a heavy load needs relatively large bandwidth. In addition, when the bandwidth is insufficient (e.g., 1 MHz), the proposed DJOA can reduce the offloading cost by 50% in Fig. 5(a), but merely 30% in Fig. 5(b). It can be concluded that more bandwidth resources need to be allocated to V2R users with heavy offloading tasks.

VI. Conclusion

This article constructed an intelligent offloading framework for 5G-enabled vehicular networks. An offloading cost minimization problem was formulated while satisfying the latency constraint of individuals. Due to the complexity of the formulated problem, it was divided into two subproblems. For the first one, a two-sided matching algorithm was proposed to allocate the unlicensed spectrum, while a DRL-based method was developed to schedule cellular channels. In order to realize distributed offloading scheduling, we simplified the system state, which can greatly decrease the communication overhead between vehicles and the macrocell. Traces of taxies in Shanghai were leveraged to illustrate the effectiveness of our solution.

References

[1] X. Wang, Z. Ning, and L. Wang, "Offloading in Internet of vehicles: A fog-enabled real-time traffic management system," *IEEE Trans. Ind. Inform.*, vol. 14, no. 10, pp. 4568–4578, Oct. 2018.

[2] Z. Ning, F. Xia, N. Ullah, X. Kong, and X. Hu, "Vehicular social networks: Enabling smart mobility," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 49–55, May 2017.

[3] A. Gulati, G. S. Aujla, R. Chaudhary, N. Kumar, and M. S. Obaidat, "Deep learning-based content centric data dissemination scheme for Internet of vehicles," in *Proc. IEEE Int. Conf. Commun.*, Kansas City, MO, USA, 2018, pp. 1–6.

[4] C. V. N. Index, *Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper*. San Jose, CA, USA: Cisco, 2017.

[5] M. Chen, Y. Qian, Y. Hao, Y. Li, and J. Song, "Data-driven computing and caching in 5G networks: Architecture and delay analysis," *IEEE Wireless Commun.*, vol. 25, no. 1, pp. 70–75, Feb. 2018.

[6] Z. Xiao, L. Zhu, J. Choi, P. Xia, and X.-G. Xia, "Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter wave communications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2961–2974, May 2018.

[7] K. Zhang, S. Leng, Y. He, S. Maharjan, and Y. Zhang, "Cooperative content caching in 5G networks with mobile edge computing," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 80–87, Jun. 2018.

[8] Z. Ning, X. Wang, F. Xia, and J. J. P. C. Rodrigues, "Joint computation offloading, power allocation, and channel assignment for 5G-enabled traffic management systems," *IEEE Trans. Ind. Inform.*, vol. 15, no. 5, pp. 3058–3067, May 2019.

[9] Z. Ning *et al.*, "A cooperative quality-aware service access system for social Internet of vehicles," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2506–2517, Aug. 2018.

[10] X. Chen, L. Jiao, W. Li, and X. Fu, "Efficient multi-user computation offloading for mobile-edge cloud computing," *IEEE/ACM Trans. Netw.*, vol. 24, no. 5, pp. 2795–2808, Oct. 2015.

[11] Z. Ning, P. Dong, X. Kong, and F. Xia, "A cooperative partial computation offloading scheme for mobile edge computing enabled Internet of Things," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4804–4814, Jun. 2019.

[12] C. You, K. Huang, H. Chae, and B.-H. Kim, "Energy-efficient resource allocation for mobile-edge computation offloading," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1397–1411, Mar. 2017.

[13] S. Sardellitti, G. Scutari, and S. Barbarossa, "Joint optimization of radio and computational resources for multicell mobile-edge computing," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 1, no. 2, pp. 89–103, Jun. 2015.

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

[15] Z. Ning *et al.*, "Deep learning in edge of vehicles: Exploring tri-relationship for data transmission," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2019.2929740.

[16] A. Jindal, G. S. Aujla, N. Kumar, R. Prodan, and M. S. Obaidat, "DRUMS: Demand response management in a smart city using deep learning and SVR," in *Proc. IEEE Global Commun. Conf.*, 2018, pp. 1–6.

[17] G. Preethi, P. V. Krishna, M. S. Obaidat, V. Saritha, and S. Yenduri, "Application of deep learning to sentiment analysis for recommender system on cloud," in *Proc. Int. Conf. Comput., Inf. Telecommun. Syst.*, 2017, pp. 93–97.

[18] A. Jindal, G. S. Aujla, N. Kumar, R. Chaudhary, M. S. Obaidat, and I. You, "SeDaTiVe: SDN-enabled deep learning architecture for network traffic control in vehicular cyber-physical systems," *IEEE Netw.*, vol. 32, no. 6, pp. 66–73, Nov./Dec. 2018.

[19] Z. Ning, F. Xia, X. Hu, Z. Chen, and M. S. Obaidat, "Social-oriented adaptive transmission in opportunistic Internet of smartphones," *IEEE Trans. Ind. Inform.*, vol. 13, no. 2, pp. 810–820, Apr. 2017.

[20] X. Wang *et al.*, "Optimizing content dissemination for real-time traffic management in large-scale Internet of vehicle systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1093–1105, Feb. 2019.

[21] B. Di, L. Song, and Y. Li, "Sub-channel assignment, power allocation, and user scheduling for non-orthogonal multiple access networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7686–7698, Nov. 2016.

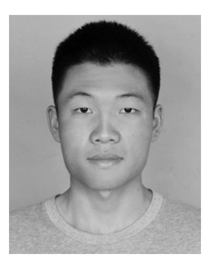
[22] L.-W. Chen and C.-C. Chang, "Cooperative traffic control with green wave coordination for multiple intersections based on the Internet of vehicles," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 7, pp. 1321–1335, Jul. 2017.

[23] Z. Ning, P. Dong, X. Wang, J. J. Rodrigues, and F. Xia, "Deep reinforcement learning for vehicular edge computing: An intelligent offloading system," *ACM Trans. Intell. Syst. Technol.*, to be published, doi: 10.1145/3317572.

[24] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

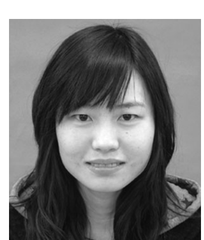
Zhaolong Ning (M'14–SM'18) received the M.S. and Ph.D. degrees in communication and information system from Northeastern University, Shenyang, China, in 2011 and 2014, respectively.

He was a Research Fellow with Kyushu University, Fukuoka, Japan. He is currently an Associate Professor of Network Engineering with the School of Software, Dalian University of Technology, Dalian, China, and an Adjunct Professor of Computer Science with School of Information Science and Engineering, Lanzhou University, Lanzhou, China. He has authored or coauthored more than 100 scientific papers in international journals and conferences. His research interests include Internet of vehicles, edge computing, and artificial intelligence.



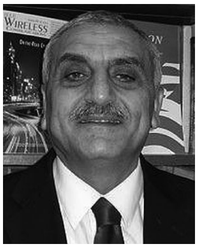
Peiran Dong received the B.S. degree in network engineering from the School of Software, Dalian University of Technology, Dalian, China, in 2018, where he is currently working toward the M.S. degree in network engineering.

His research interests include mobile edge computing, artificial intelligence, and resource management.



Xiaojie Wang received the M.S. degree from Northeastern University, Shenyang, China, in 2011, and the Ph.D. degree from the Dalian University of Technology, Dalian, China, in 2014, both in communication and information system.

From 2011 to 2015, she was a Software Engineer with NeuSoft Corporation, Shenyang. She has authored or coauthored more than 30 scientific papers in her research interests which include vehicular networks, edge computing, and resource management.



Mohammad S. Obaidat (S'85–M'86–SM'91–F'05) received the Ph.D. degree in computer engineering with a minor in computer science from the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH, USA, in 1988.

He is currently a Tenured Full Professor with the King Abdullah II School of Information Technology, The University of Jordan, Amman, Jordan; the China Ministry of Education Distinguished Overseas Professor with the University of Science and Technology Beijing, Beijing, China; and an Honorary Distinguished Professor with Amity University—A Global University, Greater Noida, India. He has chaired numerous (more than 160) international conferences and has given numerous (more than 160) keynote speeches worldwide. He founded or cofounded four international conferences. He has received extensive research funding and published to date (2018) more than 850 refereed technical articles (about half of them are journal articles), more than 65 books, and more than 65 book chapters. He is the Editor-in-Chief of three scholarly journals and an Editor of many other international journals.

Dr. Obaidat is the founding Editor-in-Chief of *Security and Privacy*. He has served on the IEEE Computer Society (CS) Fellow Evaluation Committee. He has served as the IEEE CS Distinguished Speaker/Lecturer and an Association for Computing Machinery (ACM) Distinguished Lecturer. He is a Life Fellow of the IEEE and a Fellow of the Society for Modeling & Simulation International.

Xiping Hu received the Ph.D. degree in computer science from the University of British Columbia, Vancouver, BC, Canada, in 2015.

He is currently a Professor with Lanzhou University, Lanzhou, China. He was the Co-Founder and Chief Technical Officer (CTO) of Bravolol Ltd., Hong Kong, a leading language learning mobile application company with more than 100 million users and listed as the top two language education platform globally. He has more than 70 papers published and presented in prestigious conferences and journals. His research interests include mobile cyber-physical systems, crowdsensing, and social networks.

Lei Guo received the Ph.D. degree in communication and information system from the University of Electronic Science and Technology of China, Chengdu, China, in 2006.

He is currently a Full Professor with the School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China. He has authored or coauthored more than 200 technical papers in international journals and conferences. He is an Editor for several international journals. His current research interests include communication networks, optical communications, and wireless communications.

Yi Guo received the Medical and M.Med. degrees from Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, in 1985 and 1991, respectively, and the Ph.D. degree from the University of Greifswald, Greifswald, Germany, in 1997, all in communication and information system.

He is currently the Chief of Neurology with the Second Clinical Medical College, Jinan University, Guangzhou, China, the Chairman of the Shenzhen Medical Association of Neurology and the Shenzhen Medical Association of Psychosomatic Medicine. His major research interests include cerebrovascular diseases, dementia, movement disorder diseases, sleep disorder, depression, and anxiety.

Jun Huang received the Ph.D. (Hons.) degree in communication and information system from the Institute of Network Technology, Beijing University of Posts and Telecommunications, Beijing, China, in 2012.

He is currently a Full Professor of Computer Science with the Chongqing University of Posts and Telecommunications, Chongqing, China. He has authored more than 100 publications including papers in prestigious journal and conferences. His current research interests include network optimization and control, machine-to-machine communications, and the Internet of Things.

Bin Hu received the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of sciences, Beijing, China, in 1998.

He is currently a Professor with School of Information Science and Engineering, Lanzhou University, Lanzhou, China. He has authored or coauthored more than 100 papers in peer-reviewed journals, conferences, and book chapters. He has served as Chairs/Co-Chairs in many IEEE international conferences/workshops, and an Associate Editor in peer-reviewed journals, such as the IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, *Brain Informatics*, *IET Communications*, etc.

Prof. Hu is a Fellow of the Institution of Engineering and Technology. He is the Cochair of the IEEE Systems, Man, and Cybernetics Society Technical Committee on Cognitive Computing, a Member at Large of Association for Computing Machinery (ACM) China, and the Vice-President of the International Society for Social Neuroscience (China Committee).

Ye Li (SM'15) received the B.Sc. and M.Sc. degrees from the University of Electronic Science and Technology of China, Chengdu, China, in 1995 and 2002, respectively, and the Ph.D. degree from Arizona State University, Tempe, AZ, USA, in 2006, all in electric engineering.

He is currently a Full Professor with the Research Center for Biomedical Information Technology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. He has published more than 140 papers, including 51 journal papers, as the first or corresponding author. His research interests include wearable sensing and computing, low-power wireless communication, and body area networks.

Dr. Li serves as the Area Editor of *Information Fusion* and a Guest Editor of the IEEE SENSORS JOURNAL.