Deep Learning Empowered Task Offloading for Mobile Edge Computing in Urban Informatics

Ke Zhang, Yongxu Zhu, *Member, IEEE*, Supeng Leng, *Member, IEEE*, Yejun He, *Senior Member, IEEE*, Sabita Maharjan, *Member, IEEE*, and Yan Zhang, *Senior Member, IEEE*

Abstract—Led by industrialization of smart cities, numerous interconnected mobile devices, and novel applications have emerged in the urban environment, providing great opportunities to realize industrial automation. In this context, autonomous driving is an attractive issue, which leverages large amounts of sensory information for smart navigation while posing intensive computation demands on resource constrained vehicles. Mobile edge computing (MEC) is a potential solution to alleviate the heavy burden on the devices. However, varying states of multiple edge servers as well as a variety of vehicular offloading modes make efficient task offloading a challenge. To cope with this challenge, we adopt a deep Q-learning approach for designing optimal offloading schemes, jointly considering selection of target server and determination of data transmission mode. Furthermore, we propose an efficient redundant offloading algorithm to improve task offloading reliability in the case of vehicular data transmission failure. We evaluate the proposed schemes based on real traffic data. Results indicate that our offloading schemes have great advantages in optimizing system utilities and improving offloading reliability.

Index Terms—Offloading, Q-learning, reliability, vehicular edge computing.

I. INTRODUCTION

ALONG with the advancement of Internet of Things in industrial application scenarios, urban life pattern
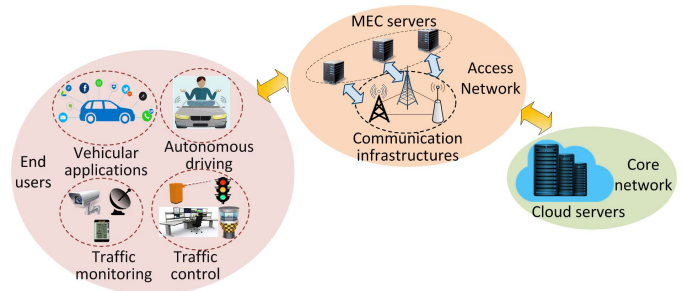
Fig. 1. MEC-enabled intelligent traffic applications.

is undergoing a tremendous change [1]. A high level of interconnections between heterogenous smart devices bring in the possibility to enable industrial automation, which improves operation and increases productivity with less or without human work [2].

Autonomous driving is one of the most attractive industrial automation applications. With large number of on-board sensors and actuators as well as advanced control systems, autonomous vehicles are capable of interpreting sensory information and identifying appropriate navigation paths. In addition, with the aid of information infrastructure in urban area and the introduction of intelligent transportation system (ITS), smart vehicles facilitate us with a pervasive and promising platform to realize a broad range of novel mobile applications, such as augmented reality, natural language processing, and interactive gaming [3]. However, the process of understanding highly dynamic and complex traffic environment while making real-time driving decisions involves processing a great volume of sensory data and requires intensive computations. Due to the constraint of on-board computation power, supporting these real-time and computationally intensive tasks and applications on vehicles is a big challenge.

Mobile edge computing (MEC), where heavy computation tasks are offloaded to the cloud resources placed at the edge of mobile networks, has emerged as a promising approach to cope with growing computing demands [4]. Fig. 1 represents a typical scenario of applying MEC technique for intelligent traffic applications. Aided with MEC technique, safety-oriented tasks in the context of autonomous driving, computation-intensive vehicular applications, and traffic sensory data analysis can be offloaded from vehicles to MEC servers on various types of

proximate wireless access infrastructure, where real-time data processing and feedback can be achieved [5].

In MEC process, the implementing efficiency is tightly coupled with the scheduling and management of computing and communication resources [6]. With the harmonization of the global deployment of Long Term Evolution (LTE) systems, LTE-V is a paradigm in vehicular communications and plays a vital role in the design of task offloading schemes. However, inherent characteristics of vehicular networks, such as high speed mobility, time-varying topology, and ephemeral interactions, bring unprecedented challenges in managing vehicular communication applied with MEC applications. In the context of vehicular networks with multiple MEC servers, the mutual effects between transmission mode selection and offloading target server determination make MEC scheduling even more complex. Furthermore, transmission may fail during vehicular offloading process, which limits the performance and applicability of MEC services. Thus, novel solutions are necessary to address this issue in order to guarantee both reliability and efficiency of the vehicular task offloading. However, very few works have investigated integrating management of computing and communication resources in a multiserver vehicular edge computing network, and task offloading reliability has not been incorporated in the recent literatures.

To bridge this gap, in this paper, we focus on task offloading in an MEC-enabled vehicular network, and present an approach to optimize MEC system performance while also improving offloading reliability. The main contributions of this paper are as follows.

1) We present an MEC-enabled LTE-V network, where the influence of various vehicular communication modes on the task offloading performance is qualitatively analyzed.
2) By applying deep Q-learning approach, we propose optimal target MEC server determination and transmission mode selection schemes, which maximize the utilities of offloading system under given delay constraints.
3) To cope with transmission failure in vehicular networks, we design an efficient redundant offloading algorithm, which also ensures offloading reliability while improving the gained utilities.

The remainder of this paper is organized as follows. In Section II, we review related work. A vehicular edge computing system model is presented in Section III. A deep Q-learning-based offloading scheme is described in Section IV. In Section V, we investigate offloading reliability. Performance evaluation is presented in Section VI. Finally, we conclude this paper in Section VII.

II. Related Work

To meet the demands from computationally intensive vehicular applications, some studies have investigated applying MEC approach in vehicular networks. Shojafar *et al.* [7] proposed an energy efficient resource allocation for vehicular fog computing centers. In [8], an MEC-based architecture was used in urban traffic management in a distributed and adaptive service manner. Sookhak *et al.* [9] designed a fog vehicular computing framework that integrates resources from both edge server and remote cloud. Hou *et al.* [10] unveiled underutilized vehicular computing resources, and put them into use for providing efficient computational support to MEC servers. To efficiently merge MEC technology in vehicular networks, Qiao *et al.* [11] introduced a collaborative task offloading and output transmission mechanism. Aissioui *et al.* [12] designed an MEC service migration scheme that ensures vehicles always connect to the nearest MEC entities. Although these studies have provided some insights about MEC-enabled vehicular applications, the effects of vehicular communication on the design of task offloading strategies have not been thoroughly investigated.

Benefiting from fast commercialization of LTE system, LTE-V has been one of the key technologies in vehicular networks. Several recent works have focused on analytical models and implementations of LTE-V. In [13], direct vehicle communication was utilized to offload data transmission from vehicles with poor quality link to infrastructures. Taking into account high mobility of vehicles, Perfecto *et al.* [14] designed a wireless link formulation mechanism, where beamwidths between vehicular communication pairs were optimized. Shah *et al.* [15] discussed key building blocks of 5G networks in the context of vehicular communications. However, joint V2I and vehicle-to-vehicle (V2V) transmission schemes in a multiple MEC server scenario have not been considered in the previous studies.

Learning is a branch of artificial intelligence, which studies systems and acquires knowledge from data. Recently, various learning techniques have been deployed for scheduling task offloading. Xu *et al.* [16] proposed an online learning-based workload offloading scheme in MEC systems with renewable power supply. In order to reduce resource consumption in task offloading, Crutcher *et al.* [17] formalized intelligent offloading metric prediction utilizing a machine learning-based approach. Deep Q-learning is a powerful tool in policy optimization, and has been utilized in various process decisions. For instance, He *et al.* [18] designed an integrated resource management scheme for connected vehicles using a deep reinforcement learning approach. Zhang *et al.* [19] used deep Q-learning in scheduling voltage and frequency for real-time systems in embedded devices. In [20], this learning approach was adopted in designing a video streaming framework. Moreover, deep Q-learning also can be used in traffic area. To relieve traffic congestion at highway junctions, Kashihara [21] applied deep Q-learning in traffic simulation study and vehicle pathway optimization. However, the potential of learning-based approaches have not been explored for designing scheduling algorithms for vehicular edge computing applications. Furthermore, mobile characteristics of vehicles and reliability of vehicular task offloading have not been considered in the previous studies.

Different from these studies, in this paper, we concentrate on task offloading in an LTE-V network, and propose optimal offloading schemes that jointly schedule vehicular communication and edge computing through a deep Q-learning approach.
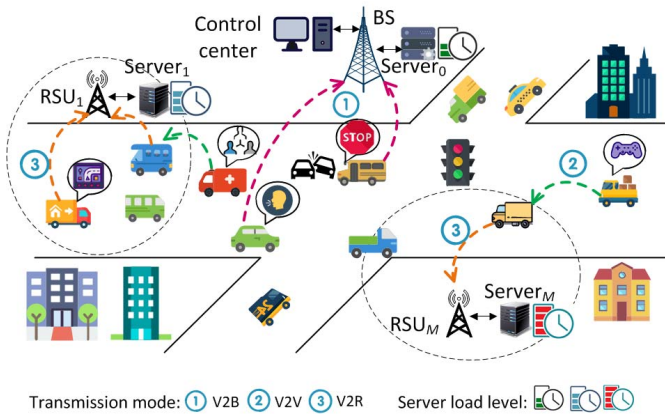
Fig. 2.   Task offloading in an MEC-enabled vehicular network.

TABLE I
MAIN VARIABLES

Variables	Description
G	Number of task types
f_i	Size of task κ_i's input
g_i	Amount of task κ_i's required computation
$t_i^{\max}$	Maximum delay tolerance of task κ_i
β_i	Probability of a task belonging to type-i
ρ	Road traffic density
M	Number of MEC servers
W_m	Computing capacity of MEC server m
c_c	Cost of a unit spectrum of cellular network
c_v	Cost of a unit spectrum of vehicular network
P_g	Probability of generating a task in a time frame
$P_{tx,b}$	Vehicle transmission power in V2B mode
$P_{tx,v}$	Vehicle transmission power in V2V and V2R modes
q_c, q_v	Amount of allocated spectrum resources through cellular network and vehicular network, respectively

III. SYSTEM MODEL

Fig. 2 shows the architecture of an MEC-enabled vehicular network in an urban area. The autonomous navigation of the vehicles requires various sensory data processing. Moreover, urban informatic infrastructure-aided mobile applications may also pose computing requirement on the vehicles. We model these data processing and computing as computation tasks. Various tasks may have different characteristics. For instance, autonomous navigation has strict delay constraints, while entertainment applications do not impose a critical delay requirement. According to this consideration, we classify these tasks into G types. A task is described in four terms as $\kappa_i = \{f_i, g_i, t_i^{\max}, \varsigma_i\}$, $i \in \mathcal{G}$ [22]. Here, f_i and g_i are the size of task input data and the amount of required computation, respectively. To provide timely response in various traffic context, the tasks are time sensitive. $t_i^{\max}$ is the maximum delay tolerance of task κ_i. Offloading system gets utility $\varsigma_i \Delta t$ from the completion of task κ_i, where Δt is the saved time in accomplishing κ_i compared to $t_i^{\max}$. The probability of the tasks belonging to type-i is denoted as β_i with $\sum_{i \in \mathcal{G}} \beta_i = 1$. The main variables used in the system model are shown in Table I.

The road is covered by a heterogeneous vehicular network. Besides the cellular network provided by a base station (BS),

there is an LTE-V network composed of mobile vehicles and M road side units (RSUs) deployed along the road. The set of these RSUs is denoted as $\mathcal{M}$. The cellular network and the vehicular network operate on different and nonoverlapping spectrum. Compared to the BS that has seamless coverage and high data transmission cost, the RSUs conversely provide spotty coverage and inexpensive access service. The costs for using a unit spectrum of the cellular network and for that belonging to the vehicular network in a unit time are c_c and c_v, respectively. We have $c_c > c_v$.

The BS is equipped with an MEC server through wired connections, which is denoted as Serv_0. In addition, each RSU hosts an MEC server. These servers are denoted as Serv_1, Serv_2, $\ldots$, Serv_M, respectively. The MEC servers get data from their attaching BS or RSUs directly. Let $\{W_0, W_1, W_2, \ldots, W_M\}$ denote the computing capacities of these servers, respectively. As the server equipped on BS can serve the vehicles located on the whole road, we consider that the capacity of Serv_0 is much higher than that of the servers deployed on the RSUs. Each MEC server is modeled as a queuing network, where the input is the offloading task. The arrived tasks is first cached in an MEC server, and then served with the first-come-first-serve policy. A server utilizes all of its computing resources to execute currently served task. The cost for tasks using a computing resource in a unit time is c_x.

Each vehicle has a cellular and an LTE-V radio interface, which work on different spectrum and enable multiple communication paradigms. In the heterogeneous network formed by overlapping of the BS and the RSUs coverage, vehicles can offload their tasks to MEC servers through multiple modes. We name the task file transmission between a vehicle and the BS as vehicle-to-BS (V2B). When a vehicle turns to the LTE-V network for task offloading, the file can be transmitted to an MEC server in a mode with joint V2V and vehicle-to-RSU (V2R) transmission.

In self-driving vehicles, real-time vehicle traffic information, such as position, speed, heading directions can be gathered by vehicular sensors [23]. Furthermore, channel state information also can be detected by these vehicles. All this information together with the description of generated vehicular tasks are transmitted to a control center through cellular networks. There is spectrum allocated for this information transmission besides the spectrum used for task offloading. Based on the collected information, the control center can utilize communication resources of the heterogeneous network as well as the computing resources of MEC servers, and efficiently schedule task offloading.

The scheduling and resource management are considered to operate in a discrete time model with fixed length time frames. The length of a frame is denoted as τ. In each time frame, a vehicle generates a computing task with probability P_g. Enabled by advanced LTE technology, we consider that the duration time of a task file transmission is within a time frame. In addition, a task offloading vehicle can only choose one transmission mode.

The communication topology between vehicles and infrastructure keeps constant during one frame. However, the topology may change in different time frames due to the mobility

of the vehicles. To facilitate the modeling of the dynamic relations, we divide the road into E segments. The position of a vehicle on the road can be denoted by the index of the segment e, where $1 \leq e \leq E$. We consider that the vehicles in the same road segment have an identical distance to a communication infrastructure.

In assessing the network performance, we focus on the upstream communication process that offloads tasks from vehicles to MEC servers in various modes. We consider that all vehicles have fixed transmission power for a given transmission mode, i.e., power $P_{tx,b}$ in V2B mode and power $P_{tx,v}$ in V2R and V2V modes. In addition, these vehicles have enough storage for caching task files.

In the case of V2B mode, the assignment of spectrum to vehicles is orthogonal, and there is no collision between V2B communication vehicles. For receiving task file from a V2B mode vehicle, the signal to noise and interference ratio (SINR) at the BS is given as

$$\gamma_{v,b} = P_{tx,b}G_r/L_0 d_{v,b}^\alpha P_w \qquad (1)$$

where $d_{v,b}$ is the distance between the transmitting vehicle and the BS. G_r is the antenna gain at the BS. L_0 and α are the path loss at a reference unit distance and path loss exponent, respectively. P_w is the power of additive white Gaussian noise.

When vehicles choose LTE-V communication in V2R or V2V modes, collisions may occur due to the spectrum reuse between communication pairs working in these modes. In such a case, the SINR at receiver r is calculated as

$$\gamma_{v,r} = \left(P_{tx,v}/L_0 d_{v,r}^\alpha\right)/\left(P_w + \sum_{j\in\mathcal{V}} P_{tx,v}/L_0 d_{j,r}^\alpha\right) \qquad (2)$$

where $\mathcal{V}$ is the set of other vehicles that communicate in the same spectrum within the interference range. The receiver r can be either an RSU or a relay vehicle.

Let γ^{min} be the minimum SINR at a receiver under the premise that the received data can be decoded. Given a static network topology and spectrum resource allocation, we can get the feasible communication pairs whose SINR is no less than γ^{min}. These pairs form the potential way to offload task files from vehicles to MEC servers.

IV. OPTIMAL OFFLOADING SCHEMES IN LEARNING APPROACH

In this section, we formulate an optimal offloading problem, and then model it as a Markov decision process. Based on deep Q-learning, an approach that incorporates deep learning algorithm with Q-functions, joint MEC server selection and offloading mode determination strategies are obtained.

A. Problem Formulation

In a given time frame, for a vehicle that is located in road segment e and generates task κ_i, we use $x_{i,e} = 1$ to indicate the task offloading to Serv_0 through V2B mode. Similarly, we use $y_{i,e,m} = 1$ and $z_{i,e,m} = 1$ to indicate the task offloading to Serv_m in a V2R mode and in a joint V2V and V2R mode, respectively. Otherwise, these indicators are set to 0.

The proposed optimal task offloading problem, which maximizes the utility of the offloading system under the constraints of task delay, is formulated as follows:

$$\max_{\{x,y,z\}} U = \sum_{l=1}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{G} \beta_i \Bigg(\varsigma_i \Big(t_i^{max} - t_{i,e_j,l}^{total} \Big)$$
$$- x_{i,e_j}^l \big(q_c c_c f_i/R_{v,b,e_j} + g_i c_x/W_0 \big)$$
$$- y_{i,e_j,m}^l \big(q_v c_v f_i/R_{v,r,e_j} + g_i c_x/W_m \big)$$
$$- z_{i,e_j,m}^l \Bigg(\sum_{h=1}^{H_{e_j}} q_v c_v f_i/R_{v,j,e_j} + g_i c_x/W_m \Bigg) \Bigg)$$

$$\text{s.t.} \quad x_{i,e_j}^l = \{0,1\}, y_{i,e_j,m}^l = \{0,1\}, z_{i,e_j,m}^l = \{0,1\}$$
$$x_{i,e_j}^l y_{i,e_j,m}^l = x_{i,e_j}^l z_{i,e_j,m}^l = y_{i,e_j,m}^l z_{i,e_j,m'}^l = 0$$
$$x_{i,e_j}^l + y_{i,e_j,m}^l + z_{i,e_j,m}^l = 1$$
$$t_{i,e_j}^{total} \leq t_i^{max}$$
$$i \in \kappa, \quad m, m' \in \mathcal{M} \qquad (3)$$

where, n is the number of tasks generated in a time frame. e_j is the road segment index of vehicle j's location. H_{e_j} is the number of transmission hops. q_c and q_v are the amount of spectrum resources allocated for each task file offloading through cellular network and LTE-V network, respectively. R_{v,b,e_j} is the transmission rate of offloading task file from vehicle at road segment e_j to the BS, which can be given as $R_{v,b,e_j} = q_c \log(1 + \gamma_{v,b})$. R_{v,r,e_j} and R_{v,j,e_j} can be calculated similarly based on the allocated spectrum q_v and SINR $\gamma_{v,r}$.

In (3), the first three constraints indicate that a task file can only be transmitted through one mode. The fourth constraint shows that the time cost for offloading task κ_i should be under its delay constraint. Here, t_{i,e_j}^{total} is the total time cost for completing a type i task generated by a vehicle located at road segment e_j. Given the task offloading strategies, t_{i,e_j}^{total} can be written as

$$t_{i,e_j}^{total} = x_{i,e_j} \Big(f_i/R_{v,b,e_j} + t_0^{wait} + g_i/W_0 \Big)$$
$$+ y_{i,e_j,m} \Big(f_i/R_{v,r,e_j} + t_m^{wait} + g_i/W_m \Big)$$
$$+ z_{i,e_j,m} \Bigg(\sum_{h=1}^{H_{e_j}} f_i/R_{v,j,e_j} + t_m^{wait} + g_i/W_m \Bigg) \Bigg) \qquad (4)$$

where t_0^{wait} and t_m^{wait} are the waiting time of the task in Serv_0 and Serv_m, respectively. The value of the waiting time will be discussed in the following section.

B. Markov Decision Approach

As each MEC server is modeled as a queuing system, the current serving state of a server may affect the time cost for accomplishing the following tasks. To choose the offloading target server efficiently, the offloading strategy taken by each task in time frame l depends on the characteristics of current vehicle network as well as the server states in frame $l - 1$. Thus, we can formulate (3) as a Markov decision process, and solve it in a Markov decision approach [24].

The state of the offloading system at time frame l is defined as $S^l = (s_0^l, s_1^l, \ldots, s_M^l)$, where s_0^l is the total computation

required by the tasks queuing in $Serv_0$ at frame l. Similarly, $s_1^l, \ldots, s_M^l$ denotes the required computation of the tasks queuing in $Serv_1$, $Serv_2$, ..., $Serv_M$ at time frame l, respectively. The actions taken by the control center at frame l can be shown as $a^l = (X^l, Y^l, Z^l)$, where $X^l = \{x_{i,e}^l\}$, $Y^l = \{y_{i,e,m}^l\}$ and $Z^l = \{z_{i,e,m}^l\}$ are the sets of task offloading strategies with various transmission modes and offloading targets for the generated tasks at frame l, respectively.

To facilitate the analysis of the effects brought by the actions to the system states, we introduce variable $\hat{c}_m^l$, $m \in \{0, 1, \ldots, M\}$, which denotes the amount of computation taken by $Serv_m$ in time frame l. We define $\hat{c}_m^l$ as

$$\hat{c}_m^l = \min\left(s_m^l + \sum_{i=1}^{G}\sum_{j=1}^{n}\left(x_{i,e_j}^l + y_{i,e_j,m}^l + z_{i,e_j,m}^l\right)g_i, \ W_m\tau\right).$$
(5)

Then, the state transitions between time frame l and $l+1$ can be written as

$$S^{l+1} = \left(s_0^l + \sum_{i=1}^{G}\sum_{j=1}^{n}x_{i,e_j}^l g_i - \hat{c}_0^l\right.$$
$$s_1^l + \sum_{i=1}^{G}\sum_{j=1}^{n}\left(y_{i,e_j,1}^l + z_{i,e_j,1}^l\right)g_i - \hat{c}_1^l, \ldots$$
$$\left. s_M^l + \sum_{i=1}^{G}\sum_{j=1}^{n}\left(y_{i,e_j,M}^l + z_{i,e_j,M}^l\right)g_i - \hat{c}_M^l\right).$$
(6)

When action a^l is taken in state S^l, the gained average utility in time frame l is

$$U^l = \sum_{j=1}^{n}\sum_{i=1}^{G}\beta_i\left(\varsigma_i\left(t_i^{\max} - t_{i,e_j,l}^{\text{total}}\right)\right.$$
$$- x_{i,e_j}^l\left(q_c c_c f_i/R_{v,b,e_j} + g_i c_x/W_0\right)$$
$$- y_{i,e_j,m}^l\left(q_v c_v f_i/R_{v,r,e_j} + g_i c_x/W_m\right)$$
$$\left.- z_{i,e_j,m}^l\left(\sum_{h=1}^{H_{e_j}}q_v c_v f_i/R_{v,j,e_j} + g_i c_x/W_m\right)\right)$$
(7)

where $t_{i,e_j,l}^{\text{total}}$ is defined in (4) with $t_0^{\text{wait}} = s_0^l/W_0$ and $t_m^{\text{wait}} = s_m^l/W_m$.

In order to maximize the utility of the offloading system, we need to obtain an optimal strategy π^*, which consists of offloading actions for various tasks in different time frames. π^* can now be expressed as

$$\pi^* = \arg\max_\pi E\left(\sum_{l=1}^{\infty}\eta^l U^l\right)$$
(8)

where η^l is a discount factor that trades off the immediate utility and the later ones, $0 < \eta^l \le 1$.

C. Deep Q-learning-Based Offloading Scheme

To derive the optimal offloading strategy π^*, we turn to reinforcement learning technology. Reinforcement learning is a main branch of machine learning, where agents take series of actions that maximize the discounted future reward with corresponding strategies in various states. Thus, a Markov decision process can be considered as a reinforcement learning problem. Under a given offloading strategy π, the gained average system utility from taking action a^l in state S^l can be expressed as a Q-function, which is shown as

$$Q_\pi\left(S^l, a^l\right) = E\left[U^l + \eta U^{l+1} + \eta^2 U^{l+2} + \cdots | S^l, a^l\right]$$
$$= E_{S^{l+1}}\left[U^l + \eta Q_\pi\left(S^{l+1}, a^{l+1}\right) | S^l, a^l\right].$$
(9)

Then, the optimal value of the Q-function will be

$$Q^*\left(S^l, a^l\right) = E_{S^{l+1}}\left[U^l + \eta \max_{a^{l+1}} Q^*\left(S^{l+1}, a^{l+1}\right) | S^l, a^l\right]$$
(10)

where the maximum utility as well as the optimal offloading strategies can be derived by value and strategy iteration. Q-learning, which is a classical algorithm of reinforcement learning technologies, can be used in modifying the iterations. In each iteration, the value of Q-function in the learning process is updated as

$$Q\left(S^l, a^l\right) \leftarrow Q\left(S^l, a^l\right) + \alpha\left[U^l + \eta \max_{a^{l+1}} Q^*\left(S^{l+1}, a^{l+1}\right)\right.$$
$$\left. - Q\left(S^l, a^l\right)\right]$$
(11)

where α is the learning rate.

However, the states of the offloading system consist of the amount of required computation queuing in the MEC servers, whose value is continuous. It is hard to find the optimal solution through discretizing the state space. Thus, Q-learning approach cannot be directly implemented in solving our proposed Markov decision problem. To address this issue, we turn Q-function into a function approximator, which is a function form easy to be handled in optimal action acquisition process. Here, we choose a multilayered neural network as a nonlinear approximator that is able to capture complex interaction among various states and actions. Based on the Q-function estimation, we utilize deep Q-learning technology to obtain the optimal offloading strategies π^* [25].

We refer to the proposed neural network-based approximator as Q-network, where θ is the set of parameters of the network. With the help of Q-network, the Q-function in (9) can be estimated as $Q(S^l, a^l) \approx Q'(S^l, a^l; \theta)$ [26]. Q' is trained to converge to real Q values over iterations. Based on Q', the optimal offloading strategies in each state is derived from the actions that lead to the maximum utility. The chosen action at frame l can now be written as $a^{l*} = \arg\max_{a^l} Q'(S^l, a^l; \theta)$.

In the learning process, experience replay technique is utilized to improve the learning efficiency, where the learning experience at each time frame is stored in a replay memory [25]. The experience consists of observed state transitions as well as gained utilities led by actions. The experience gained at time frame l is expressed as (S^l, a^l, U^l, S^{l+1}). During Q-learning updates, a batch of stored experience drawn randomly from the replay memory is used as samples in training the parameters of Q-network. The goal of the training is to

Algorithm 1 Deep Q-Learning-Based Task Offloading

Initialization:

Initialize Q-network with weights θ, action-value function Q, and experience replay buffer.

1: **For** a given steady vehicular traffic flow **Do**
2:     Observe the initial state S^0;
3:     **For** time frames $l = 0, \ldots, L_{\max}$ **Do**
4:         Select a random action a^l with probability ε, otherwise choose action $a^l = \arg\max_a Q(S^l, a; \theta)$;
5:         Execute action a^l, derive the next state S^{l+1} and obtain offloading utility U^l according to (6) and (7);
6:         Store the experience (S^l, a^l, U^l, S^{l+1}) into the experience replay buffer;
7:         Get a batch of samples from the replay memory, and calculate loss function $Loss(\theta^l)$ according to (12);
8:         Calculate the gradient of $Loss(\theta^l)$ with respect to θ^l according to (14);
9:         Update θ^l according to (15);
10:    **End For**
11: **End For**

---

minimize the difference between $Q(S^l, a^l)$ and $Q'(S^l, a^l; \theta)$. We define a loss function to denote the difference as

$$\text{Loss}\left(\theta^l\right) = E\left[\frac{1}{2}\left(Q_{\text{tar}}^l - Q'\left(S^l, a^l; \theta^l\right)\right)^2\right] \quad (12)$$

where θ^l is the parameters of Q-network at time l. Q_{tar}^l is a learning target, which denotes the optimal value of the Q-function in frame l and can be shown as

$$Q_{\text{tar}}^l = U^l + \eta Q\left(S^l, \arg\max_{a^{l+1}} Q'\left(S^{l+1}, a^{l+1}; \theta^l\right)\right). \quad (13)$$

We deploy a gradient descent approach to modify θ. The gradient derived through differentiating $\text{Loss}(\theta^l)$ is calculated as

$$\nabla_{\theta^l}\text{Loss}\left(\theta^l\right) = E\left[\nabla_{\theta^l}Q'\left(S^l, a^l; \theta^l\right)\left(Q'\left(S^l, a^l; \theta^l\right) - Q_{\text{tar}}^l\right)\right]. \quad (14)$$

Then, θ^l is updated according to

$$\theta^l \leftarrow \theta^l - \varpi \nabla_{\theta^l}\text{Loss}\left(\theta^l\right) \quad (15)$$

where ϖ is a scalar step size.

In order to avoid local optimum while balancing exploration and exploitation in the learning process, we adopt ε-greedy policy. Random selected actions are taken with probability ε to explore better offloading strategies, otherwise optimal actions are chosen in exploitation the replay memory with probability $1 - \varepsilon$, where $0 < \varepsilon < 1$. The proposed deep Q-learning-based offloading scheme is shown in Algorithm 1.

V. RELIABLE OFFLOADING IN PRESENCE OF TRANSMISSION FAILURE

Although the deep Q-learning-based schemes provide an optimal way for offloading tasks with maximum utility, task file transmission failure may seriously undermine the offloading efficiency. Unlike V2B transmission mode, where vehicles communicate in orthogonal spectrum, the vehicle communication pairs in V2V mode may work with the same spectrum resources and cause serious interference between them. In addition, in V2V transmission, wireless interface of some chosen relay vehicles may be suddenly occupied by vehicular applications newly generated in these vehicles during the file delivery. Task transmission as well as offloading process may be endangered or even interrupted by the interference.

In this section, we focus on offloading reliability in a joint V2V and V2R mode. At first, we investigate the number of transmission hops in this joint mode offloading process. Recall that $\gamma_{\min}$ is the minimum SINR at a receiver under the premise that the received data can be decoded, and the maximum distance of one hop vehicular communication without co-channel interference can be calculated as

$$d_{v,r}^{\max} = \left(\frac{P_{tx,v}}{\gamma_{\min}L_0 P_w}\right)^{-\alpha}. \quad (16)$$

Theorem 1: Given traffic density of the vehicles on the road as ρ and the probability of a vehicle generating a task as P_g, the mean value of the farthest communication distance between two vehicles can be approximated as $d_{v,r}'^{\max} = d_{v,r}^{\max} - \Delta d/2$. Δd is distance reduction, and $\Delta d > \ln \sigma/2(1 - P_g)\rho$, where σ is a coefficient and $0 < \sigma \ll 1$.

Proof: See Appendix A. ∎

For a vehicle located in road segment e, when it chooses MEC Serv$_m$ as its offloading target, according to Theorem 1, the number of the transmission hops from the original vehicle to RSU m is

$$H_{e,m} = \lceil d_{e,m}/d_{v,r}^{\max} \rceil. \quad (17)$$

Among the total number of transmission hops, $H_{e,m} - 1$ hops are for V2V communication and the last one hop is for V2R communication.

Next, we analyze the interference brought by vehicles operating in the same frequency to the V2V and V2R communications. As stated above, each transmitting vehicle randomly chooses q_v spectrum from Q_v resources for file delivery. The total number of orthogonal spectral resource blocks for vehicular communication is $n_r = \lfloor Q_v/q_v \rfloor$. We consider vehicle v sends computing task file to receiver r. Here, r can be either a vehicle in V2V mode or an RSU in V2R mode. Let $d_{v,r}$ denote the distance between v and r, and $d_{j,r}$ be the distance between r and its jth nearest interference vehicle. The following lemma presents the statistical distribution of interference distance $d_{j,r}$.

Lemma 1: For vehicular communication, the probability distribution of distance between receiving vehicle r and its jth nearest interference vehicle is shown as

$$f\left(d_{j,r}\right) = \frac{\left(n_r/P_g\rho\right)^j}{(j-1)!}d_{j,r}^{j-1}e^{-n_r d_{j,r}/P_g\rho}. \quad (18)$$

Proof: See Appendix B. ∎

According to Theorem 1 and (2), the total time consumption of file transmission from vehicle v, which generates a type-i,

to RSU m can be calculated as

$$t_{v,m,i}^{\text{total}} = \frac{(H_{e,m} - 1)f_i}{q_v \log_2\left(1 + \frac{P_{tx,v}/L_0(d'^{\text{max}}_{v,r})^\alpha}{P_w + \sum_{j \in \mathcal{V}} P_{tx,v}/L_0 d_{j,r}^\alpha}\right)}$$
$$+ \frac{f_i}{q_v \log_2\left(1 + \frac{P_{tx,v}/L_0(d_{v,m} - (H_{e,m}-1)d'^{\text{max}}_{v,r})^\alpha}{P_w + \sum_{j \in \mathcal{V}} P_{tx,v}/L_0 d_{j,r}^\alpha}\right)}. \quad (19)$$

In (19), interference distance $d_{j,r}$ is a random variable as shown in Lemma 1. Now, we derive the probability distribution of $t_{v,m,i}^{\text{total}}$. To facilitate the analysis of the impact of co-channel vehicular communication on $t_{v,m,i}^{\text{total}}$, we only consider the two nearest interfering vehicles in (19). It is noteworthy that the analysis approach can be extended to a scenario with multiple interfering vehicles.

Lemma 2: x and y are independent random variables with probability distribution $f_x(x)$ and $f_y(y)$, respectively. b_0 and c_0 are positive constants. Let $z = b_0(x^{-\alpha} + y^{-\alpha}) + c_0$. The probability distribution of z can be presented as

$$f_z(z) = \int_{-\infty}^{+\infty} f_x\left[\left(\frac{h - c_0}{b_0}\right)^{-\frac{1}{\alpha}}\right]\left|\frac{-(h - c_0)^{-\frac{1}{\alpha}-1}}{\alpha b_0^{-\frac{2}{\alpha}-1}}\right|$$
$$\times f_y\left[\left(\frac{z - h}{b_0}\right)^{-\frac{1}{\alpha}}\right]\left|\frac{-(z - h)^{-\frac{1}{\alpha}-1}}{\alpha b_0^{-\frac{2}{\alpha}-1}}\right|dh. \quad (20)$$

Proof: See Appendix C. ∎

Let $z = \sum_{j=1}^{2} P_{tx,v}L_0^{-1}d_{j,r}^{-\alpha} + P_w$. Through using Lemmas 1 and 2, we obtain

$$f_z(z) = \left|\frac{1}{\alpha\left(P_{tx,v}L_0^{-1}\right)^{-\frac{2}{\alpha}-1}}\left|\int_{-\infty}^{+\infty} f_{d_{1,r}}\left[\left(\frac{h - P_w}{P_{tx,v}L_0^{-1}}\right)^{-\frac{1}{\alpha}}\right]\right.\right.$$
$$\times \left|-(h - P_w)^{-\frac{1}{\alpha}-1}\right|f_{d_{2,r}}\left[\left(\frac{z - h}{P_{tx,v}L_0^{-1}}\right)^{-\frac{1}{\alpha}}\right]$$
$$\times \left|-(z - h)^{-\frac{1}{\alpha}-1}\right|dh. \quad (21)$$

Lemma 3: Let z be a random variable with probability density of $f_z(z)$ and $w = 1/\log_2(1 + [a_0/z])$, where a_0 is a positive constant. The probability distribution of w is given as

$$f_w(w) = f_z\left(\frac{a_0}{2^{1/w} - 1}\right)\left|\frac{-a_0 2^{1/w} \ln(1/w)}{w^2(2^{1/w} - 1)^2}\right|. \quad (22)$$

Proof: As $z = h(w) = (a_0/[2^{1/w} - 1])$, we have $f_w(w) = f_z[h(w)]|h'(w)|$. ∎

Let $L_1 = \log_2(1 + ([P_{tx,v}/L_0(d'^{\text{max}}_{v,r})^\alpha]/[P_w + \sum_{j=1}^{2} P_{tx,v}/L_0 d_{j,r}^\alpha]))$, and $L_2 = \log_2(1 + ([P_{tx,v}/L_0(d_{v,m} - (H_{e,m} - 1)d'^{\text{max}}_{v,r})^\alpha]/[P_w + \sum_{j=1}^{2} P_{tx,v}/L_0 d_{j,r}^\alpha]))$. According to Lemma 3, we get

$$f_{L_1}(l_1) = f_z\left(\frac{P_{tx,v}L_0^{-1}(d'^{\text{max}}_{v,r})^{-\alpha}}{2^{1/l_1} - 1}\right)$$
$$\times \left|\frac{-P_{tx,v}L_0^{-1}(d'^{\text{max}}_{v,r})^{-\alpha}2^{1/l_1}\ln(1/l_1)}{l_1^2(2^{1/l_1} - 1)^2}\right| \quad (23)$$

and

$$f_{L_2}(l_2) = f_z\left(\frac{P_{tx,v}L_0^{-1}(d_{v,m} - (H_{e,m} - 1)d'^{\text{max}}_{v,r})^{-\alpha}}{2^{1/l_2} - 1}\right)$$
$$\times \left|\frac{-P_{tx,v}L_0^{-1}(d_{v,m} - (H_{e,m} - 1)d'^{\text{max}}_{v,r})^{-\alpha}2^{1/l_2}\ln(1/l_2)}{l_2^2(2^{1/l_2} - 1)^2}\right|. \quad (24)$$

Based on (23) and (24), we get the following theorem.

Theorem 2: The probability distribution of $t_{v,m,i}^{\text{total}}$ can be expressed as

$$f_{t_{v,m,i}^{\text{total}}}(t) = \frac{q_v^2}{f_i^2(H_{e,m} - 1)}\int_{-\infty}^{+\infty} f_{L_1}\left(\frac{q_v l}{f_i(H_{e,m} - 1)}\right)$$
$$\times f_{L_2}\left(t - \left(\frac{q_v l}{f_i}\right)\right)dl. \quad (25)$$

Considering that a computing task needs to be accomplished under its delay constraint, we define a reliable link as its task file transmission is completed within a given time threshold. For type i task, the time threshold is given as $t_i^{\text{max}} - t_m^{\text{wait}} - g_i/W_m$. Thus, the reliability of transmission type i task is defined as follows:

$$\text{Pr}_{f,i} = \text{Pr}\left\{t_{v,m,i}^{\text{total}} \leq t_i^{\text{max}} - t_m^{\text{wait}} - g_i/W_m\right\}. \quad (26)$$

Let ϑ be a reliability threshold. We aim to keep the task file transmission reliability no less than ϑ, i.e., $\text{Pr}_{f,i} \geq \vartheta$. To enhance the reliability, we use redundant offloading schemes. For each computing task offloading through a joint V2V and V2R mode, the vehicle that generates the task sends some copies of the task file to the target MEC server in multiple separated transmission paths. The offloading is accomplished when any copy reaches the MEC server and is executed there. We consider a scenario, where a vehicle located at road segment e offloads a type-i task to Serv_m. Let $N_{i,e,m}$ denote the number of paths required for offloading the task to Serv_m. From the perspective of a single task while ignoring the interference between multiple redundant transmissions, to assure reliability, $N_{i,e,m}$ should be no less than $N_{\text{low},i,e,m}$, which is given as

$$N_{\text{low},i,e,m} = \left\lceil \log_{(1-\text{Pr}_{f,i})}(1 - \vartheta) \right\rceil. \quad (27)$$

However, in the offloading system, several tasks may be generated simultaneously. The average number of tasks generated on vehicles located in one road segment at the same time frame will be

$$\bar{N}_{\text{task}} = \sum_{k=1}^{\infty} P_g k(\rho e)^k \exp(-\rho e)/k!. \quad (28)$$

According to our proposed deep Q-learning-based task offloading scheme, the tasks of the same type and generated in the vehicles at the same road segment are offloaded in an identical approach. Thus, the average number of type-i tasks, which are generated at road segment e and offloaded to Serv_m through a joint V2V and V2R mode, is given as

$$\bar{N}'_{\text{task}} = \beta_i \bar{N}_{\text{task}} z_{i,e,m}. \quad (29)$$

If all these $\bar{N}'_{\text{task}}$ tasks choose to offload their tasks with $N_{i,e,m}$ redundant transmission paths, serious interference may occur between the paths. Affected by this interference, task transmission reliability decreases while transmission delay increases. Consequently, offloading utility decreases. Thus, there exists a maximum number of redundant file transmission for each type of tasks in a given road segment. As it is hard to derive a closed-form expression of the relations between $N_{i,e,m}$ and $t^{\text{total}}_{v,m,i}$, we take a search approach to obtain the optimal number of redundant transmission path. To improve the searching efficiency while excluding invalid value, we present the upper and lower bounds of the redundant number. We first analyze the lower bound and give the following theorem.

Theorem 3: In the presence of interference caused by multiple redundant task file transmissions, $N_{\text{low},i,e,m}$ given in (27) can act as the lower bound of $N_{i,e,m}$.

Proof: From (19), we can find that as $N_{i,e,m}$ increases, more interference may be brought to task file transmission and $t^{\text{total}}_{v,m,i}$ becomes longer. Thus, the transmission reliability shown in (26) decreases. Let $\text{Pr}'_{f,i}$ and $N'_{\text{low},i,e,m}$ denote the reliability and the minimum number of paths required for offloading transmission in the scenario with redundant transmission interference, respectively. Based on the above analysis, we have $\text{Pr}'_{f,i} < \text{Pr}_{f,i}$. Since $0 < 1 - \text{Pr}_{f,i} < 1$, according to (27), we get $N_{\text{low},i,e,m} < N'_{\text{low},i,e,m}$. Thus, $N_{\text{low},i,e,m}$ can act as the lower band of the searching range. ∎

Next, we investigate the upper bound of $N_{i,e,m}$. In a given system state, to maximize the total offloading utility shown in (3), type-i tasks at road segment e should choose the offloading mode that brings them higher utility, where $i \in [1, G]$ and $e \in [1, E]$. Motivated by this consideration, the maximum redundant number of task transmission, $N_{\text{max},i,e,m}$ can be obtained through comparing the gained utilities through various offloading modes. Considering that the computing capacity of Serv_0 at the BS is much powerful compared to the servers equipped at the RSUs, and that no interference between transmission vehicles exists in the cellular network, although the communication cost of V2B mode is higher than that of V2V mode, offloading tasks to Serv_0 in V2B mode can be taken as a backup approach. Thus, we focus on comparing the utilities gained from offloading type-i tasks to Serv_0 and Serv_m.

Considering that $N_{i,e,m}$ redundant task transmissions may be generated simultaneously, in this scenario, the density of interfering vehicles is given as $P_g\rho(1 - \beta_i + N_{i,e,m}\beta_i)$. According to Lemma 1, the probability distribution of the interference distance can be shown as

$$f(d_{i,r}) = \frac{\left(n_r/P_g\rho\left(1 - \beta_i + N_{i,e,m}\beta_i\right)\right)^i}{(i-1)!}d_{i,r}^{i-1}$$
$$\times e^{-\frac{n_r}{P_g\rho\left(1-\beta_i+N_{i,e,m}\beta_i\right)}d_{i,r}}. \tag{30}$$

Based on (30) and Theorem 2, we get the average time cost for transmitting type i task from a vehicle located in road segment e to RSU m, which can be presented as

$$\bar{t}^{\text{total}}_{v,m,i} = \int_0^{+\infty} f'_{t^{\text{total}}_{v,m,i}}(t)dt. \tag{31}$$

Algorithm 2 Reliable Offloading in Vehicular Edge Computing Networks

Initialization:
    Offloading strategies obtained from Algorithm 1.
1: Sort road segments according to $\gamma_{v,b}$ from high to low;
2: **For** each road segment with $\gamma_{v,b}$ from high to low **Do**
3:    **For** each type of tasks generated in this road segment to be offloaded to Serv_m, $m = \{1, 2, ..., M\}$ **Do**
4:        Compute type i task's transmission reliability $\text{Pr}_{f,i}$ according to (26);
5:        **if** $\text{Pr}_{f,i} < \vartheta$ **then**
6:            Compute $N_{\text{min},i,e,m}$ and $N_{\text{max},i,e,m}$ according to (27) and (33), respectively;
7:            Search $N_{\text{opt},i,e,m} = \arg\max U_r$ according to (32), where $N_{\text{min},i,e,m} \leq N_{\text{opt},i,e,m} \leq N_{\text{max},i,e,m}$;
8:            Compute reliability $\text{Pr}'_{f,i}$ with redundant transmission;
9:            **if** $\text{Pr}'_{f,i} \geq \vartheta$ **then**
10:               Offload type-i tasks generated on segment e to Serv_m in $N_{\text{opt},i,e,m}$ redundant paths;
11:            **else**
12:               Offload type-i tasks generated on segment e to Serv_0 through V2B mode;
13:               Update the states of Serv_0 and Serv_m;
14:            **end if**
15:        **end if**
16:    **End For**
17: **End For**

According to the utility definition in (3), we have the utility gained from offloading a type-i task to Serv_m with redundant transmission as

$$U_r = \varsigma_i\left(t_i^{\text{max}} - t_{i,r}^{\text{total}}\right) - N_{i,e,m}H_{e,m}q_v c_v \bar{t}^{\text{total}}_{v,m,i} - g_i c_x/W_m \tag{32}$$

where total time cost $t_{i,r}^{\text{total}} = \bar{t}^{\text{total}}_{v,m,i} + t_m^{\text{wait}} + g_i/W_m$.

From (30) and (32), we can see that more redundant transmission paths lead to more serious interference and higher offloading cost. Since redundant task file transmission is taken under the condition that U_r is higher than what is gained from offloading in V2B mode, the maximum number of redundant paths for type-i tasks generated at road segment e to Serv_m will be

$$N_{\text{max},i,e,m} = R_r\left(q_c c_c f_i/R_c - \varsigma_i\left(t_{i,r}^{\text{total}} - t_{i,c}^{\text{total}}\right)\right.$$
$$\left. - g_i c_x\left(\frac{1}{W_m} - \frac{1}{W_0}\right)\right)/H_{e,m}q_v c_v f_i. \tag{33}$$

Based on the lower and upper bounds in (27) and (33), respectively, the optimal redundant transmission number can be searched in a limited interval and be obtained efficiently. The reliable offloading scheme that prevents offloading failure through optimal redundant transmission is illustrated in Algorithm 2.
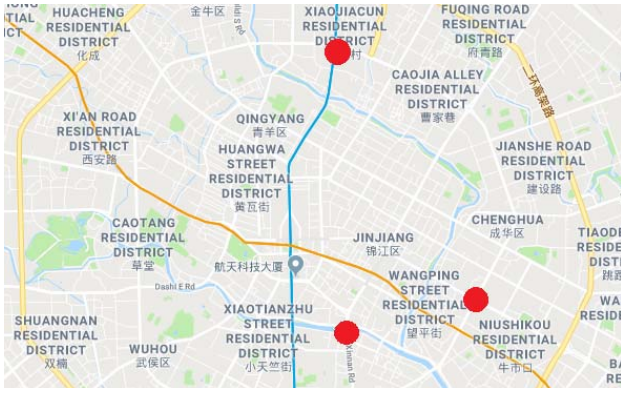
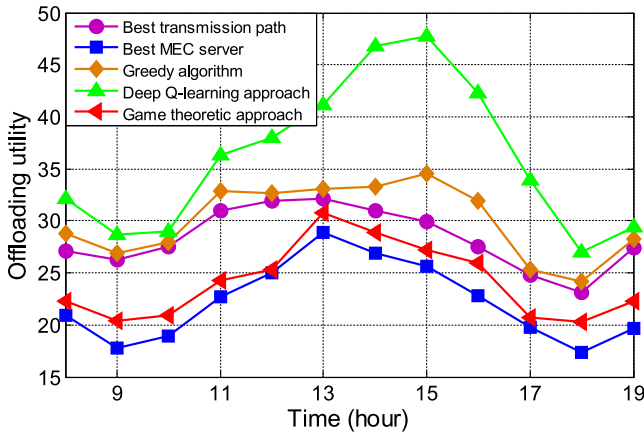Fig. 3. Snapshot of the selected roads in the urban area.

Fig. 4. Average utilities with different offloading schemes.

Fig. 5. Average utilities with different offloading schemes and various numbers of BSs.

VI. NUMERICAL RESULTS

In this section, we evaluate the performance of the proposed task offloading schemes based on real traffic data, which consists of 1.4 billion GPS traces of more than 14 000 taxis recorded for 21 days in a city of China [27]. In the recorded urban area, we select three roads, which are shown in Fig. 3. The corresponding geographical information is obtained from Google map. To increase the scale of the simulation, background traffic flow is added to the roads.

We consider a scenario that, where are one BS and five RSUs on each selected road. We set computing capacity $W_0 = 1000$ units, and the capacities of the MEC servers equipped on the RSUs are randomly taken from [100, 200] units. The amount of allocated spectrum resources q_c is 20MHz and q_v is 10MHz [28]. L_0 and α are 47.86 dB and 2.75, respectively. The transmission power $P_{tx,b}$ and $P_{tx,v}$ are set as 33 dBm and 30 dBm, respectively [29].

Fig. 4 presents the impact of real traffic on the obtained average utilities of a task with different offloading schemes. It is clear that our proposed deep Q-learning scheme yields higher offloading utility compared to other schemes, especially in the nonrush period from 12:00 to 16:00. The reason is that our scheme jointly considers transmission efficiency as well as load states of the MEC servers. However, the offloading scheme that chooses target server according to the vehicles'
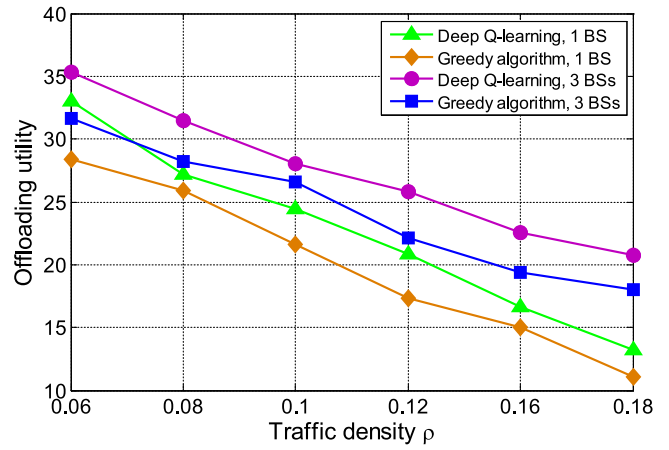
best transmission path and the scheme that selects MEC server according to server state only take one factor into account. The ignored factor may seriously affect the offloading efficiency. In the game theoretic approach, the vehicles running on a road segment act as the players that compete for task offloading services to get higher utility. Since each vehicle independently determines its offloading strategy from the perspective of its own interests, and ignores the cooperation between multiple vehicles, the system performance is getting worse. In the greedy algorithm, each vehicle chooses its offloading strategy in a distributed manner. For each time frame, a vehicle determines its optimal strategy according to currently observed communication environment and the service states of MEC servers. From Fig. 4, we see that the greedy algorithm gets less average offloading utility compared to our proposed learning approach. Task execution on an MEC server takes a certain amount of time, especially when it serves a large number of tasks. Thus, offloading strategies adopted in previous time frames may affect system performance of subsequent frames. Although the greedy algorithm jointly optimizes file transmission path and MEC server selection in current frame, it ignores the follow-up effects. In contrast, our proposed learning scheme takes both of these effects into the design of offloading strategies, thus leading to better performance.

Fig. 5 compares the performance of our proposed deep Q-learning offloading approach and greedy algorithm, which are implemented in a road with different numbers of BSs and various traffic densities. From the figure, we can clearly see that our proposed deep Q-learning approach yields better utility compared to the greedy algorithm in both scenarios: a single BS and three BSs. The reason is that our approach takes into system performance relations between subsequent frames into the design of the offloading strategies. We further find that the difference between the utilities gained by the same approach in different scenarios with various numbers of BSs becomes larger as traffic density grows. Since the MEC servers equipped on the BSs only take parts of vehicular computing tasks and the transmission cost of V2B mode is higher than that of V2V mode, a number of tasks will be preferentially offloaded to the MEC servers equipped on RSUs through V2V
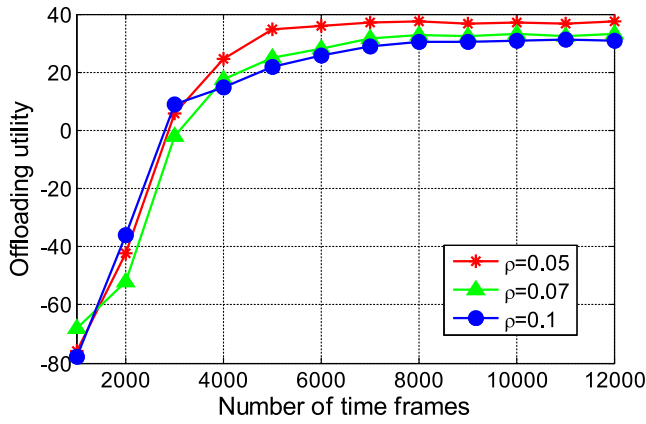
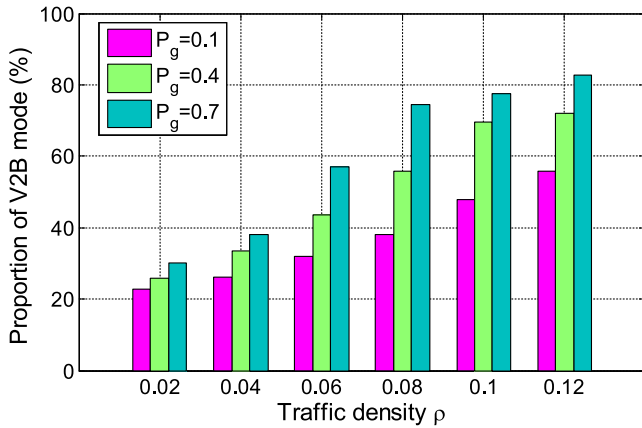Fig. 6. Convergence of the proposed deep Q-learning-based offloading scheme.

Fig. 8. Average reliability with different offloading schemes.

Fig. 7. Proportion of V2B mode offloading.

Fig. 9. Average utilities with different offloading schemes.

transmissions. However, as traffic density increases, V2V communication interference gets worse, and more vehicles choose V2B mode to offload their tasks.

Fig. 6 shows the convergence of our proposed deep Q-learning-based offloading scheme. From this figure, we see that the learning process takes about 8000 time frames to reach the optimal offloading strategies with different vehicle density ρ. It is noteworthy that in practice, the proposed learning-based offloading scheme is executed offline, which means the actual duration of execution has little to do with the performance of the task offloading application of our interest. For a given steady vehicular traffic flow and computing capacity of the MEC servers, we may obtain the optimal offloading strategies for vehicles with various states in advance. The strategy set is stored in the control center, and can be obtained and applied on vehicles directly.

In Fig. 7, we compare the impacts of the task generation probability P_g on proportion of V2B mode offloading with various traffic density. It can be seen that adopting our proposed offloading scheme, the proportion of task transmission in V2B mode becomes higher as P_g increases. Offloading a large number of generated tasks through vehicular communication may bring serious interference and impair offloading efficiency. Consequently, more parts of tasks are offloaded to
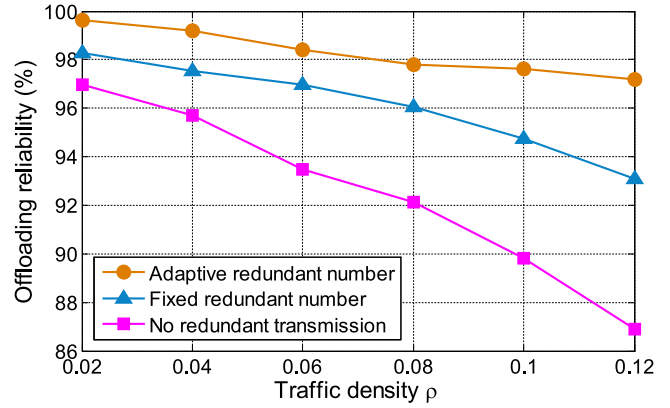
$Serv_0$ through V2B mode with the increase of P_g. Fig. 7 illustrates that comparing to the performance of the case when $P_g = 0.1$, the proportion with $P_g = 0.7$ raises faster with the increase of traffic density when ρ is small, and changes slowly when ρ is large. This result indicates that offloading in V2B mode may alleviate the transmission burden on vehicular networks when the number of generated tasks is relatively small. However, due to the resource constraints of $Serv_0$, this approach cannot continuously improve offloading utility when the number of generated tasks is high.

Fig. 8 shows the comparison of average reliability for different offloading schemes with various traffic density. Our proposed offloading scheme with adaptive redundant transmission path number brings the highest offloading reliability. Although compared to offloading scheme without redundant data transmission, applying offloading scheme with fixed number of redundant path yields higher reliability, its performance gets poor as ρ increases, especially when ρ is higher than 0.08. Due to the interference between vehicle communication pairs, in the scenario with high traffic density, too much redundant transmission may further aggravate the interference, and worsen offloading reliability.

Fig. 9 illustrates average utilities of a task for different offloading schemes with various traffic density. As higher traffic density leads to worse vehicular communication

performance, more redundant transmission and corresponding communication cost are required to ensure the offloading reliability. As a result, the utility of our proposed adaptive redundant scheme falls down with increase of ρ. However, due to the adaptive number of redundant path as well as the cooperation of multiple transmission modes, which help reduce the offloading cost, our scheme still gets higher utility. It is noteworthy that when ρ is above 0.08, the utility of the scheme with fixed number of redundant paths is lower than that of the offloading scheme without any redundant transmission. The reason is that in high traffic density scenario, fixed number offloading scheme results in significant interference in vehicular communication, which reduces data delivery rate while increasing offloading delay. Therefore, the fixed number scheme improves offloading reliability at the cost of great utility loss.

VII. CONCLUSION

In this paper, we have studied task offloading in a heterogeneous vehicular network with multiple MEC servers. To maximize task offloading utility, we first design an optimal offloading scheme with joint MEC server selection and transmission mode determination in a deep Q-learning approach. Then, we focus on reliable offloading in presence of task transmission failure, and propose an adaptive redundant offloading algorithm to ensure offloading reliability while improving system utility. The analytic results illustrate the gain from the proposed schemes for offloading vehicular tasks with optimal utility as well as respecting the constraints of reliability and latency.

APPENDIX A
PROOF OF THEOREM 1

Consider an interval $(-d_{v,r}^{\max}, d_{v,r}^{\max})$ centered at the origin. When a point process consisting of V points distributed in this interval, the Euclidean distance d_E between the origin and the farthest point follows the following conditional probability distribution [30]:

$$f_{d_E}(d|V) = V d^{V-1}/d_{v,r}^{\max}, \quad 0 \le d \le d_{v,r}^{\max}. \quad (34)$$

The traffic density of the vehicles running on the road is ρ and task generating probability is P_g. As the vehicles generating tasks cannot receive data from other ones, the density of potential V2V relay vehicles is $(1 - P_g)\rho$. Then, the probability of υ vehicles located in road segment $(-d_{v,r}^{\max}, d_{v,r}^{\max})$ is calculated as

$$\Pr\{V = \upsilon\} = \frac{\left(2(1 - P_g)\rho d_{v,r}^{\max}\right)^\upsilon e^{-2(1 - P_g)\rho d_{v,r}^{\max}}}{\upsilon!}. \quad (35)$$

According to (34) and (35), we obtain the probability distribution of d_E as

$$f_{d_E}(d) = \sum_{\upsilon=1}^{\infty} f_{d_E}(d|V = \upsilon) \Pr\{V = \upsilon\}$$
$$= 2(1 - P_g)\rho e^{-2(1 - P_g)\rho\left(d_{v,r}^{\max} - d\right)}, \quad 0 \le d \le d_{v,r}^{\max}. \quad (36)$$

Let Δd be distance reduction from $d_{v,r}^{\max}$, and v_f be the farthest vehicle that can be communicated with by the vehicle placed at the origin. The probability that vehicle v_f is on road segment $(d_{v,r}^{\max} - \Delta d, d_{v,r}^{\max})$ can be given as

$$\Pr\{d \in \left(d_{v,r}^{\max} - \Delta d, d_{v,r}^{\max}\right)\}$$
$$= \Pr\{d < d_{v,r}^{\max}\} - \Pr\{d < d_{v,r}^{\max} - \Delta d\}$$
$$= 1 - e^{-2(1 - P_g)\rho\Delta d}. \quad (37)$$

Given coefficient σ, where $0 < \sigma \ll 1$, when $\Delta d > \ln \sigma/2(1 - P_g)\rho$, vehicle v_f is located in the road segment with a probability close to 1. Thus, the mean value of the distance between v_f and the vehicle located at the origin is $d_{v,r}^{\max} - \Delta d/2$.

APPENDIX B
PROOF OF LEMMA 1

For vehicular communication from vehicle v to receiver r, as each transmitting vehicle randomly chooses a spectrum block from n_r blocks, the average number of vehicles between r and its nearest interfering vehicle, which works in the same spectrum as r, is presented as

$$\bar{h} = \left(1 - \frac{1}{n_r}\right)\frac{1}{n_r} + \left(1 - \frac{1}{n_r}\right)^2 \frac{1}{n_r} \cdot 2$$
$$+ \cdots + \left(1 - \frac{1}{n_r}\right)^{h-1} \frac{1}{n_r} \cdot (h - 1) + \cdots$$
$$= n_r - 1. \quad (38)$$

Considering that the location of vehicles on the road follows Poisson distribution with density ρ, the probability distribution of the distance between the nearest interfering vehicle and r is given as

$$f(d_{1,r}) = \begin{cases} \frac{n_r}{P_g \rho} e^{-\frac{n_r}{P_g \rho} d_{1,r}}, & d_{1,r} \ge 0 \\ 0, & d_{1,r} < 0. \end{cases} \quad (39)$$

Let $\{d_{1,r}, d_{2,r}, \ldots\}$ denote the distance between r and the interfering vehicles from near to far, respectively. Since the location of these vehicles is independent and identically distributed, and the distance between two adjacent vehicles follows exponential distribution, the probability of $d_{j,r}$, where $j = \{1, 2, \ldots\}$, follows Erlang distribution and can be shown as:

$$f(d_{j,r}) = \begin{cases} \frac{n_r/(P_g \rho)^j}{(j-1)!} d_{j,r}^{j-1} e^{-\frac{n_r}{P_g \rho} d_{j,r}}, & d_{j,r} \ge 0 \\ 0, & d_{j,r} < 0. \end{cases} \quad (40)$$

APPENDIX C
PROOF OF LEMMA 2

Let $k = b_0 x^{-\alpha} + c_0$ and $s = b_0 y^{-\alpha}$. We have

$$f_k(k) = f_x\left[\left(\frac{k - c_0}{b_0}\right)^{-\frac{1}{\alpha}}\right]\left|\frac{-(k - c_0)^{-\frac{1}{\alpha}-1}}{\alpha b_0^{-\frac{2}{\alpha}-1}}\right|$$

and

$$f_s(s) = f_y\left[\left(\frac{s}{b_0}\right)^{-\frac{1}{\alpha}}\right]\left|\frac{-(s)^{-\frac{1}{\alpha}-1}}{\alpha b_0^{-\frac{2}{\alpha}-1}}\right|.$$

As k and s are independent random variables, there is

$$f_z(z) = \int_{-\infty}^{+\infty} f_k(h)f_s(z-h)dh. \tag{41}$$

Integrating $f_k(k)$ and $f_s(s)$ into (41), we can derive (20).

References

[1] M. Wollschlaeger, T. Sauter, and J. Jasperneite, "The future of industrial communication: Automation networks in the era of the Internet of Things and industry 4.0," *IEEE Ind. Electron. Mag.*, vol. 11, no. 1, pp. 17–27, Mar. 2017.

[2] W. Zhong, K. Xie, Y. Liu, C. Yang, and S. Xie, "Auction mechanisms for energy trading in multi-energy systems," *IEEE Trans. Ind. Informat.*, vol. 14, no. 4, pp. 1511–1521, Apr. 2018.

[3] Z. Zhou *et al.*, "Computation resource allocation and task assignment optimization in vehicular fog computing: A contract-matching approach," *IEEE Trans. Veh. Technol.*, to be published.

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

[5] K. Zhang, S. Leng, Y. He, S. Maharjan, and Y. Zhang, "Mobile edge computing and networking for green and low-latency Internet of Things," *IEEE Commun. Mag.*, vol. 56, no. 5, pp. 39–45, May 2018.

[6] Y. Dai, D. Xu, S. Maharjan, and Y. Zhang, "Joint computation offloading and user association in multi-task mobile edge computing," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 12313–12325, Dec. 2018.

[7] M. Shojafar, N. Cordeschi, and E. Baccarelli, "Energy-efficient adaptive resource management for real-time vehicular cloud services," *IEEE Trans. Cloud Comput.*, vol. 7, no. 1, pp. 196–209, Mar. 2019.

[8] J. Liu *et al.*, "High-efficiency urban traffic management in context-aware computing and 5G communication," *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 34–40, Jan. 2017.

[9] M. Sookhak *et al.*, "Fog vehicular computing: Augmentation of fog computing using vehicular cloud computing," *IEEE Veh. Technol. Mag.*, vol. 12, no. 3, pp. 55–64, Sep. 2017.

[10] X. Hou *et al.*, "Vehicular fog computing: A viewpoint of vehicles as the infrastructures," *IEEE Trans. Veh. Technol.*, vol. 65, no. 6, pp. 3860–3873, Jun. 2016.

[11] G. Qiao, S. Leng, K. Zhang, and Y. He, "Collaborative task offloading in vehicular edge multi-access networks," *IEEE Commun. Mag.*, vol. 56, no. 8, pp. 48–54, Aug. 2018.

[12] A. Aissioui, A. Ksentini, A. M. Gueroui, and T. Taleb, "On enabling 5G automotive systems using follow me edge-cloud concept," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 5302–5316, Jun. 2018.

[13] P. Luoto *et al.*, "Vehicle clustering for improving enhanced LTE-V2X network performance," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2017, pp. 1–5.

[14] C. Perfecto, J. D. Ser, and M. Bennis, "Millimeter-wave V2V communications: Distributed association and beam alignment," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 2148–2162, Sep. 2017.

[15] S. A. A. Shah, E. Ahmed, M. Imran, and S. Zeadally, "5G for vehicular communications," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 111–117, Jan. 2018.

[16] J. Xu, L. Chen, and S. Ren, "Online learning for offloading and autoscaling in energy harvesting mobile edge computing," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 3, pp. 361–373, Sep. 2017.

[17] A. Crutcher *et al.*, "Hyperprofile-based computation offloading for mobile edge networks," in *Proc. IEEE 14th Int. Conf. Mobile Ad Hoc Sensor Syst.*, Oct. 2017, pp. 525–529.

[18] Y. He, N. Zhao, and H. Yin, "Integrated networking, caching, and computing for connected vehicles: A deep reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 67, no. 1, pp. 44–55, Jan. 2018.

[19] Q. Zhang, M. Lin, L. T. Yang, Z. Chen, and P. Li, "Energy-efficient scheduling for real-time systems based on deep Q-learning model," *IEEE Trans. Sustain. Comput.*, vol. 4, no. 1, pp. 132–141, Jan./Mar. 2019.

[20] M. Gadaleta, F. Chiariotti, M. Rossi, and A. Zanella, "D-DASH: A deep Q-learning framework for DASH video streaming," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 703–718, Dec. 2017.

[21] K. Kashihara, "Deep Q learning for traffic simulation in autonomous driving at a highway junction," in *Proc. IEEE Int. Conf. Syst. Man Cybern. (SMC)*, Oct. 2017, pp. 984–988.

[22] K. Zhang, Y. Mao, S. Leng, Y. He, and Y. Zhang, "Mobile-edge computing for vehicular networks: A promising network paradigm with predictive off-loading," *IEEE Veh. Technol. Mag.*, vol. 12, no. 2, pp. 36–44, Jun. 2017.

[23] F. Malandrino, C. Casetti, C.-F. Chiasserini, and M. Fiore, "Content download in vehicular networks in presence of noisy mobility prediction," *IEEE Trans. Mobile Comput.*, vol. 13, no. 5, pp. 1007–1021, May 2014.

[24] S. Xie, W. Zhong, K. Xie, R. Yu, and Y. Zhang, "Fair energy scheduling for vehicle-to-grid networks using adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 8, pp. 1697–1707, Aug. 2016.

[25] V. Mnih *et al.*, "Playing Atari with deep reinforcement learning," *arXiv: 1312.5602*, Dec. 2013. [Online]. Available: https://arxiv.org/abs/1312.5602

[26] H. van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in *Proc. Nat. Conf. Artif. Intell.*, 2016, pp. 2094–2100.

[27] *Taxi GPS Data*. Accessed: Aug. 2014. [Online]. Available: http://www.pkbigdata.com/common/zhzgbCmptData Details.html

[28] K. I. Pedersen, G. Berardinelli, F. Frederiksen, P. Mogensen, and A. Szufarska, "A flexible 5G frame structure design for frequency-division duplex cases," *IEEE Commun. Mag.*, vol. 54, no. 3, pp. 53–59, Mar. 2016.

[29] A. Bazzi, B. M. Masini, A. Zanella, and I. Thibault, "On the performance of IEEE 802.11p and LTE-V2V for the cooperative awareness of connected vehicles," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10419–10432, Sep. 2017.

[30] S. Srinivasa and M. Haenggi, "Distance distributions in finite uniformly random networks: Theory and applications," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 940–949, Feb. 2010.

Ke Zhang received the Ph.D. degree from the University of Electronic Science and Technology of China, Chengdu, 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 current research interests include scheduling of mobile edge computing, design and optimization of next-generation wireless networks, smart grid, and the Internet of Things.
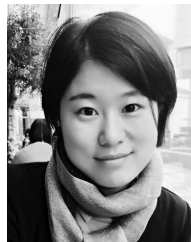
Yongxu Zhu (GS'16–M'19) received the M.S. degree from the Beijing University of Posts and Telecommunications, Beijing, China, and Dublin City Univeristy, Dublin, Ireland, in 2012 and 2013, respectively, and the Ph.D. degree in electrical engineering from University College London, London, U.K., in 2017.

She is currently a Research Associate with Loughborough University, Loughborough, U.K. Her current research interests include UAV communications, wireless edge caching, millimeter-wave communications, heterogeneous cellular networks, and physical-layer security.

Supeng Leng (M'06) received the Ph.D. degree from Nanyang Technological University (NTU), Singapore.

He is a Full Professor and a Vice Dean with the School of Information and Communication Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu, China, where he is also the Leader of the Research Group with Ubiquitous Wireless Networks. He has been a Research Fellow with Network Technology Research Center, NTU. He has authored or co-authored over 150 research papers in recent years. His current research interests include resource, spectrum, energy, routing, and networking in Internet of Things, vehicular networks, broadband wireless access networks, smart grid, and the next generation mobile networks.

Dr. Leng 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.

Yejun He (SM'09) received the Ph.D. degree in information and communication engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2005.

He is currently a Full Professor with the College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China, where he is also the Director of Guangdong Engineering Research Center of Base Station Antennas and Propagation and the Shenzhen Key Laboratory of Antennas and Propagation. His current research interests include wireless mobile communication and antennas and radio frequency.

Dr. He is currently the IEEE Antennas and Propagation Society Shenzhen Chapter Chair. He is a Fellow of the IET.

Sabita Maharjan (M'09) received the Ph.D. degree in networks and distributed systems from University of Oslo, and Simula Research Laboratory, Norway, in 2013.

She is currently a Senior Research Scientist in the Simula Metropolitan Center for Digital Engineering, Oslo, Norway, and an Associate Professor in University of Oslo, Norway. She has been published regularly in prestigious journals in her field such as in the IEEE TRANSACTIONS ON SMART GRID, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, *IEEE Communications Magazine*, *IEEE Network*, *IEEE Wireless Communications Magazine*, and the IEEE INTERNET OF THINGS JOURNAL. Her current research interests include vehicular networks and 5G, network security and resilience, smart grid communications, Internet of Things, massive machine type communication, software defined wireless networking, and advanced vehicle safety.

Dr. Maharjan serves/has served on the Technical Program Committee of conferences including top conferences like IEEE INFOCOM and IEEE IWQoS.

Yan Zhang (M'05–SM'10) received the Ph.D. degree from the School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore.

He is currently a Full Professor with the Department of Informatics, University of Oslo, Oslo, Norway. His current research interests include next generation wireless networks leading to 5G, and green and secure cyber-physical systems (e.g., smart grid, healthcare, and transport).

Dr. Zhang is an Associate Technical Editor for *IEEE Communications Magazine*, an Editor of *IEEE Network*, the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, the IEEE INTERNET OF THINGS JOURNAL, *IEEE Vehicular Technology Magazine*, and an Associate Editor for IEEE ACCESS. He serves as the Chair of a number of conferences, including IEEE CloudCom 2015, IEEE SmartGridComm 2015, IEEE PIMRC 2016, IEEE CloudCom 2016, IEEE ICCC 2016, IEEE CCNC 2016, IEEE GLOBECOM 2017, and IEEE VTC-Spring 2017. He is a Senior Member of the IEEE ComSoc, IEEE CS, IEEE PES, and IEEE VT societies. He serves as a TPC member for numerous international conferences, including IEEE INFOCOM, IEEE ICC, IEEE GLOBECOM, and IEEE WCNC. He is the IEEE Vehicular Technology Society Distinguished Lecturer.