IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Artificial Intelligence-Empowered Edge of Vehicles: Architecture, Enabling Technologies, and Applications

HONGJING JI[1,2], OSAMA ALFARRAJ[3], AND AMR TOLBA.[3,4]

[1]School of Software, Dalian University of Technology, 116620, Dalian, China
[2]School of Software, Taiyuan University of Technology, 030024, Taiyuan, China
[3]Computer Science Department, Community College, King Saud University, Riyadh 11437, Saudi Arabia
[4]Mathematics and Computer Science Department, Faculty of Science, Menoufia University, Shebin-El-kom 32511, Egypt

Corresponding author: Osama Alfarraj (e-mail: oalfarraj@ksu.edu.sa).

This work was funded by the Researchers Supporting Project No. (RSP-2019/102) King Saud University, Riyadh, Saudi Arabia.

ABSTRACT With the proliferation of mobile devices and a wealth of rich application services, the Internet of vehicles (IoV) has struggled to handle computationally intensive and delay-sensitive computing tasks. To substantially reduce the latency and the energy consumption, application work is offloaded from a mobile device to a remote cloud or a nearby mobile edge cloud for processing. Compared with remote clouds, mobile edge clouds are located at the edge of the network. Therefore, mobile edge computing (MEC) has the advantages of effectively utilizing idle computing and storage resources at the edge of the network and reducing the network transmission delay. In addition, mobile devices are increasingly moving toward intelligence. To satisfy the service experience and service quality requirements of mobile users, the vehicle Internet is transforming into the intelligent vehicle Internet. Artificial intelligence (AI) technology can adapt to rapidly changing dynamic environments to provide multiple task requirements for resource allocation, computational task scheduling, and vehicle trajectory prediction. On this basis, combined with MEC technology and AI technology, computing and storage resources are placed on the edge of the network to provide real-time data processing while providing more efficient and intelligent services. This article introduces IoV from three aspects, namely, MEC, AI and the advantages of combining the two, and analyzes the corresponding architecture and implementation technology. The application of MEC and AI in IoV is analyzed and compared with current approaches. Finally, several promising future directions in the field of IoV are discussed.

INDEX TERMS
Internet of Vehicles (IoV), Mobile Edge Computing (MEC), Artificial Intelligence (AI)

I. INTRODUCTION

WITH the rapid development of Internet of things (IoT) technology and the increasing number of vehicle networks, the traditional vehicle ad hoc network (VANET) is gradually being integrated into the Internet of vehicles (IoV). IoV is a new model that combines VANETs and vehicle remote information processing to connect vehicles, people and things [1]. In addition, it is a highly important field in intelligent transportation systems (ITSs), as it covers intelligent transportation, cloud computing, vehicle information services, logistics transportation services [2] [3], modern wireless technology, Internet access and communication and other technologies and applications [4]. According to the forecast report from Cisco, the global monthly mobile data usage in 2021 will be approximately 49 exabytes, and the number of mobile devices will be 11.6 billion, increasing about approximately seven times between 2016 and 2021 [5]. With the explosion of mobile data, mobile phones are increasingly being used for various computation-intensive applications, such as augmented reality; natural language processing; face, hand gestures, and object recognition; and various forms of user configurations used for recommendation [6]; hence, mobile users enjoy a rich experience in the service network. Faced with the surge of mobile data flow, reducing the delay of data transmission between vehicles and improving the throughput of data transmission between vehi-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2983609, IEEE Access

cles are urgent problems [7]. Therefore, the vehicle network must adopt advanced communication technology and data acquisition technology to improve the safety and efficiency of the traffic system, reduce accidents and reduce traffic congestion [8]. Generally speaking, public communication interfaces are divided into wireless networks (such as bluetooth and wi-fi) and cellular networks (such as 3G, 4G and 5G) [9]. However, the limited network bandwidth in traditional cellular networks limits the fast growth of the data transmission rate. In the emerging 5G network, the application of D2D (device to device) communication technology promises to substantially improve the spectrum efficiency to support data transmission between caching vehicles and mobile users [5]. The federal communications commission (FCC) authorized the 75-mhz band for the provision of vehicle-to-vehicle wireless communications as dedicated short-range communications (DSRC). In addition, IEEE standardizes the entire communication stack according to IEEE 802.11p as a wireless access to the vehicle environment (WAVE) to support the interconnection between vehicles and between vehicles and roads [10]. In addition, various communication modes coexist in IoV, which include vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), vehicle-to-sensor (V2S), vehicle-to-pedestrian (V2P), and vehicle-to-network (V2N) [11] communications, which form a dynamic mobile communication system. Fig. 1 illustrates the architecture of IoV. This enables the sharing and collection of information about vehicles, roads and their surroundings. While the development of communication technology can alleviate a certain amount of traffic congestion, the limited ability of the infrastructure to communicate, compute, and store resources can lead to long delays and massive data transmission problems. In order to overcome this challenge, combined with the deployment of resources on the edge of the wireless network, the proposed vehicle edge network has attracted wide attention.

Mobile edge computing (MEC) technology can overcome the challenges of traditional mobile cloud computing (MCC). For example, (1) centralized cloud servers are located far away from the terminal devices, thereby resulting in low efficiency in computation-intensive environments; (2) the offloading of computing to the cloud consumes energy, thereby reducing the service life of mobile batteries; and (3) providing mobile users with complex memory-utilization applications and higher data storage capacity is difficult [12]. Reference [13] studied the multi-user computing offloading problem of mobile edge cloud computing under multi-channel wireless interference, and put forward a distributed computing offloading algorithm by using the game theory method. In addition, MEC can provide substantial value to mobile operators, service providers and end users. The application scenarios of MEC span multiple fields, such as augmented reality, online games, big data analysis and health monitoring in the medical Internet of things [14] [15].

With the emergence of IoV and vehicle intelligence, vehicles are transforming from transport tools to intelligent terminals [16]. In addition, the variety and quantity of on-

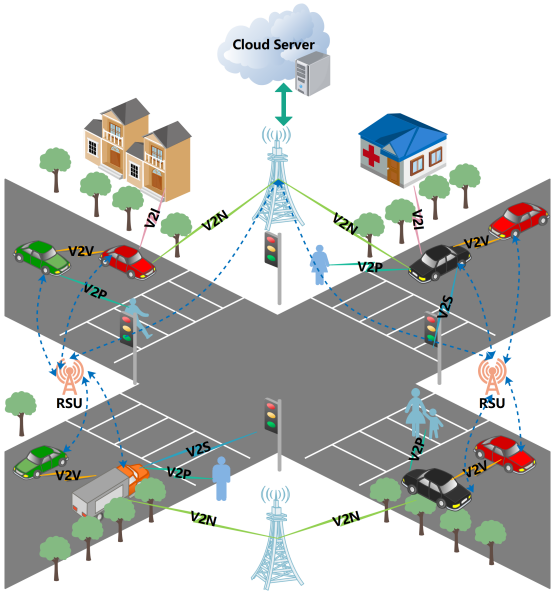


FIGURE 1: The architecture of IoV.

board equipment are increasing, and people's requirements for automobile service quality are constantly increasing [17]. In the age of IoV, vehicle-mounted intelligent modules can provide intelligent vehicle control, traffic management, accident prevention and navigation capabilities, along with rich multimedia and mobile Internet application services and many emerging interactive applications [12] that improve the user experience, reduce operating costs and promote a safe driving environment. Artificial intelligence (AI) can substantially improve the cognitive performance and intelligence of vehicle networks, thereby contributing to the optimal allocation of resources for problems with diverse, time-varying and complex characteristics [18]. Reinforcement learning (RL) is an important branch of machine learning. It refers to the process of realizing objectives via multiple steps and suitable decisions in a series of scenarios, which can be regarded as a multi-step sequential decision problem [8]. To overcome the problem of decentralized management of connected vehicles in a distributed intelligent transportation system, reference [19] designed an ant colony optimization algorithm that is based on swarm intelligence (SI) for dynamic decision-making of networked vehicles, which enables vehicles to automatically and adaptively identify the best path to the destination. In [20], an intelligent resource management strategy for joint communication mode selection, resource block allocation and power control in D2D-V2V communication vehicle networks is proposed. The model-free participant critical learning framework is used to effectively improve the learning efficiency and identify the optimal strategy to ensure that the vehicle-to-vehicle link satisfies the communication requirements of ultra-reliability and low latency while maximizing the overall network capacity.

The main contributions of this article are as follows:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2983609, IEEE Access

- We introduce the architecture of IoV, the deployment of V2X in vehicle-mounted communication, and the application of MEC and AI in IoV. We describe the advantages and development history of MEC and the relationship between AI and DRL, and We analyze the previous research on the application of AI to vehicle edge networks.
- We study the architecture of MEC-based IoV and discuss the use of MEC in IoV. In addition, the characteristics of MEC, FC and MCC are analyzed, and the key technologies for supporting MEC are introduced. In addition, the previous studies on efficient MEC calculation for IoV are analyzed.
- We consider the theoretical characteristics of AI; analyze DRL, which is a key method for realizing AI, and demonstrate the architecture of AI in IoV. In addition, We introduce the effective key AI algorithms for calculating the offload and resource allocation in IoV, and We analyze the previous AI research on IoV.
- We combine the application of AI and MEC technology in IoV and analyze the key technologies that support the application of AI in vehicle edge networks. In addition, the previous studies on edge caches and on joint computing resources and caches are introduced. Finally, the future development directions and research challenges of IoV are discussed.

This article reviews the architecture, implementation technology and application of IoV that is based on AI and MEC. We explore the characteristics of IoV development, the communication mode, and the impact of combining AI and MEC technology on the construction of intelligent IoV. This article is divided into the following parts: Section II describes the architecture of MEC in IoV and introduces the development history of edge computing and the characteristics of MEC. In addition, the research on computing offloading of MEC in IoV is analyzed. Section III mainly studies the application of AI In IoV, expounds on the architecture in which AI is combined with IoV, and discusses DRL, which is an important technique for realizing AI. The key algorithms and applications of AI in IoV are analyzed. In Section IV, the significance of the combination of AI and MEC technology in IoV is discussed, the key technologies in AI-based vehicle edge networks are studied, and relevant studies on IoV are analyzed. In Section V, the challenges that are faced by IoV and the future development directions are discussed. Finally, Section VI summarizes the study.

II. MOBILE EDGE COMPUTING IN INTERNET OF VEHICLES

With the proliferation of mobile devices in the IoV, there are stringent computing and processing requirements for computation-intensive applications and delay-sensitive applications. The combination of IoV and MEC has emerged as a promising approach for addressing the growing demand for computing by shifting heavy computing tasks to cloud resources on the edges of mobile networks. In this part, we describe the development history of MEC and the MEC architecture in IoV, and we explain the advantages of MEC in IoV. Then, the key MEC technology of IoV is introduced. Finally, the research status of MEC-based computational offloading in IoV is discussed.

A. ARCHITECTURE

With the continuous improvement of the number and intelligence of mobile devices, increasingly many mobile applications require many computing tasks. However, due to the limited computing power and battery capacity of the user's device, it is difficult to handle computationally intensive tasks locally. The emergence of cloud computing as a potential solution formally initiated the third Internet revolution. Based on the concept and advantages of cloud computing, mobile cloud computing (MCC) was proposed in 2009 and refers to a centralized cloud computing platform that migrates data processing, storage and other tasks of intensive applications from the original mobile device terminals to the cloud.

For applications that are closely involved in data-intensive and delay-sensitive computing tasks, MCC has difficultly satisfying the stringent requirements of real-time operations. Therefore, a new computing paradigm, namely, fog computing (FC), is extended from cloud computing. A fog can be described as a cloud that is closer to the ground, which pushes computing resources and application services to the edge of data generation and processing. In reference [21], the author considers the mobility of fog nodes. The task assignment process between fixed and mobile fog nodes is regarded as a two-objective optimization problem in which the service latency and quality loss must be balanced. An event-triggered dynamic task assignment framework that is based on linear-programming-based optimization (LBO) and binary particle swarm optimization (BPSO) is proposed for solving joint optimization problems. In reference [22], a real-time traffic management unloading scheme in IoV systems that is based on fog computing is proposed, which can minimize the average response time of vehicle reporting events. Although fog computing has the advantages of location-awareness and low latency, ubiquitous connectivity and ultra-low latency requirements pose challenges to real-time traffic management in smart cities [23].

To extend the cloud computing capacity to the edge of the network, to enable the end users to use cloud computing services more quickly and efficiently, and to improve the user experience, in 2013, mobile edge computing (MEC) was proposed for the provision of IT and cloud computing capabilities for wireless access networks by deploying common servers on the wireless access side. MEC is not a replacement for MCC but an extension of cloud computing that relaxes the transmission bandwidth and delay requirements. Compared with MCC, MEC has the following characteristics [24]: (1) low delay and low energy consumption: data generation and processing are conducted close to the data sources and users, thereby reducing the data transmission delay and energy consumption; (2) diversity: edge devices with various com-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2983609, IEEE Access

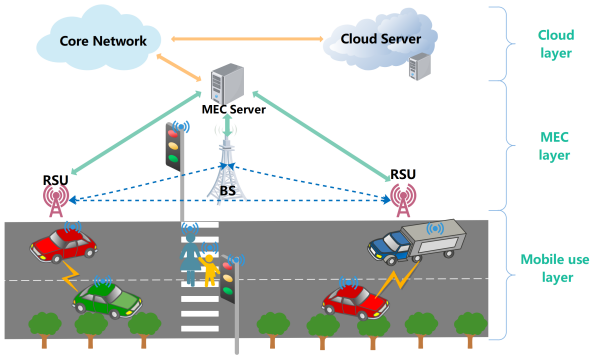


FIGURE 2: The architecture of MEC for IoV.

puting capabilities, such as roadside units (RSUs), vehicles and WiFi hotspots, coexist; and (3) resource limitation: the computing power of edge nodes is typically lower than that of cloud servers.

With the progress of MEC standardization, the focus has gradually shifted from targeting 3GPP mobile networks to supporting non-3GPP networks (Wi-Fi and wired networks) and even 5G networks. The name is also modified from moving edge calculation to multi-access edge calculation. The multi-access edge computing technology can realize the interconnection of multiple wireless access technologies to enable the computing/storage tasks of multiple servers to be conducted cooperatively. [25] The emergence of MEC servers enables wireless access networks to flexibly use computing and storage resources while providing time-sensitive, computation-intensive and highly reliable application services. In reference [26], a scenario in which co-driving vehicles and free-driving vehicles that are facilitated by HD map interconnection via a wireless network co-existence is designed, and a multi-access edge computing architecture that is based on SDN and NFV technology is proposed. The joint optimization of computing/storage resource management between MEC servers and bandwidth resource slices between Base Satations (BSs) effectively improves the utilization of computing/storage resources. MEC, FC and MCC are compared in TABLE I.

Fig. 2 shows that the architecture of vehicle edge computing can be divided into a mobile user layer, an MEC layer and a cloud layer. Communication between mobile users generates a large amount of mobile data. By offloading to RSUs, BSs, and other relay nodes with computing and storage capabilities, tasks that require more computational processing can be offloaded to MEC servers or even to remote cloud servers to fully utilize the computing resources to provide efficient computing services.

B. TECHNOLOGIES

1) Network Features Virtualization(NFV) and Software-Defined Networks(SDN)

NFV enables the abstraction of physical network resources and the flexible sharing of resources between isolated users [27]. Virtualization technology is a key technology of MEC and realizes the separation of the service layer and the physical resource layer of edge computing, and can assign tasks to various physical resources, thereby efficiently utilizing resources. By integrating NFV into the MEC server, virtualized computing and storage resources can support the functions of various applications and services and can be applied to the server for functional programming to support a variety of application services, thereby enhancing the flexibility of the server and reducing the cost function supply [26]. SDN is a new network mode that was proposed by the CLean State research group of Stanford University in the United States. It is an implementation method of NFV. By separating the control surface of network devices from the data surface and opening the programmability, the logic centralized control of distributed network nodes and mobile devices can be realized. Reference [28] studies SDN in super-dense network task offloading problems and designs the edge of a cloud or offloads tasks on a local process scheme; the main calculation and control function is separated from the distributed small unit base station, which is integrated into the centralized SD UDN in the macrocell base station controller. Based on the decision of the SD UDN controller, it is decided whether the mobile device should perform tasks locally or offload tasks to the edge cloud for processing, and the computing resources should be optimally allocated to each task to realize the objectives of minimizing the delay and preserving the user's device battery life.

2) Collaborative Mobile Edge Cloud Computing

Collaborative mobile edge cloud computing combines the advantages of MEC and MCC, which is of substantial significance for ensuring the full utilization of MEC and cloud computing resources. While cloud computing may produce long delays during offloading, it can provide sufficient cloud computing resources. Although an MEC server outperforms cloud computing in reducing the communication delay and the energy consumption, with the increasing use of computation-intensive applications, the limited computing resources of the MEC server cannot fully satisfy all the uninstallation requirements. With the increasing number of computing tasks, the resource bottleneck problem of the MEC server becomes increasingly prominent. Therefore, cloud computing and MEC should be highly complementary. Reference [29] proposed a collaborative offloading scheme for vehicle-to-vehicle networks that is based on mobile edge cloud computing and cloud computing, and it developed a distributed computing offloading and resource allocation algorithm for computational offloading optimization in vehicle-to-vehicle networks. Reference [30] proposes a design framework for edge computing in wireless broadband access networks that supports smart cities by embedding a green, viable virtual network. A suitable resource partitioning approach is used for each virtual network embedding, and backing up edge devices by using heuristic policies to determine the number and geographic location is recommended.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2983609, IEEE Access

TABLE 1: Comparison of MEC,FC and MCC

	MEC	FC	MCC
Deployment Location	Network edge	Near edge	Remote network center
Ownership	Mobile operators	Enterprise private	Cloud provider
Location awareness	High	Medium	Low
Distance from the user	Near		Far
Transmission Lelay	Low		High
Server Hardware	Small data center, medium computing resources		Large data centers with a large number of high-performance computing servers
Network Architecture	Multi-level, Distributed		Centralized
Scalability	High		General
Application scenario	Delay-sensitive applications	Widely distributed mobile applications that require low latency and a small amount of calculation	Applications with moderate delay requirements but heavy calculations

In reference [31], Ning et al. designed an iterative heuristic MEC resource allocation algorithm for making unloading decisions dynamically. In reference [32], Hu et al. considered the collaborative calculation of the offloading, combined power and time distribution. They proposed a capture-unload protocol that is based on the block-time-division mechanism for minimizing the transmission power of wireless access points.

3) Content Distribution

In the context of mass content delivery, a suitable content distribution scheme can facilitate the avoidance of repeated content transmission by the network. In addition, the application of the content distribution framework in heterogeneous IoV systems can improve the message accuracy and reduce the communication overhead between vehicles and the infrastructure [37]. Current mobile users have consumed a substantial amount of the capacity, and the demand for in-vehicle infotainment services is still growing rapidly. To improve the network performance and the user quality of service (QoS), content distribution is often combined with content caching technology and data prefetching technology to further reduce the data access latency. In reference [38], a content propagation box that is based on edge calculation is proposed. First, a two-stage relay selection algorithm is designed to facilitate edge computing devices in the selective transmission of content via V2I communication. Then, the vehicle that is selected by the edge computing device relays the content via V2V communication to the vehicle that is interested in the content during the trip to the destination. Reference [39] proposed a content distribution framework that utilizes 5G edge network caching and wireless link time slot scheduling. The wireless resource allocation and return link utilization of vehicle-to-roadside-unit communication at each information station are considered. To maximize the throughput, wireless links are dynamically allocated to vehicles using time slots. In reference [40], we studied the impacts of the storage cost and the retention time of content storage on cache optimization in mobile scenarios. In addition, a cache problem in a vehicle network is modeled, and its complexity is analyzed. For symmetric cases, an optimal dynamic programming algorithm with polynomial time complexity is developed. For general cases, a multi-helper caching algorithm with low complexity and effective retention perception is proposed, which can obtain the best caching solution.

C. APPLICATIONS

For satisfying the strict requirements of limited mobile terminal resources and computation-intensive and delay-sensitive applications, computational offloading technology is regarded as a key technology. Within the framework of MEC, the mobile terminal can offload a task to the nearby edge computing server for processing and feed back the calculation results to the mobile terminal, thereby effectively overcoming the resource limitation and reducing the power consumption of the terminal during local calculation. Offloading decision, computational resource allocation and mobility management of computational offloading are three key issues in the field of MEC-based computational offloading. In the following, we analyze the previous research on MEC in IoV.

1) Offloading Decision

In offloading decision-making, data transfer between dependent tasks is typically considered. Mobile terminal computing offloading methods mainly include local computing, offloading to the MEC server for execution and offloading to the cloud server for execution. Many studies have been conducted on offloading decision-making, such as studies on whether to offload, the quantity and location of offloading, service type, user perfection, access technology, network traffic, device performance, and edge node property [41]. The offloading method is mainly based on the resource size, the calculation and return time and the power consumption of the calculation. The main influencing factors are the delay and the energy consumption. To minimize the cost in terms of communication and computing resources, the author of reference [25] proposed a task diversion mechanism in the edge computing network of vehicles under the condition of high mobility of the vehicles. The task offloading scheme is analyzed in the scenario of an independent mobile edge computing device server and in the scenario of a collabo-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2983609, IEEE Access

TABLE 2: Comparison of Computing Offloading in MEC

Ref.	The key technology			User Numbers	Offloading Methods	Number of Compute Nodes		Compute nodes	Scheme
	Offloading Decision	Resource Allocation	Mobility Management			Single node	Multi -node		
[6]	√	√		multi-user	Partial		√	Cloud server APs BSs	A distributed algorithm for computing polynomial complexity of equal distribution of wireless and cloud resources
[12]		√		multi-user	Partial		√	MEC server, Cloud server, RUs	A greedy heuristic algorithm for an on-board cloud edge system
[13]	√			multi-user	Partial		√	BSs	A game theory method and a distributed computing offloading algorithm
[21]	√	√	√	multi-user	Partial		√	Fog node	A dynamic task assignment framework based on programming optimization and binary particle swarm optimization
[24]		√	√	multi-user	Partial		√	Macro cell RSUs	A joint power control and channel allocation scheme
[25]	√	√		multi-user	Partial		√	MEC server	Scheme 1: in the independent MEC server scenario, a task offloading scheme based on mobility. Scheme 2: in the collaborative MEC server scenario, a location-based task unloading scheme .
[28]	√	√		multi-user	Partial		√	BSs with edge cloud servers	A task unloading framework for computing moving edges in software -defined ultra-intensive networks
[29]	√	√		multi-user	Partial		√	MEC server, cloud server	A distributed collaborative computing algorithm for offloading and resource allocation
[33]	√	√		multi-user	Partial		√	MEC server	A multi-user multi-task offload scheduling scheme in a mobile edge cloud system that can be updated
[34]	√	√		multi-user	Partial	√		Cloud server	An efficient heuristic algorithm based on semidefinite relaxation and a new random mapping method.
[35]	√	√		multi-user	Partial		√	MEC server	A joint optimization algorithm for the selection and unloading of vehicle edge computing servers
[36]		√	√	multi-user	Partial		√	MEC server	A multi-task two-layer computing offloading framework for heterogeneous networks

rative mobile edge computing device server. Reference [6] proposed a polynomial-complexity algorithm for computing the equal distribution of wireless and cloud resources in dense wireless networks to minimize the computing costs. The resource allocation problem of offloading computing to the mobile cloud by mobile users is considered, where a single mobile device can stream computations to the mobile cloud via multiple access points or a base station.

2) Computational Resource Allocation

The objective of computing resource allocation is to minimize the cost of task processing so that resources can be fully and reasonably utilized. It consists of two processes: task assignment, namely, the assignment of tasks that can be executed in parallel to specified resources, and resource allocation. The execution order of tasks is determined according to the pre-established resource allocation strategy. In the MEC scenario, computing resource allocation is also used to improve the overall system performance and to reduce the overall execution time and resource consumption. Computing resources are often considered in conjunction with offloading decisions, which can be divided into single-node computing resource allocation and multi-node computing resource allocation according to the numbers of users and computing nodes, where in single-node allocation, a base station can only serve one computing task in a time interval, and in multi-node allocation, a base station in a time gap can serve more than one computing task. In the multi-node computing resource allocation scenario, the main problems focus on communication interference between users and resource competition [41]. In reference [33], the author proposed a multi-user and multi-task offloading scheduling scheme in the updatable mobile edge cloud system. Considering the energy arrival of the mobile edge cloud and the dynamics of task arrival of various mobile devices, an energy acquisition strategy is proposed by combining energy acquisition with mobile edge cloud computing. To maximize the system utility and match the offloading energy consumption of the mobile edge cloud with the acquired energy, a task offloading scheduling scheme is proposed for mapping the computing workload of a mobile device to multiple wireless devices. In reference [34], a mobile cloud computing system that consists of multiple users, a computing access point and a remote cloud server is studied. An efficient heuristic algo-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2983609, IEEE Access

rithm is proposed for handling the joint task of loading and allocating computing and communication resources to minimize the energy consumption, the calculated weighted total cost, and the maximum latency among all users. In reference [35], an algorithm is designed for making joint selection decisions and calculating resources and the offload rate. A comprehensive task processing delay is used to develop the system utilities, which considers both the transmission and computation times. This scheme substantially improves the performance of load balancing and maximizes the system availability.

3) Mobility Management of Computational Offloading

Mobility management of computing resource offloading is of substantial significance to the integrity of the user computing offloading process. Due to the mobility of the user, it is inevitable that the user will be disconnected from the base station. Mobile cellular networks ensure the continuity and quality of service by switching among base stations. For scenarios with low user mobility, during the process of offloading the application to the MEC server, the power of the current base station can be adjusted adaptively to ensure uninterrupted service. If the user switches to a new service base station, virtual machine migration of the compute node is used to solve the problem. In reference [24], a joint task allocation, subchannel allocation and power allocation problem is formulated. Aiming at maximizing the total offloading rate, a hybrid computing shunt management framework for real-time traffic in 5G networks is proposed. A joint power control and channel allocation scheme is designed based on non-orthogonal multiple access and mobile edge computing. MEC can reduce the computing limitations and extend the service life of mobile devices; however, it will lead to the dense distribution of MEC servers. Although MEC servers are close to the mobile users, they face user-related challenges, which will affect the computing shunt. Reference [36] focuses on the joint computing of offloading and multi-task user correlation, and it studies the scheduling problem in distributed MEC systems with densely distributed MEC servers. To reduce the energy consumption or improve the performance, an efficient algorithm for calculating offloading is proposed by considering the distribution of the computational resources and the transmitted power. A comparison of computing offloading in MEC is presented in TABLE II.

III. ARTIFICIAL INTELLIGENCE IN INTERNET OF VEHICLES

DRL is an essential technology for realizing AI. DRL utilizes the advantages of deep neural networks (DNNs) to train the learning process, thereby improving the learning speed and performance of the RL algorithm and overcoming the unsuitability of reinforcement learning for large-scale networks. In this part, we introduce the development of AI, analyze the relationship between AI and DRL, discuss the theory and architecture of AI, and analyze the application of AI in IoV research.

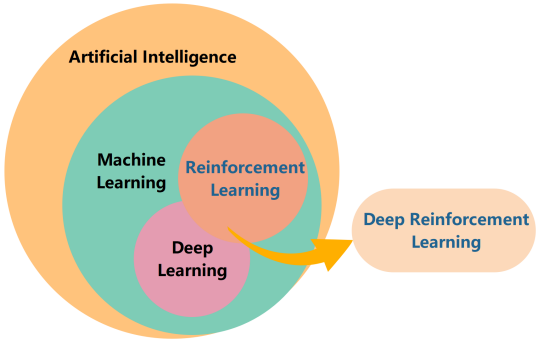


FIGURE 3: The relationship between AI, ML, RL, DL and DRL.

A. ARCHITECTURE

The objective of artificial intelligence (AI) is to endow machines with human intelligence. Machine learning (ML) is a method for implementing AI by using algorithms to parse data, learn from data, and make decisions and predictions regarding real-world events. Deep learning (DL) is a technology for realizing ML, which enables ML to realize many applications and expands the scope of AI. Reinforcement learning (RL), which is also known as evaluation learning, is a technique of ML. Deep reinforcement learning (DRL) is the combination of DL and RL. It aims at realizing the optimization objective of RL with the operation mechanism of DL to advance toward general AI. Fig. 3 illustrates the relationship among AI, ML, RL, DL and DRL.

AI is a promising approach for making vehicle networks intelligent. RL is a powerful tool in ML. In contrast to traditional ML, RL does not have an immediate end result; only a temporary reward (set primarily according to human experience) is observed. Therefore, RL can also be regarded as delayed supervised learning [16]. In the case of small state space and behavior space, RL technology can be used to enable network entities to identify the optimal strategy for decision-making or behavior. However, in a complex large-scale network, for improving the learning efficiency, a learning method that combines RL with DL, namely, DRL, is regarded as a potential solution [42]. The three key elements of RL are the system status, the system actions, and the rewards. In RL, the environment is typically represented as a Markov decision process (MDP). Agents interact with the unknown environment through repeated observation, action and reward to construct the optimal strategy [8]. Due to the limited data that are obtained from outside, DRL systems often rely on their own experience to learn by themselves. Via this approach, knowledge is acquired and solutions are adapted to the environment. For the spatial-temporal coverage problem in mobile crowdsensing systems, reference [43] proposes a vehicle selection scheme that is based on DRL.

Fig. 4 illustrates the architecture of AI in IoV, where the agent observes its current environmental state, takes

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2983609, IEEE Access

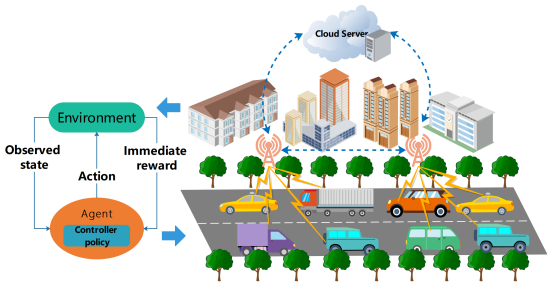


FIGURE 4: Architecture for AI in IoV.

action, and receives its immediate reward along with the new state. The observed information, which includes immediate rewards and the new status, is used to adjust the agent's strategy, and the process is repeated until the agent's strategy approaches the optimal strategy.

B. TECHNOLOGIES

1) Markov Decision Processes(MDP)

The theoretical basis of MDP is a Markov chain (MC), which is a stochastic process in a discrete index set and state space. MDP provides a mathematical framework for modeling decision problems in which the results are partially random and controlled by the decision maker or agent. MDP is used to model RL problems in ML, which facilitates the study of dynamic programming and optimization problems that can be solved via RL technology [42]. In reference [44], the author proposed an architecture that combines a satellite network with 5G cloud on-board Internet and designed a joint optimization problem of computation offloading under time-delay and cost constraints that is based on an incentive mechanism. The solution uses the Markov chain Monte Carlo and simulated annealing algorithms to effectively support seamless coverage and global resource management. To overcome the inability of mobile IP to cope with high-speed and frequent vehicle movements, reference [45] adopts a switching management scheme that is based on machine learning in a two-tier intelligent transportation network. In the first layer, the recursive neural network model is used to predict the received signal strength of the access point to obtain the switch trigger decision. In the second layer, the random Markov model is used to predict the next access point using the vehicle flow.

2) Q-learning and Deep Q-Learning(DQL)

Q-learning is a typical time-difference RL algorithm. The Q-function is defined for the evaluation of the long-term return of the strategy, and a neural network is used instead. For each event, Q-learning makes a decision that is based on the Q-value, which evaluates the selected operation in the current scenario [16].When state space and action space are small, Q-learning algorithm can effectively obtain the optimal strategy. However, in practical applications, these Spaces are often large due to the complexity of the system model. In this case, the Q-learning algorithm may not be able to find the optimal strategy. Therefore, the introduction of DQL algorithm can overcome this shortcoming [42].Using Q-learning or DQL algorithms can intelligently control the use of network resources in IoV [20].Reference [46] established a generic, green, intelligent, and scalable scheduling strategy for resource distribution, which is used to adapt to the randomness of the traffic environment, to learn from high-dimensional input scheduling policies using the depth of the Q-network, to support the efficient operation of the vehicle network and balance the IoT gateway of the available energy, and to minimize the total cost.

3) Long Short-Term Memory(LSTM)

LSTM enables a recurrent neural network (RNN) to evolve into one of many network topologies. It is a time-cycling neural network that can remember features in data at any time interval. LSTM is composed of forward components and backward components. LSTM solves the vanishing gradient problem of RNN by explicitly introducing a storage unit. LSTM can be used to create large recursive networks to facilitate the solution of difficult ordering problems in machine learning and to obtain the latest results [11]. In reference [47], the author used a Markov decision process to model the content caching problem in the Internet of vehicles and proposed an active caching strategy of Q-learning which is based on LSTM. In a service scenario in which Non-Orthogonal Multiple Access (NOMA) users are randomly deployed by a BS, reference [48] proposes a method that is based on deep learning, the NOMA technology with LSTM integration, a framework that can be automatically and completely learned via the method of offline learning in an unknown channel environment, end-to-end processing of a NOMA wireless channel, and optimization that is based on NOMA user activity and data detection.

C. APPLICATIONS

Compared with the traditional DRL-based centralized approach, the DRL-based distributed approach can learn information from the environment more quickly and can substantially reduce the communication overhead of vehicles. In reference [49], an intelligent unloading framework of a vehicle-mounted network that supports 5G is built. To balance the transmission load, the cellular channel and the multiplexed sub-channel are used for the task transmission. According to the bilateral matching algorithm, all users are divided into V2R and V2I users to allocate the unauthorized spectrum. Then, a distributed deep reinforcement learning algorithm is proposed for scheduling the cellular channel, which can minimize the unloading cost under the premise of satisfying individual delay constraints. Reference [50] uses an online learning algorithm based on reinforcement learning to propose a collaborative online caching strategy to achieve content caching and updating. In reference [51], the author proposes an RSU cloud, which is an infrastructure that

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2983609, IEEE Access

supports computing and communication in the Internet of vehicles, that utilizes the dynamic programmability of SDN and the cost analysis method of reconstruction. Modeling cloud resource management as a multi-objective optimization problem, with a heuristic algorithm and reinforcement learning approach for the selection of the configuration that minimizes the cost of reconfiguration, may yield high virtual machine mobility immediately, but in the long run, the opposite may occur. Most studies focus on the optimization of mobile edge networks in consideration of the network, communication and computing costs and cache. To satisfy the requirements of system resource management scheduling and system performance optimization, this is considered a promising solution for improving the predictive performance of channel state information in an edge computing network. In reference [11], a channel prediction model that is based on LSTM is proposed, and the associated algorithm, which is based on deep learning, can predict future channel parameters based on past and present channel parameters. Basic methods of machine learning often incur a large training cost. Samples are difficult to obtain in practice, and if the network parameters change, task mismatch easily occurs. Reference [7], the authors designed a transmission strategy that is based on deep learning by considering the social characteristics of the edge of vehicle equipment and physical properties, and they established a connection framework for assessing interactions, in which a clustering algorithm that is based on triangular patterns is used to control the network size and a discovery algorithm that is based on a convolutional neural network is used for data sharing with partners.

IV. ARTIFICIAL INTELLIGENCE EMPOWERED EDGE OF VEHICLES

In the age of intelligent IoV, the application of AI to vehicle edge networks is a promising approach for the development of intelligent transportation. In this part, we introduce the advantages of AI in applications to vehicle edge networks and the architecture of AI-based vehicle edge networks. Then, the related key technologies are described. Finally, we introduce the previous research on the application of AI to vehicle edge networks.

A. ARCHITECTURE

Traditional data sources are typically transferred remotely to the cloud center, and services that are based on the mobile cloud cannot guarantee the satisfaction of low-latency requirements for content transfer [52]. Therefore, mobile edge computing has the potential to overcome this challenge. According to [53], not only can MEC reduce the communication latency, but MEC nodes can also use the potential resources in the network to reduce the workload of the central base station. In reference [54], a joint communication, caching and computing (3C) model is proposed for the provision of infotainment services in smart cars. It minimizes the latency of access to infotainment services under resource constraints. The problem of mixed-integer, nonlinear and non-convex optimization is transformed into a linear programming problem via the relaxation technique, and its convergence is demonstrated. In addition, according to [55], by using the mobile edge cache to store the content on the edge of the network, the content can be transmitted directly via wireless transmission without the need for backhauling or core network transmission, thereby reducing the end-to-end delay and the backhaul pressure.

As the applications of mobile users become richer and more intelligent, they are faced with the requirements of massive data processing, delay-sensitivity, and location awareness, among others. In recent years, artificial intelligence (AI)-based vehicle edge computing has attracted substantial attention. DRL is a tool of machine learning. The available DRL technology can be applied to image processing, pattern recognition, natural language processing and computation-intensive applications. The integration of DRL technology and vehicle edge computing is used to construct the intelligent computing shunt system, which faces such problems as high mobility of vehicles and difficulty finding continuous image sequences. In reference [53], the authors use the finite-state Markov chain, DRL and the calculation integration vehicle edge to build an intelligent offloading system, and they develop a joint optimization of task scheduling and resource allocation problem in a traffic network, which is decomposed into two sub-optimization problems: task scheduling among multiple vehicles and the allocation of resources. The former is solved via a bilateral matching algorithm, while the latter is handled by an integrated DRL method. In addition, shared edge computing services can be provided by mobile edge servers that are deployed on the edge of the network to improve the user quality of service. However, due to the unevenness of space and the dynamics of time, the distribution of vehicles is unbalanced. Therefore, an unbalanced communication load of the mobile edge server is generated. Reference [56] proposes an active load balancing method, namely, an end-to-end load balancer, which uses a deep CNN to learn spatiotemporal correlations and predict road traffic conditions. A new framework that is based on CNN is used to address the optimization problem of NLP, to fine-tune the network from end to end, and to implement the efficient collaborative scheduling of cached data between mobile edge servers. In reference [57], the hybrid computation offloading and intelligent cache problems in layered IoV with edge intelligence are studied. However, to satisfy the demand of real-time analysis of heterogeneous data from an intelligent vehicle network and its environment, deep reinforcement learning still faces many challenges [58].

Fig. 5 illustrates the architecture of a vehicle edge network that is based on AI. The mobile vehicle communicates with a roadside unit (RSU) that is equipped with an MEC server via an on-board unit (OBU). RSUs have computing and storage capabilities, and multiple RSUs can communicate. Computationally heavy tasks can be offloaded to the base station (BS), and the collected data can be used to make intelligent decisions (such as predicting the direction of vehicle

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2983609, IEEE Access

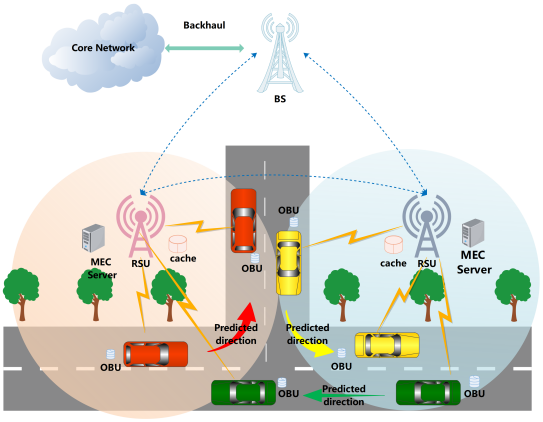


FIGURE 5: The architecture of vehicle edge network based on AI.

movement) with the help of high computing power, and can be used to support deep learning.

B. TECHNOLOGIES

1) Collaborative edge caching

Collaborative edge caching combines the advantages of mobile edge caching and collaborative caching. By actively storing files in the base station (BS), the mobile edge cache can provide content directly without remote file extraction, which reduces the end-to-end latency and the backhaul stress. Simultaneously, to effectively utilize the limited cache size, collaborative caching can be used to improve the diversity of the cache. Under the new user-centered network architecture, multiple base stations can serve users. In addition, collaborative caching can improve the cache hit performance to overcome the moderate cache hit performance bottleneck that is caused by the relatively small cache storage on a single BS. In reference [55], collaborative edge caching in large-scale user-centered clustered mobile networks is studied, and a greedy content layout algorithm that is based on optimal bandwidth allocation is proposed for minimizing the cache size and the average file transfer rate under bandwidth constraints. In reference [59], the author proposed an edge network cache replacement strategy that is based on deep learning using a deep LSTM network. The joint framework is used to merge the smart cache replacement algorithm and the corresponding collaboration mechanism. The cache strategy is automatically learned in real time from the request sequence to reduce the transmission latency and the backhaul data traffic.

2) Multi Armed Bandit

MAB is a reinforcement learning method. It has been extensively studied for addressing the key tradeoff between exploration and development in sequential decision-making under uncertain conditions. The original k-armed bandit problem assumed that one option was selected from k options repeatedly. The option is the arm. Each time an option is selected, a reward is obtained as feedback, and the action selection is repeated to focus the action on the best arm to maximize the expected total reward in a period of time. To cope with the unknown service requirements in the changing user groups, reference [60] proposed a combined context bandit learning problem. A spatiotemporal edge service placement algorithm is used to solve the problem. Multiple learners are considered, and each learner can maintain a distinct location-specific context space. The context information of connected users is collected according to users' preferences, and location-awareness and context-awareness are realized for renting computing resources flexibly and economically in the shared edge computing platform. Reference [61] proposed a distributed adaptive task offloading algorithm that is based on learning that is based on multi-arm bandit theory. It enables the vehicle to learn the offloading delay performance of an adjacent vehicle while offloading the calculation task, eliminates the need for frequent state exchanges, and increases the input and occurrence awareness for adaptation to the dynamic environment.

3) Nonorthogonal multiple access (NOMA)

As an emerging technology in 5G networks, NOMA has advantages in terms of its spectrum, connectivity, energy efficiency and other aspects, thereby enabling multiple users to reuse frequency resources nonorthogonally. NOMA technology not only has advantages in increasing the system throughput and supporting large-scale connections but can also be used to eliminate multi-user interference in multi-user detection systems by assigning power levels according to users' channel conditions. Reference [65] studied the cache-assisted non-orthogonal multiplexing access of the onboard network that supports 5G. Considering the full-file cache and split-file cache, the optimization problem of the overall probability of decoding the files successfully in each vehicle is formulated and solved in the first scenario. In the second scenario, a joint power distribution optimization problem is proposed for determining the power distribution between vehicles and individual files. In reference [66], the author considers task offloading and user selection between macro units and edge devices. A moving edge algorithm that is based on non-orthogonal multiple access is proposed, and a heuristic algorithm is designed from the aspects of offloading decision, channel allocation and power control to improve the transmission rate gain and the discharge offloading efficiency.

C. APPLICATIONS

The emerging 5G mobile network has the advantages of high bandwidth and low latency. By expanding the antenna scale, the 5G wireless network can improve frequency reuse and increase the capacity of the cellular network via network densification. However, in the face of massive data, the traditional caching strategy has encountered a bottleneck. Therefore, the proposal of a moving edge cache is extremely important in the 5G network, which can provide higher service quality for many new applications. Caching content in the base station can significantly reduce the network

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2983609, IEEE Access

TABLE 3: Comparison of Edge Caching Based on AI Algorithm

Cache Type	Ref.	Optimization Objective	Program	Result
Non-cooperative cache	[52]	Minimize content transfer latency	Content caching strategy based on two-dimensional Markov chain	It is verified that the prediction accuracy increases and the delay of content transmission gradually decreases
	[62]	Minimize the long-term cost of acquiring IoT data	Framework of Internet of Things system based on edge cache	This solution can reduce the long-term cost for users to obtain IoT data
Cooperative caching	[18]	Maximize system availability	An edge computing and caching scheme based on AI algorithm	This solution greatly improves the practicability of the system
	[27]	Reduce latency, network burden and communication redundancy	An integrated framework that dynamically coordinates network, cache, and computing resources	The framework provides network functions, caching functions and computing functions
	[55]	Minimize the transmission rate under bandwidth constraints	A greedy content layout algorithm based on optimal bandwidth allocation	Reduced average file transfer latency to 45%
	[59]	Reduce transmission latency and backhaul network load	An edge network cache replacement strategy based on deep learning	14% to 22% reduction in overall transmission latency and 15% to 23% savings in backhaul data traffic
	[63]	Reduce latency due to backhaul bandwidth consumption	Convolutional neural network analysis method and Multi-Layer Perceptron method to predict cache content	The accuracy of predicting infotainment content to be cached can reach 99.28%
	[64]	Minimize system cost under latency constraints	A multi-time-scale framework based on artificial intelligence	The scheme effectively mitigates the harmful effects of limited backhaul capacity and low BS computing resources

latency, whereas caching content on the edge can reduce the data traffic in the core network and conserve bandwidth for the Internet [59]. In addition, edge caching can improve the spectrum efficiency and reduce the energy consumption due to device heterogeneity and dense deployment [67]. IoT data are transient; for example, the popularity distribution of data may vary with the time and location, and static-based caching strategies have difficulty satisfying the various requirements of IoT services, such as mobility and geographically distributed support. In reference [62], for caching temporary data on the edge of the IoT, the author proposed a framework of the IoT system that is based on the edge cache. Considering both the "data freshness" and the "edge cache", the cache strategy of the deep reinforcement learning method can make smart cache decisions without assuming the data popularity or user request distribution. In reference [63], the author uses a convolutional neural network to predict and obtain the user's age and gender characteristics. By deploying multiple-access edge computing servers on roadside units, WiFi access points and acer stations for caching infotainment content in and around self-driving cars, fog computing extends the infrastructure of traditional cloud computing to the edge of the network, thereby substantially reducing the long-distance latency from the terminal to the cloud server. Since edge servers are distributed in the surrounding area, fog computing is expected to improve the data transmission efficiency. Reference [52] proposed a vehicle edge collaborative filtering content transmission scheme that is based on fog calculation. A collective filtering algorithm and a two-dimensional Markov chain are used to combine positional awareness, content caching, and decentralized computing for content precaching at the edge of the vehicle network. Due to the highly dynamic network environment and the uncertainty of mobile users, reference [5] proposed the concept of vehicle caching, which uses vehicle mobility to improve the service scope and cache capacity. The interaction between a cached vehicle and the mobile user is modeled as a two-dimensional Markov process. On this basis, an online vehicle cache design scheme that is based on network energy efficiency optimization is proposed. It is proved to outperform the available scheme in terms of the hit ratio, energy efficiency, cache utilization and system gain. Machine learning is also an emerging tool for solving caching, computing and communication problems in 5G wireless communication. Various studies, such as [18] and [27], have investigated the joint optimization of computing resources and caching. In reference [18], large amounts of data and popular content are produced by computation-intensive applications, time-delay-sensitive applications, and on-board sensors. This paper discusses the resource processing and storage of vehicles with limited resources in the Internet of vehicles. An AI-based algorithm is proposed for dynamically orchestrating the architecture of edge computing and cache resources, and a novel resource management scheme is developed, which uses deep reinforcement learning. In contrast to other studies, this study uses a two-layer cross-layer offloading model that combines a heterogeneous network and mobile edge computing to realize dynamic resource allocation. In reference [27], the principle of programmable control of the network that is defined by software, the principle of caching in information and communication technology and the principle of network virtualization are used to construct the framework of the dynamic arrangement of integrated network, cache and computing resources. The main disadvantage is the lack of consideration of the energy consumption. In references [64] and [8], the joint optimization of resource allocation

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2983609, IEEE Access

for caching, computing and communication is considered. In reference [64], an algorithm that is based on an AI multi-temporal framework is designed, which facilitates the configuration of cache placement and the calculation of the parameters of resource allocation. For cost minimization under the constraints of limited RSU storage capacity, dynamic fluctuations in computing resources, vehicle mobility, and strict end-to-end delay limits, reference [8] proposed a deep reinforcement learning method that is based on a multi-time-scale framework. Vehicle mobility is leveraged to enhance the caching and computing strategies. A long-time-scale model of motion perception reward estimation is proposed to reduce the complexity that is due to large action spaces. Resource allocation and computational offloading are inextricably linked. Reference [68] studied the optimal utility task offloading scheme in a heterogeneous vehicle network with multiple mobile edge computing servers under constraints on the reliability and waiting time and proposed an adaptive redundancy offloading algorithm that is based on deep Q-learning to ensure the reliability of offloading and to improve the practicability of the system. Reference [16] proposed an energy-saving task offloading scheme that is based on DRL and combined it with fog computing technology. Considering load balancing and time delay constraints, an optimization problem was formulated for minimizing the energy consumption of traffic offloading, which was decomposed into two parts: flow redirection and offloading decision. Algorithms that are based on Edmonds-Karp and DRL were developed for solving the problem. In reference [69], a distributed dynamic computing offloading strategy that is based on DRL is proposed for dynamic task offloading control of multi-user MEC systems to minimize the long-term average computing cost consumption and the task buffer delay in the power. A comparison of edge caching that is based on AI algorithms is presented in TABLE III.

V. RESEARCH CHALLENGE AND OPEN ISSUES

In the previous sections, we reviewed the architecture and related technologies of MEC, AI, and AI-based vehicle edge networks in IoV. In addition, we analyzed the previous research from three aspects. However, the future IoV still faces challenges. In this section, we will discuss several possible research challenges and propose several promising research directions.

A. SECURITY AND PRIVACY

In recent years, security and privacy issues in IoV have received extensive attention. Mobile vehicles collect information via V2V communication between vehicles and via V2I communication from vehicles to roadside infrastructure. Due to the high mobility of vehicles, communication is often interrupted, thereby resulting in frequent failures of communication links. In addition, hackers' security attacks on communication channels and sensor tampering will lead to severe privacy invasion. In addressing these security and privacy issues, challenges in the solution of identity privacy, data privacy and location privacy issues will be encountered. Potential solutions include communication authentication, MEC and access control of cloud computing servers [70]. Reference [71] proposes an architecture of edge auxiliary network connecting vehicles. To solve the problem of location privacy, a location-based differential privacy protection service framework is proposed for ensuring location privacy within the coverage of the edge nodes. Li et al. [72] proposed an online double auction scheme for k-anonymous location privacy protection, which could solve the problems of optimal charging scheduling for electric vehicles and location privacy protection for owners of electric vehicles. Chen et al. [73] designed a data trading method for the Internet of vehicles that is based on block chain. An iterative dual auction mechanism is used to protect the privacy of both parties in data transaction, to reduce the data transmission cost and to improve the system stability.

B. GREEN ENERGY SAVING

Green energy saving has a profound impact on the construction of a green IoV. Automobile exhaust emission is the main factor that affects the human environment and the air environment. To alleviate the current environmental pollution scenario while adapting to the highly dynamic traffic environment, it is highly important to use RSUs to communicate with nearby vehicles to realize efficient task scheduling to satisfy vehicle communication requirements. Energy saving in RSU scheduling and RSU energy collection are essential for solving the problem of energy consumption. The implementation of a wind or solar RSU in an energy-constrained vehicle environment can increase the network capacity and promote energy recovery. In addition, the minimum number of active RSUs can be set to maintain the network operation and connectivity [74]. To minimize the total energy consumption of RSUs under the delay constraint, reference [75] constructed an MEC-based IoV energy-saving scheduling framework for balancing the computing tasks among RSUs. A heuristic algorithm is designed that considers the task scheduling among MEC servers and the energy consumption of the RSU downlink. In addition, electric vehicles, which are powered by electric engines instead of internal combustion engines, which are powered by fossil fuels, can effectively reduce the carbon footprint and play an important role in realizing efficient energy management [76].

C. HIGH MOBILITY

Mobility is an important feature of vehicle networks. With the rapid increase of the road traffic density, high speed and frequent vehicle movements are the main factors that render the network topology dynamic. The high mobility of intelligent vehicles not only adds considerable complexity in co-optimizing the allocation of computing and cache resources but also hinders the provision of stable and reliable wireless communication [77]. First, the data transmission distance is constantly changing due to vehicle movements. Therefore, the data rate and the effective duration of channel

transmission in V2X communication will also be affected. Second, the changes in the vehicle speed and direction over time will lead to frequent handovers between edge servers. Active communication management is transferred from one RSU or BS to another RSU or BS. The duration for which a vehicle remains within the coverage area of the RSU or BS also varies. Due to the widespread use of various Global Positioning System (GPS) devices and mobile Internet in daily life, vehicle trajectory data can be easily obtained on a large scale [78]. Therefore, addressing the high mobility of IoV by predicting the vehicle movement direction and studying data routing distribution protocols is a feasible solution [79].

D. INTELLIGENT COMPUTATION

With many edge nodes deployed in 5G networks, edge computing has the advantage of reducing the traffic load and the backhaul pressure, but edge devices still face the challenge of real-time processing. Edge cognitive computing has become a new paradigm. By analyzing and interpreting the available data and information in cyberspace, the intelligence of machines can be increased for the prediction and generation of new information, thereby providing more intelligent cognitive services. Reference [80] proposed an architecture of edge cognitive computing by combining edge computing and cognitive computing. Considering the elastic distribution of cognitive computing services and the mobility of users, a dynamic cognitive service migration mechanism that is based on edge cognitive computing is designed. It integrates the communication, computing, storage and application on the edge network, improves the user experience and realizes rational resource allocation and cognitive information circulation.

VI. CONCLUSION

In this study, two key technologies, namely, MEC and AI, were analyzed by focusing on the development of intelligent IoV and the previous research on combining the two technologies. First, the communication mode and architecture of the traditional Internet of vehicles were introduced, along with the advantages of the emerging 5G network. In addition, MEC, FC and MCC were compared by studying the development history of MEC. The advantages of MEC were analyzed, the MEC, FC and MCC were compared by studying the development history of MEC. The advantages of MEC were analyzed, the key technologies of MEC were evaluated, and several key technologies for calculating the unloading in MEC were studied. Then, the differences and connection between AI and DRL in realizing intelligent IoV were discussed, with a focus on the characteristics and application status of DRL in dynamic vehicle networks. Then, combining the advantages of MEC and AI technologies, the previous research on the application of AI to vehicle edge networks was analyzed. Finally, the possible future research directions of IoV were discussed.

REFERENCES

[1] L. Mendiboure, M. A. Chalouf, and F. Krief, "Towards a 5g vehicular architecture," in International Workshop on Communication Technologies for Vehicles. Springer, 2019, pp. 3–15.

[2] S. Guo, C. Chen, J. Wang, Y. Liu, X. Ke, Z. Yu, D. Zhang, and D.-M. Chiu, "Rod-revenue: Seeking strategies analysis and revenue prediction in ride-on-demand service using multi-source urban data," IEEE Transactions on Mobile Computing, 2019.

[3] C. Chen, D. Zhang, X. Ma, B. Guo, L. Wang, Y. Wang, and E. Sha, "Crowddeliver: planning city-wide package delivery paths leveraging the crowd of taxis," IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 6, pp. 1478–1496, 2016.

[4] P. Lin, Q. Song, and A. Jamalipour, "Multidimensional cooperative caching in comp-integrated ultra-dense cellular networks," IEEE Transactions on Wireless Communications, 2019.

[5] Y. Zhang, C. Li, T. H. Luan, Y. Fu, W. Shi, and L. Zhu, "A mobility-aware vehicular caching scheme in content centric networks: Model and optimization," IEEE Transactions on Vehicular Technology, vol. 68, no. 4, pp. 3100–3112, 2019.

[6] S. Jošilo and G. Dán, "Selfish decentralized computation offloading for mobile cloud computing in dense wireless networks," IEEE Transactions on Mobile Computing, vol. 18, no. 1, pp. 207–220, 2018.

[7] Z. Ning, Y. Feng, M. Collotta, X. Kong, X. Wang, L. Guo, X. Hu, and B. Hu, "Deep learning in edge of vehicles: Exploring trirelationship for data transmission," IEEE Transactions on Industrial Informatics, vol. 15, no. 10, pp. 5737–5746, 2019.

[8] R. Q. Hu et al., "Mobility-aware edge caching and computing in vehicle networks: A deep reinforcement learning," IEEE Transactions on Vehicular Technology, vol. 67, no. 11, pp. 10 190–10 203, 2018.

[9] C. Chen, Y. Ding, Z. Wang, J. Zhao, B. Guo, and D. Zhang, "Vtracer: When online vehicle trajectory compression meets mobile edge computing," IEEE Systems Journal, 2019.

[10] Z. Ning, F. Xia, N. Ullah, X. Kong, and X. Hu, "Vehicular social networks: Enabling smart mobility," IEEE Communications Magazine, vol. 55, no. 5, pp. 16–55, 2017.

[11] G. Liu, Y. Xu, Z. He, Y. Rao, J. Xia, and L. Fan, "Deep learning-based channel prediction for edge computing networks toward intelligent connected vehicles," IEEE Access, vol. 7, pp. 114 487–114 495, 2019.

[12] I. Sorkhoh, D. Ebrahimi, R. Atallah, and C. Assi, "Workload scheduling in vehicular networks with edge cloud capabilities," IEEE Transactions on Vehicular Technology, vol. 68, no. 9, pp. 8472–8486, 2019.

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

[14] Z. Ning, P. Dong, X. Wang, X. Hu, L. Guo, B. Hu, Y. Guo, T. Qiu, and R. Y. Kwok, "Mobile edge computing enabled 5g health monitoring for internet of medical things: A decentralized game theoretic approach," IEEE Journal on Selected Areas in Communications, 2020.

[15] P. Dong, Z. Ning, M. S. Obaidat, P. Jiang, Y. Guo, X. Hu, B. Hu, and B. Sadoun, "Edge computing based healthcare systems: Enabling decentralized health monitoring in internet of medical things," IEEE Network, 2020.

[16] Z. Ning, P. Dong, X. Wang, L. Guo, J. J. Rodrigues, X. Kong, J. Huang, and R. Y. Kwok, "Deep reinforcement learning for intelligent internet of vehicles: An energy-efficient computational offloading scheme," IEEE Transactions on Cognitive Communications and Networking, vol. 5, no. 4, pp. 1060–1072, 2019.

[17] Z. Ning, X. Hu, Z. Chen, M. Zhou, B. Hu, J. Cheng, and M. S. Obaidat, "A cooperative quality-aware service access system for social internet of vehicles," IEEE Internet of Things Journal, vol. 5, no. 4, pp. 2506–2517, 2017.

[18] Y. Dai, D. Xu, S. Maharjan, G. Qiao, and Y. Zhang, "Artificial intelligence empowered edge computing and caching for internet of vehicles," IEEE Wireless Communications, vol. 26, no. 3, pp. 12–18, 2019.

[19] K.-H. N. Bui and J. J. Jung, "Aco-based dynamic decision making for connected vehicles in iot system," IEEE Transactions on Industrial Informatics, vol. 15, no. 10, pp. 5648–5655, 2019.

[20] H. Yang, X. Xie, and M. Kadoch, "Intelligent resource management based on reinforcement learning for ultra-reliable and low-latency iov communication networks," IEEE Transactions on Vehicular Technology, vol. 68, no. 5, pp. 4157–4169, 2019.

[21] C. Zhu, J. Tao, G. Pastor, Y. Xiao, Y. Ji, Q. Zhou, Y. Li, and A. Ylä-Jääski, "Folo: Latency and quality optimized task allocation in vehicular

fog computing," IEEE Internet of Things Journal, vol. 6, no. 3, pp. 4150–4161, 2018.

[22] X. Wang, Z. Ning, and L. Wang, "Offloading in internet of vehicles: A fog-enabled real-time traffic management system," IEEE Transactions on Industrial Informatics, vol. 14, no. 10, pp. 4568–4578, 2018.

[23] Z. Ning, J. Huang, and X. Wang, "Vehicular fog computing: Enabling real-time traffic management for smart cities," IEEE Wireless Communications, vol. 26, no. 1, pp. 87–93, 2019.

[24] Z. Ning, X. Wang, J. J. Rodrigues, and F. Xia, "Joint computation offloading, power allocation, and channel assignment for 5g-enabled traffic management systems," IEEE Transactions on Industrial Informatics, vol. 15, no. 5, pp. 3058–3067, 2019.

[25] C. Yang, Y. Liu, X. Chen, W. Zhong, and S. Xie, "Efficient mobility-aware task offloading for vehicular edge computing networks," IEEE Access, vol. 7, pp. 26 652–26 664, 2019.

[26] H. Peng, Q. Ye, and X. S. Shen, "Sdn-based resource management for autonomous vehicular networks: A multi-access edge computing approach," IEEE Wireless Communications, vol. 26, no. 4, pp. 156–162, 2019.

[27] Y. He, F. R. Yu, N. Zhao, V. C. Leung, and H. Yin, "Software-defined networks with mobile edge computing and caching for smart cities: A big data deep reinforcement learning approach," IEEE Communications Magazine, vol. 55, no. 12, pp. 31–37, 2017.

[28] M. Chen and Y. Hao, "Task offloading for mobile edge computing in software defined ultra-dense network," IEEE Journal on Selected Areas in Communications, vol. 36, no. 3, pp. 587–597, 2018.

[29] J. Zhao, Q. Li, Y. Gong, and K. Zhang, "Computation offloading and resource allocation for cloud assisted mobile edge computing in vehicular networks," IEEE Transactions on Vehicular Technology, vol. 68, no. 8, pp. 7944–7956, 2019.

[30] W. Hou, Z. Ning, and L. Guo, "Green survivable collaborative edge computing in smart cities," IEEE Transactions on Industrial Informatics, vol. 14, no. 4, pp. 1594–1605, 2018.

[31] 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 of Things Journal, vol. 6, no. 3, pp. 4804–4814, 2018.

[32] X. Hu, K.-K. Wong, and K. Yang, "Wireless powered cooperation-assisted mobile edge computing," IEEE Transactions on Wireless Communications, vol. 17, no. 4, pp. 2375–2388, 2018.

[33] W. Chen, D. Wang, and K. Li, "Multi-user multi-task computation offloading in green mobile edge cloud computing," IEEE Transactions on Services Computing, vol. 12, no. 5, pp. 726–738, 2018.

[34] M.-H. Chen, M. Dong, and B. Liang, "Resource sharing of a computing access point for multi-user mobile cloud offloading with delay constraints," IEEE Transactions on Mobile Computing, vol. 17, no. 12, pp. 2868–2881, 2018.

[35] Y. Dai, D. Xu, S. Maharjan, and Y. Zhang, "Joint load balancing and offloading in vehicular edge computing and networks," IEEE Internet of Things Journal, vol. 6, no. 3, pp. 4377–4387, 2018.

[36] Y. Dai, D. Xu, S. Maharjan, and Y. Zhang, "Joint computation offloading and user association in multi-task mobile edge computing," IEEE Transactions on Vehicular Technology, vol. 67, no. 12, pp. 12 313–12 325, 2018.

[37] X. Wang, Z. Ning, X. Hu, L. Wang, B. Hu, J. Cheng, and V. C. Leung, "Optimizing content dissemination for real-time traffic management in large-scale internet of vehicle systems," IEEE Transactions on Vehicular Technology, vol. 68, no. 2, pp. 1093–1105, 2018.

[38] Y. Hui, Z. Su, T. H. Luan, and J. Cai, "Content in motion: An edge computing based relay scheme for content dissemination in urban vehicular networks," IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 8, pp. 3115–3128, 2018.

[39] H. Zhou, N. Cheng, J. Wang, J. Chen, Q. Yu, and X. Shen, "Toward dynamic link utilization for efficient vehicular edge content distribution," IEEE Transactions on Vehicular Technology, vol. 68, no. 9, pp. 8301–8313, 2019.

[40] T. Deng, P. Fan, and D. Yuan, "Optimizing retention-aware caching in vehicular networks," IEEE Transactions on Communications, vol. 67, no. 9, pp. 6139–6152, 2019.

[41] Q. Qi, J. Wang, Z. Ma, H. Sun, Y. Cao, L. Zhang, and J. Liao, "Knowledge-driven service offloading decision for vehicular edge computing: A deep reinforcement learning approach," IEEE Transactions on Vehicular Technology, vol. 68, no. 5, pp. 4192–4203, 2019.

[42] N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y.-C. Liang, and D. I. Kim, "Applications of deep reinforcement learning in communications and networking: A survey," IEEE Communications Surveys & Tutorials, vol. 21, no. 4, pp. 3133–3174, 2019.

[43] C. Wang, X. Gaimu, C. Li, H. Zou, and W. Wang, "Smart mobile crowdsensing with urban vehicles: A deep reinforcement learning perspective," IEEE Access, vol. 7, pp. 37 334–37 341, 2019.

[44] M. LiWang, S. Dai, Z. Gao, X. Du, M. Guizani, and H. Dai, "A computation offloading incentive mechanism with delay and cost constraints under 5g satellite-ground iov architecture," IEEE Wireless Communications, vol. 26, no. 4, pp. 124–132, 2019.

[45] N. Aljeri and A. Boukerche, "A two-tier machine learning-based handover management scheme for intelligent vehicular networks," Ad Hoc Networks, vol. 94, p. 101930, 2019.

[46] R. F. Atallah, C. M. Assi, and M. J. Khabbaz, "Scheduling the operation of a connected vehicular network using deep reinforcement learning," IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 5, pp. 1669–1682, 2018.

[47] L. Hou, L. Lei, K. Zheng, and X. Wang, "A q-learning-based proactive caching strategy for non-safety related services in vehicular networks," IEEE Internet of Things Journal, vol. 6, no. 3, pp. 4512–4520, 2018.

[48] G. Gui, H. Huang, Y. Song, and H. Sari, "Deep learning for an effective nonorthogonal multiple access scheme," IEEE Transactions on Vehicular Technology, vol. 67, no. 9, pp. 8440–8450, 2018.

[49] Z. Ning, P. Dong, X. Wang, M. S. Obaidat, X. Hu, L. Guo, Y. Guo, J. Huang, B. Hu, and Y. Li, "When deep reinforcement learning meets 5g-enabled vehicular networks: A distributed offloading framework for traffic big data," IEEE Transactions on Industrial Informatics, vol. 16, no. 2, pp. 1352–1361, 2020.

[50] P. Lin, Q. Song, J. Song, A. Jamalipour, and F. R. Yu, "Cooperative caching and transmission in comp-integrated cellular networks using reinforcement learning," IEEE Transactions on Vehicular Technology, 2020.

[51] M. A. Salahuddin, A. Al-Fuqaha, and M. Guizani, "Software-defined networking for rsu clouds in support of the internet of vehicles," IEEE Internet of Things Journal, vol. 2, no. 2, pp. 133–144, 2014.

[52] X. Wang, Y. Feng, Z. Ning, X. Hu, X. Kong, B. Hu, and Y. Guo, "A collective filtering based content transmission scheme in edge of vehicles," Information Sciences, vol. 506, pp. 161–173, 2020.

[53] Z. Ning, P. Dong, X. Wang, J. J. Rodrigues, and F. Xia, "Deep reinforcement learning for vehicular edge computing: An intelligent offloading system," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 10, no. 6, p. 60, 2019.

[54] S. A. Kazmi, T. N. Dang, I. Yaqoob, A. Ndikumana, E. Ahmed, R. Hussain, and C. S. Hong, "Infotainment enabled smart cars: A joint communication, caching, and computation approach," IEEE Transactions on Vehicular Technology, vol. 68, no. 9, pp. 8408–8420, 2019.

[55] S. Zhang, P. He, K. Suto, P. Yang, L. Zhao, and X. Shen, "Cooperative edge caching in user-centric clustered mobile networks," IEEE Transactions on Mobile Computing, vol. 17, no. 8, pp. 1791–1805, 2017.

[56] J. Li, G. Luo, N. Cheng, Q. Yuan, Z. Wu, S. Gao, and Z. Liu, "An end-to-end load balancer based on deep learning for vehicular network traffic control," IEEE Internet of Things Journal, vol. 6, no. 1, pp. 953–966, 2018.

[57] Z. Ning, K. Zhang, X. Wang, L. Guo, X. Hu, J. Huang, B. Hu, and R. Y. Kwok, "Edge intelligence for internet of vehicles: A joint hybrid computation offloading and intelligent caching algorithm," IEEE Transactions on Intelligent Transportation Systems, 2020.

[58] A. Ferdowsi, U. Challita, and W. Saad, "Deep learning for reliable mobile edge analytics in intelligent transportation systems: An overview," IEEE Vehicular Technology Magazine, vol. 14, no. 1, pp. 62–70, 2019.

[59] H. Pang, J. Liu, X. Fan, and L. Sun, "Toward smart and cooperative edge caching for 5g networks: A deep learning based approach," in 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). IEEE, 2018, pp. 1–6.

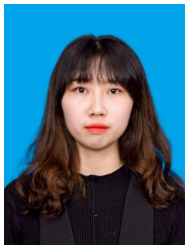
[60] L. Chen, J. Xu, S. Ren, and P. Zhou, "Spatio–temporal edge service placement: A bandit learning approach," IEEE Transactions on Wireless Communications, vol. 17, no. 12, pp. 8388–8401, 2018.

[61] Y. Sun, X. Guo, J. Song, S. Zhou, Z. Jiang, X. Liu, and Z. Niu, "Adaptive learning-based task offloading for vehicular edge computing systems," IEEE Transactions on Vehicular Technology, vol. 68, no. 4, pp. 3061–3074, 2019.

[62] H. Zhu, Y. Cao, X. Wei, W. Wang, T. Jiang, and S. Jin, "Caching transient data for internet of things: A deep reinforcement learning approach," IEEE Internet of Things Journal, vol. 6, no. 2, pp. 2074–2083, 2018.

[63] A. Ndikumana and C. S. Hong, "Self-driving car meets multi-access edge computing for deep learning-based caching," in 2019 International Conference on Information Networking (ICOIN). IEEE, 2019, pp. 49–54.

[64] R. Q. Hu, L. Hanzo et al., "Twin-timescale artificial intelligence aided mobility-aware edge caching and computing in vehicular networks," IEEE Transactions on Vehicular Technology, vol. 68, no. 4, pp. 3086–3099, 2019.
[65] S. Gurugopinath, P. C. Sofotasios, Y. Al-Hammadi, and S. Muhaidat, "Cache-aided non-orthogonal multiple access for 5g-enabled vehicular networks," IEEE Transactions on Vehicular Technology, vol. 68, no. 9, pp. 8359–8371, 2019.
[66] Z. Ning, X. Wang, and J. Huang, "Mobile edge computing-enabled 5g vehicular networks: Toward the integration of communication and computing," IEEE Vehicular Technology Magazine, vol. 14, no. 1, pp. 54–61, 2018.
[67] Z. Ning, X. Kong, F. Xia, W. Hou, and X. Wang, "Green and sustainable cloud of things: Enabling collaborative edge computing," IEEE Communications Magazine, vol. 57, no. 1, pp. 72–78, 2018.
[68] K. Zhang, Y. Zhu, S. Leng, Y. He, S. Maharjan, and Y. Zhang, "Deep learning empowered task offloading for mobile edge computing in urban informatics," IEEE Internet of Things Journal, vol. 6, no. 5, pp. 7635–7647, 2019.
[69] Z. Chen and X. Wang, "Decentralized computation offloading for multi-user mobile edge computing: A deep reinforcement learning approach," arXiv preprint arXiv:1812.07394, 2018.
[70] J. A. Onieva, R. Rios, R. Roman, and J. Lopez, "Edge-assisted vehicular networks security," IEEE Internet of Things Journal, vol. 6, no. 5, pp. 8038–8045, 2019.
[71] L. Zhou, L. Yu, S. Du, H. Zhu, and C. Chen, "Achieving differentially private location privacy in edge-assistant connected vehicles," IEEE Internet of Things Journal, vol. 6, no. 3, pp. 4472–4481, 2018.
[72] D. Li, Q. Yang, D. An, W. Yu, X. Yang, and X. Fu, "On location privacy-preserving online double auction for electric vehicles in microgrids," IEEE Internet of Things Journal, vol. 6, no. 4, pp. 5902–5915, 2018.
[73] C. Chen, J. Wu, H. Lin, W. Chen, and Z. Zheng, "A secure and efficient blockchain-based data trading approach for internet of vehicles," IEEE Transactions on Vehicular Technology, vol. 68, no. 9, pp. 9110–9121, 2019.
[74] X. Wang, Z. Ning, X. Hu, L. Wang, L. Guo, B. Hu, and X. Wu, "Future communications and energy management in the internet of vehicles: Toward intelligent energy-harvesting," IEEE Wireless Communications, vol. 26, no. 6, pp. 87–93, 2019.
[75] Z. Ning, J. Huang, X. Wang, J. J. Rodrigues, and L. Guo, "Mobile edge computing-enabled internet of vehicles: Toward energy-efficient scheduling," IEEE Network, vol. 33, no. 5, pp. 198–205, 2019.
[76] J. James, W. Yu, and J. Gu, "Online vehicle routing with neural combinatorial optimization and deep reinforcement learning," IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 10, pp. 3806–3817, 2019.
[77] Z. Ning, F. Xia, X. Hu, Z. Chen, and M. S. Obaidat, "Social-oriented adaptive transmission in opportunistic internet of smartphones," IEEE Transactions on Industrial Informatics, vol. 13, no. 2, pp. 810–820, 2016.
[78] C. Chen, Y. Ding, X. Xie, S. Zhang, Z. Wang, and L. Feng, "Trajcompressor: an online map-matching-based trajectory compression framework leveraging vehicle heading direction and change," IEEE Transactions on Intelligent Transportation Systems, 2019.
[79] Z. Ning, L. Liu, F. Xia, B. Jedari, I. Lee, and W. Zhang, "Cais: A copy adjustable incentive scheme in community-based socially aware networking," IEEE Transactions on Vehicular Technology, vol. 66, no. 4, pp. 3406–3419, 2016.
[80] M. Chen, W. Li, G. Fortino, Y. Hao, L. Hu, and I. Humar, "A dynamic service migration mechanism in edge cognitive computing," ACM Transactions on Internet Technology (TOIT), vol. 19, no. 2, pp. 1–15, 2019.



HONGJING JI received the B.Sc. degrees in Software Engineering in 2020 from the Taiyuan University of Technology, Taiyuan, China. She is currently working toward the M.Sc. degree in School of Software, Dalian University of Technology, Dalian, China. Her research interests include edge computing, internet of vehicle and resource management.

OSAMA ALFARRAJ received the master's and Ph.D. degrees in information and communication technology (ICT) from Griffith University, in 2008 and 2013, respectively. His doctoral dissertation investigates the factors influencing the development of Government in Saudi Arabia, and it is a qualitative investigation of the developers' perspectives. He is currently an Associate Professor with ICT, King Saud University, Riyadh, Saudi Arabia. His research interests include electronic commerce, M-government, the Internet of Things, cloud computing, AI, and big data analytics.

AMR TOLBA received the M.Sc. and Ph.D.degrees from the Faculty of Science, MenoufiaUniversity, Egypt, in 2002 and 2006, respec-tively. He is currently an Associate Professor withthe Faculty of Science, Menoufia University. Heis on leave from Menoufia Univesity with theComputer Science Department, Community Col-lege, King Saud University, Saudi Arabia. He hasauthored/co-authored over 30 scientific papers ininternational journals and conference proceedings.His main research interests include socially aware network, Internet ofThings, intelligent systems, big data, recommender systems, and cloudcomputing. He serves as a technical program committee member in severalconferences.