IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 19 August 2024, accepted 29 August 2024, date of publication 5 September 2024, date of current version 13 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3454810

RESEARCH ARTICLE

Task Offloading and Resource Allocation in an RIS-Assisted NOMA-Based Vehicular Edge Computing

ABDUL-BAAKI YAKUBU[1], AHMED H. ABD EL-MALEK[1], (Senior Member, IEEE), MOHAMMED ABO-ZAHHAD[1,2], (Senior Member, IEEE), OSAMU MUTA[3], (Member, IEEE), AND MAHA M. ELSABROUTY[1], (Senior Member, IEEE)

[1]Department of Electronics and Communications Engineering, Egypt-Japan University of Science and Technology, New Borg El-Arab City, Alexandria 21934, Egypt
[2]Department of Electrical Engineering, Faculty of Engineering, Assiut University, Asyut 71515, Egypt
[3]Faculty of Information Science and Electrical Engineering, Kyushu University, Fukuoka 812-0053, Japan

Corresponding author: Abdul-Baaki Yakubu (abdulbaaki.yakubu@ejust.edu.eg)

This work was supported in part by the Seventh Tokyo International Conference on African Development (TICAD7) African Scholarship, in part by Egypt-Japan University of Science and Technology (E-JUST) and the Egyptian Ministry of Higher Education, and in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant 24K07490.

ABSTRACT With the rise of intelligent transportation (ITS), autonomous cars, and on-the-road entertainment and computation, vehicular edge computing (VEC) has become a primary research topic in 6G and beyond communications. On the other hand, reconfigurable intelligent surfaces (RIS) are a major enabling technology that can help in the task offloading domain. This study introduces a novel VEC architecture that incorporates non-orthogonal multiple access (NOMA) and reconfigurable intelligent surfaces (RIS), where vehicles perform binary or partial computation offloading to edge nodes (eNs) for task execution. We construct a vehicle-to-infrastructure (V2I) transmission model by considering vehicular interference and formulating a joint task offloading and resource allocation (JTORA) problem with the goal of reducing total service latency and energy usage. Next, we decompose this problem into task offloading (TO) problem on the vehicle side and resource allocation (RA) problem on the eN side. Specifically, we describe offloading decisions and offloading ratios as a decentralized partially observable Markov decision process (Dec-POMDP). Subsequently, a multi-agent distributed distributional deep deterministic policy gradient (MAD4PG) is proposed to solve the TO problem, where every vehicular agent learns the global optimal policy and obtains individual decisions. Furthermore, a whale optimization algorithm (WOA) is used to optimize the phase shift coefficient of the RIS. Upon receiving offloading ratios and offloading decisions from vehicles, edge nodes utilize the Lagrange multiplier method (LMM) and Karush-Kuhn-Tucker (KKT) conditions to address the RA problem. Finally, we design a simulation model based on real-world vehicular movements. The numerical results demonstrate that, compared to previous algorithms, our proposed approach reduces the overall delay and energy consumption more effectively.

INDEX TERMS Reconfigurable intelligent surface, non-orthogonal multiple access, real-time task offloading, vehicular edge computing, multi-agent deep reinforcement learning.

I. INTRODUCTION

Recently, the combination of vehicular communication and edge computing has sparked revolutionary developments in vehicular networks, fostering several types of innovative applications ranging from real-time traffic management to autonomous driving. This has led to the development of the modern era of intelligent transport systems (ITS) with unprecedented capabilities. However, this requires high data transmission and intensive computation, with

The associate editor coordinating the review of this manuscript and approving it for publication was Olutayo O. Oyerinde.

stringent latency requirements. The traditional cloud-centric computing paradigm faces never-before-seen issues owing to the spread of data-intensive automotive services and the strict latency requirements for future applications. Virtual servers provide distant storage capacity and processing resources to users via cloud computing. Without requiring vehicles to have considerable storage space or processing power, stored data can be retrieved anywhere. Therefore, users can exchange large amounts of data across vehicles. Nonetheless, the exponential rise in the quantity of devices, including vehicles, has resulted in a substantial surge in data generation. These devices send an escalating volume of data for processing in the main cloud, thereby placing additional strain on cloud resources. Furthermore, some small devices find the communication cost with the cloud to be too expensive. Moreover, the exchange of data between automobiles and cloud servers requires large bandwidth. Additionally, a high traffic load increases the energy consumption of different wireless devices, significantly increasing the bandwidth cost. Thus, keeping up with the growing advancements in vehicle applications while satisfying the demands of computation and communication has become increasingly difficult.

Nevertheless, traditional cloud or terminal computing approaches cannot satisfy the various quality of service (QoS) or quality of experience (QoE) requirements of vehicular applications. Edge computing [1] is a desirable solution for enhancing computational capabilities in vehicular environments. Vehicular edge computing (VEC) [2] has emerged as a potentially useful approach for streamlining task execution at the edge of automotive networks. Considerable work has been devoted to developing VEC [3], [4], in which vehicle-to-infrastructure (V2I) communication enables edge nodes (eN)[1] to conduct tasks using data uploaded by vehicles. With the help of these computing models, computational workloads may be effectively offloaded to the edge, improving the overall network performance and lowering latency and energy usage.

In addition, to satisfy the diverse communication requirements of connected vehicles, non-orthogonal multiple access (NOMA) has evolved into a potentially useful technology for enhancing spectral efficiency and facilitating extensive connectivity to maximize resource utilization. It enables several users to transmit and receive concurrently within the same communication bandwidth by employing successive interference cancellation (SIC) and superposition coding [5]. NOMA has been acknowledged as a crucial communication technique to support multi-access edge computing (MEC) for fast data rates and large capacity, better than orthogonal multiple access (OMA). In VEC, NOMA can enhance the reliability and effectiveness of wireless communication between vehicles, roadside infrastructure, and edge-computing nodes [6]. Furthermore, it enhances throughput and facilitates real-time data exchange, cooperative sensing, and collaborative computing of tasks. Despite its potential, a vehicular scene may suffer from obstacles that can hinder or weaken the direct communication between distant vehicles in the cell and edge nodes. Therefore, reconfigurable intelligent surfaces (RIS) have been suggested as a viable method to enhance the wireless network performance [7]. An RIS is composed of numerous reflecting elements that can be adjusted to alter the phase and amplitude of incoming signals. It can significantly increase the coverage and link quality by modifying the phase-shift variables and amplitude-reflection coefficients as necessary [8].

Despite the substantial potential of RIS, several technical challenges must be overcome before they can be seamlessly integrated into vehicular edge computing systems. These obstacles include developing effective RIS design algorithms, optimizing phase-shift control strategies, coordinating RIS operations with vehicular mobility patterns, and integrating RIS with current communication protocols and standards. Additionally, the use of RIS-assisted NOMA-based offloading does not solve the high communication overhead in VEC owing to poor decision-making regarding offloading. As such, informed decisions are essential, and the effectiveness of the combination of RIS and NOMA depends on key metrics such as delay and energy consumption [9].

The significant challenges facing existing research on task offloading and resource optimization are poor real-time decision-making and low decision-making accuracy [10]. Two offloading mechanisms have been explored in previous studies, namely: binary and partial. Earlier studies have employed iterative techniques (such as game theory and heuristic algorithms) [11], [12]. However, real-world edge networks frequently feature time-varying system factors (such as the amount of data, vehicle states, and channel conditions), complexity, and uncertainty. Current methods make greedy decisions at every time frame, disregarding the optimization of the overall performance of the system, which may result in suboptimal decisions. For these reasons, the researchers in [13], [14], and [15] used reinforcement learning (RL) and deep reinforcement learning (DRL) to make offloading decisions for the computational offloading of VEC. However, none of these studies have investigated both resource optimization and offloading in VEC. A few studies have developed a joint optimization framework that combines resource allocation and task offloading [16], [17]. Multi-agent deep reinforcement learning (MADRL) [18] has surfaced as a distributed approach for unmanned area vehicles (UAV) and vehicular applications [19], [20], [21], [22]. Nonetheless, these solutions rely on centralized training and decentralized execution (CTDE) methods. CTDE-based methods result in problems such as *centralized-decentralized mismatch* (CDE) [23] and *multi-agent credit assignment* (MACA) [24]. Therefore, they cannot be directly applied for real-time resource allocation in vehicular networks. In contrast, [25] used a multi-agent distributed distributional deep deterministic policy gradient (MAD4PG) and game theory to address CTDE-based problems. However, their

[1] An eN in this context contains a base station (BS) / roadside unit (RSU), with edge server.

work did not investigate RIS capabilities in the VEC environment.

For the above reasons, this study examines a joint task offloading and resource optimization (JTORA) problem for dynamic and fast-moving vehicles to realize real-time, accurate, and decentralized joint task offloading and resource allocation. The proposed JTORA problem aims to minimize long-term service delays and energy usage. We separated the problem into a task offloading (TO) problem on the vehicular side and a resource allocation (RA) problem on the eN side. Next, MAD4PG, together with a whale optimization algorithm (WOA) termed (MAD4PG-RIS), is proposed to address the TO problem and *RIS phase shift optimization*. In particular, it combines distributional RL and *value decomposition functions* (VDF) [23] to speed up learning and facilitate convergence. Each eN then uses the *Lagrange multiplier method* (LMM) and *Karush-Kuhn-Tucker* (KKT) conditions to address the RA problem after receiving the TO decisions of each vehicle. The main contributions of this study are as follows:

1) We propose an RIS-assisted NOMA-based VEC architecture that deploys multiple RISs to enhance the communication links between vehicles and edge nodes (eNs). Typically, vehicles must increase their transmission power to maintain their signal strength over longer distances, leading to higher energy consumption. Our architecture mitigates this issue using RISs. Vehicles offload tasks to eNs using the same communication bandwidth, with tasks varying in terms of their computational requirements and deadlines. In addition, we enable vehicles to compute tasks locally by combining the strengths of binary and partial offloading.
2) We formulated the JTORA problem to minimize the overall service delay and energy cost by considering binary and partial offloading. The model for V2I transmission incorporates intra and inter-edge interferences using the NOMA principle. This optimization is a long-run mixed-integer nonlinear programming (LR-MINP) problem that is simplified by dividing it into two subproblems: task offloading (TO), including RIS phase-shift optimization on the vehicle side, and resource allocation (RA) on the edge node (eN) side. The TO problem is a decentralized, partially observable Markov decision process (Dec-POMDP) owing to the agile and partially perceived characteristics of vehicles. The RA problem, given the offloading decisions and ratios, is treated as a separate concave programming (SCP) problem, which is straightforward to solve.
3) Given the complex and mixed action space of each vehicular agent in Dec-POMDP, we utilize a multi-agent distributed distributional deep deterministic policy gradient (MAD4PG). To improve the convergence, we used Kullback-Leibler (KL) divergence regularization to address the convergence bias between the distributions of the behavior and target policies. This approach uses distributional functions to manage continuous state-action returns and reduces imprecise Q-value estimations. In addition, to effectively address MACA and CDM concerns, the algorithm employs value decomposition functions (VFD) to divide the centralized critic network into a proportional total of distinct critic networks, automatically determining each vehicular agent's local Q-value function.
4) We optimized the RIS phase shift parameters by employing the *whale optimization algorithm* (WOA). Hence, MAD4PG-RIS jointly makes offloading decisions and offloading ratios and optimizes the RIS phase-shift parameters. Subsequently, we designed a simulation using real-world vehicular movement.

The remainder of this paper is organized as follows. Section II reviews the related work. Section III describes the proposed system model. Section IV describes the JTORA problem. In Section V, the MAD4PG-RIS and WOA are proposed, and we derive the optimal solution for the RA problem. In Section VI, numerical results are presented. Finally, the conclusions are presented in Section VII.

Lowercase and uppercase boldface letters indicate vectors and matrices, respectively. $\mathbb{C}^{n\times m}$ represents the complex domain with space $n \times m$ and $\overline{f(x)}$ denotes the lower bound of $f(x)$. The operator $[\cdot]^T$ represents the transpose, and $|\cdot|$ and $\|\cdot\|$ denote the absolute value and Euclidean norm, respectively. Table 1 summarizes the key symbols and notations used in this study.

II. RELATED WORK

Over the years, significant efforts have been dedicated to task offloading and resource optimization in VEC. Mainstream research investigations are often concerned with minimizing latency and energy consumption. Various strategies have been implemented to address challenges related to optimizing decisions and resources for task offloading. Some of these studies used heuristics, meta-heuristics, greedy algorithms, game theory, reinforcement learning, etc. Nguyen et al. [12] used a meta-heuristic algorithm to maximize the acceptance ratio and optimize task execution latency for parked vehicles (PVs). Their approach suggests the use of Kubernetes-based container orchestration, which intends to effectively offload online processing to parked vehicles during peak traffic hours. Raza et al. [26] employed a game-theoretical method to reduce the delay and energy in VEC. Manogaran et al. [31] formulated blockchain-assisted offloading to maximize data availability. They proposed Naïve Bayes classification to determine offloading instances to avoid unnecessary backlogs. However, these iterative and probabilistic methods cannot be used for real-time task offloading decisions in VEC owing to the dynamic and frequent time-varying factors in vehicular networks. Therefore, researchers have explored reinforcement learning methods to determine the optimal offloading policies in vehicular environments.

TABLE 1. Summary of primary symbols and notations.

Symbol	Description
$\mathbb{E}$	Set of eNs, $\mathfrak{e}_e \in \mathbb{E}$ with indexed $e \in \{1, 2, \ldots, E\}$
$\mathbb{V}$	Set of vehicles, $\nu_v \in \mathbb{V}$ with indexed $v \in \{1, 2, \ldots, \mathcal{V}\}$
$\mathbb{R}$	Set of RISs
$\mathbb{T}$	Set of time slots, $t \in \mathbb{T} = \{1, 2, \ldots, \mathcal{T}\}$
K	Set of requested tasks by vehicles.
$\boldsymbol{\Theta}$	RIS diagonal phase shift matrix
k_v^t	Requested tasks by vehicle v at time t.
T_{k_v}	Deadline of task k_v^t
d_k	Tasks k_v^t data size.
C	CPU cycles for processing a single bit of task data size.
$\boldsymbol{\hbar}_{v,e}^t$	Channel gain of vehicle v communicating with eN e
$d_{v,e}$	Distance between vehicle v and eN e
$d_{v,r}$	Distance between vehicle v and RIS r
r_e	Communication range of eN e
b	Communication bandwidth
p_e^t	The highest V2I communication power at eN e
$p_{v,e}^t$	Vehicle transmission power at time t allocated by eN e
p_v	Power consumed by vehicles for local computation
$p_e^{t,exec}$	Power allocated by eN e for computing task at the edge
F_e	Maximum computation capability of eN e
$f_{v,e}^t$	Allocated computation capability by eN e for computing vehicle's v task at time t.
f_v	Computing capability of a vehicle v for local processing
$q_{v,e}^t$	Binary decision for task offloading
$\alpha_{v,l}, \alpha_{v,e}$	Ratio of vehicular task computed locally and eN e, respectively
$\mathbf{K}_v^t$	Set of uploaded tasks of vehicles within eN e at time t
$\mathbf{K}_{v_e}^t$	Set of offloaded tasks in eN e at time t
$\mathbb{V}_e^t$	Set of vehicles within eN e at time t

An in-depth analysis of RL and DRL methods for task offloading in VEC and vehicular cloudlets was presented in [14]. Their work briefly highlights algorithms such as deep Q networks (DQN), deep deterministic policy gradient (DDPG), proximal policy optimization (PPO), and many others used in VEC. Liu et al. [16] formulated a semi-Markov process and proposed two RL methods (Q-learning and DRL) to optimize computation offloading policies and resource allocation. Their objective was to optimize the total utility of the VEC network. Zhan et al. [13] employed a PPO algorithm to schedule offloads in vehicular networks. They carefully designed an overall computation cost function that sought to minimize the delay and energy for traveling vehicle terminals (VT), intending to offload their task for computation. In another study, the authors of [15] developed a PPO strategy that makes offloading decisions and uses alternate radio access technologies to reduce latency. However, none of these methods employ NOMA to improve the spectral efficiency of the VEC architecture.

Patel et al. [6] examined NOMA-based vehicular network communication capabilities, and the results revealed that NOMA performed approximately 20% better than the traditional OMA. Du et al. [32] utilized a heuristic algorithm for resource optimization and offloading in NOMA-enabled VEC. They aimed to decrease processing delay by performing partial task offloading. Zhou et al. [33] presented an improved WOA (IWOA) for multi-step offloading considering security and NOMA in ultra-dense IoT networks. Their model obtained a binary association decision for IoT mobile devices (IMDs) and performed partial offloading to a nearby base station (BS). In another study [34], they proposed a further improved WOA (FIWOA) and genetic WOA (GWOA) to mitigate interference in ultra-dense networks through joint multi-step secure offloading and resource allocation utilizing NOMA. Zhu et al. [35] introduced an efficient power allocation strategy designed to optimize long-term power usage and latency while considering random task arrival and channel variation. Liu et al. [36] suggested an alternative directional algorithm based on NOMA to allocate power to autonomous vehicles. Nevertheless, these studies did not investigate the use of RIS to achieve spectral efficiency.

Agrawal et al. [37] evaluated the effectiveness of RIS-aided UAV-enhanced vehicular communications. Their work analyzed the performance by considering multiple nonidentical interferers with regard to the bit error rate (BER), coverage probability, and goodput using Monte Carlo simulations. In addition, Javed et al. [38] presented the advantages of using an RIS in cybertwin VEC networks, including Doppler lessening, overcoming blockages, and improving localization. They demonstrated that RIS could increase the proportion of direct visibility in cars to 25% in an area with traffic density reaching 60 cars/km. In another study, Zhu et al. [28] illustrated a technique for dynamic scheduling in RIS-aided vehicular networks with the goal of maximizing the computing throughput. They formulated a mathematical model for task offloading policies and task scheduling considering vehicle movement taken into account.

Furthermore, researchers explored the synergy between RIS and NOMA in VEC. Li et al. [30] investigated an RIS-assisted cell with several user scenarios in a NOMA MEC system. They jointly optimize the phase shift, transmit data size, and transmission rate to minimize total energy usage. Xie et al. [29] adopted a heuristic algorithm to jointly optimize offloading decisions and computation resources while using an iterative method to optimize the RIS phase-shift coefficients. In contrast, Mirza et al. [27] used an advantage actor-critic (A2C) technique to determine offloading decisions in zero-energy RIS-aided vehicular networks, which aimed to reduce the overall delay and energy costs. However, the DRL of current studies, such as iterative approaches, faces the challenge of poor real-time decision-making because their foundation lies in centralized scheduling, which entails a significant communication burden and complex decision-making in a VEC.

Previous studies have suggested the use of MADRL as a distributed approach for VEC task offloading. For instance, Ju et al. [22] employed a double Q-learning algorithm to obtain secure offloading decisions aimed at minimizing the processing delay. They collaboratively optimized the spectrum selection and transmission power of the VEC networks. Liao et al. [39] presented a blockchain and semi-distributed learning with DRL used for training local data whereas federated averaging was employed for collaboration in a

TABLE 2. Comparing our work with related studies and emphasizing the main distinctions and contributions.

Reference	Offloading mechanisms		Radio access	RIS		NOMA	Objectives		Solution type
	Binary	partial		Usage	Phase opt.		Energy	Delay	
Raza et al. [26]	✓	✗	V2X	✗	✗	✗	✓	✓	Game theoretic
Liu et al. [16]	✓	✓	V2X	✗	✗	✗	✗	✓	Q-learning and DRL
Zhan et al. [13]	✓	✗	V2I	✗	✗	✗	✓	✓	PPO based RL
Mirza, al. [27]	✓	✓	V2X	✓	✗	✗	✓	✓	A2C based RL
Zhu et al. [28]	✗	✓	V2I	✓	✗	✗	✗	✓	Dynamic scheduling algorithm
Xie et al. [29]	✓	✗	V2I	✓	✓	✓	✗	✓	Heuristics and Iterative method
Li et al. [30]	✗	✓	-	✓	✓	✓	✓	✗	Block coordinate descent algorithm
Ju et al. [22]	✓	✗	V2V	✗	✗	✗	✗	✓	Double DQN
Zhang et al. [20]	✗	✓	V2V	✗	✗	✗	✗	✓	MADDPG
Kumar et al. [17]	✓	✗	V2X	✗	✗	✗	✓	✓	L-MADDPG
Alam et al. [21]	✓	✗	V2X	✗	✗	✗	✓	✓	MADRLHA
Zhang et al. [10]	✗	✓	-	✗	✗	✗	✓	✓	IS-DMDRL
Xu et al. [25]	✓	✗	V2I	✗	✗	✓	✗	✓	MAD4PG
This work	✓	✓	V2I	✓	✓	✓	✓	✓	MAD4PG and WOA

secure computation offloading setting. Zhang et al. [20] proposed the use of MADDPG as a resource-allocation technique to reduce the cost of offloading vehicular tasks while maintaining stringent delay limits. Kumar et al. [17] developed a Lyapunov MADDPG (L-MADDPG) to jointly optimize task distribution and resources to reduce delay and energy. In another study, Alam et al. [21] formulated a bipartite graph maximum matching problem and proposed a multi-agent DRL-based Hungarian algorithm (MADRLHA) to address the offloading of dynamic tasks. Nevertheless, this approach results in issues, such as MACA and CDM. Zhang et al. [10] proposed an improved soft actor-critic (SAC)-based decentralized MADRL (IS-DMDRL) algorithm to obtain decentralized task offloading policies in Industrial Internet of Things (IIOT) devices. Similarly, Xu et al. [25] utilized MAD4PG and game theory to optimize offloading decisions to achieve a high service ratio for VEC networks.

Therefore, it is necessary to address the complex decision-making dilemma in VEC while including cutting-edge technologies such as NOMA and RIS for spectral efficiency and reliable communication links. Therefore, none of the existing solutions can be employed directly for diverse resource distributions and complex real-time task assignments in VEC.

In this study, we employed NOMA and RIS for joint task offloading and resource optimization. As highlighted in Table 2, our proposed solution differs from the existing literature in that it combines several advantages: it considers both binary and partial offloading in V2I communication and RIS usage and optimizes the phases of the RIS for spectral efficiency. The synergistic benefits of these techniques for task offloading and resource optimization are demonstrated by comparing the results presented in Section VI. It is worth mentioning that, while we focus on optimizing communication efficiency, energy consumption, and task offloading in RIS-assisted NOMA-based VEC architectures, secure offloading and blockchain technologies can be integrated by employing the methods in [31], [33], [34], and MADRL converges to [39]. However, this aspect is beyond the scope of the present study.

III. SYSTEM MODEL

We propose an RIS-assisted NOMA-based architecture for joint task offloading and resource allocation in VEC. As illustrated in Fig. 1, infrastructures like roadside units (RSUs) or base stations (BSs) are installed along roads, which serve as eNs for vehicles to communicate using V2I communication. Vehicles randomly receive varying tasks, which are partially offloaded to the near eNs. The power control adopted in this study is a base-station-centric approach. The edge node centrally determines the power levels of all vehicles based on their QoS requirements, channel conditions, and the overall network state. In vehicular communication networks, distant vehicles often experience significant path loss, leading to weak channels between vehicles and edge nodes. To compensate for this path loss and ensure reliable communication, distant vehicles typically need to increase their transmission power. A higher transmission power can reduce the transmission delay by improving signal strength and reliability. A higher transmission power also significantly increases the energy consumption, which is undesirable, especially for battery-powered vehicles. Therefore, we deployed an RIS consisting of passive elements that do not require a power-hungry transmitter. This passive nature significantly reduces the overall energy consumption compared with active relays or higher transmission power from vehicles. Vehicles transmit through the same bandwidth and the signals are decoded using the NOMA principle.

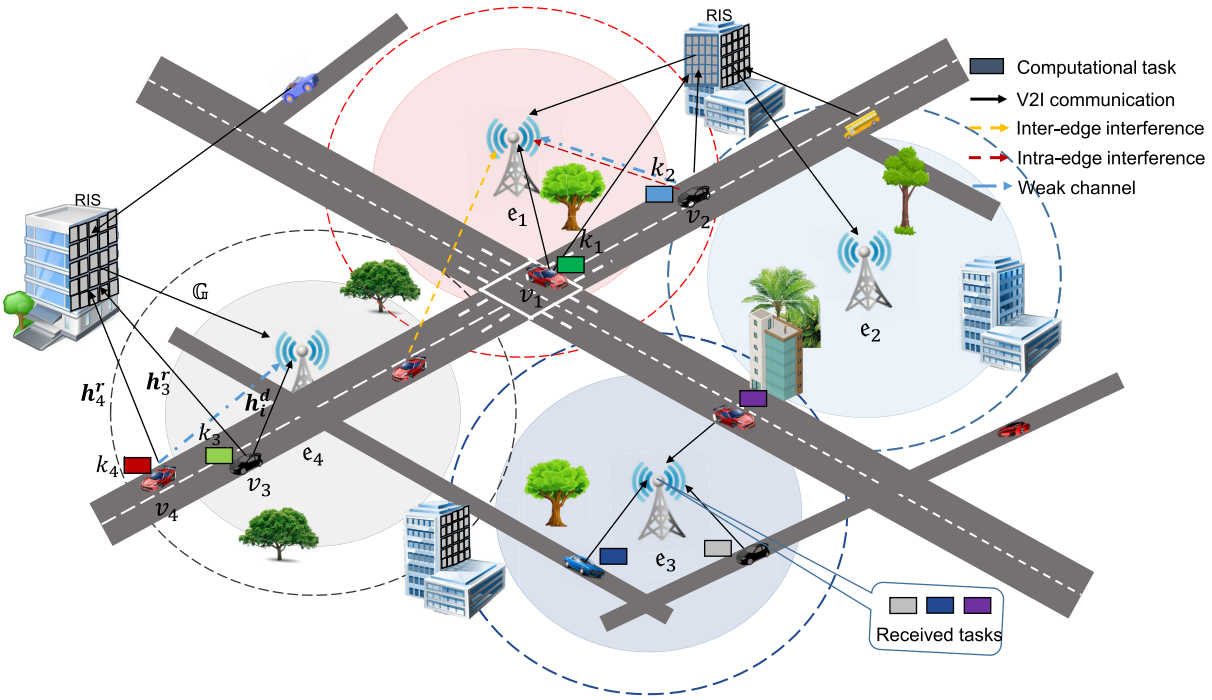


FIGURE 1. RIS-assisted NOMA-based VEC architecture.

The eNs are represented as $\mathfrak{e}_e \in \mathbb{E}$ with the index $e \in \{1, 2, \ldots, E\}$, which contains several computing units (CPU chips) to enhance the computational tasks offloaded by mobile vehicles. $\mathfrak{e}_e = (p_e, f_e, r_e, \ell_e)$ describes eN. The highest V2I communication power is denoted by p_e, $f_e \in [0, F_e]$ is the computational capacity, r_e is the eN coverage, and ℓ_e represents the location of eN. The vehicles are represented by the set $v_v \in \mathbb{V}$ with index $v \in \{1, 2, \ldots, \mathcal{V}\}$. We consider multiple RIS with M passive elements, where the RIS is represented as $\mathfrak{r}_r \in \mathbb{R}$ with index $r \in \{1, 2, \ldots, E\}$. Where each RIS is associated with a specific eN. Consequently, each RIS is dedicated to serving a paired edge node. The RIS diagonal phase-shift matrix is defined as [40] $\boldsymbol{\Theta} = diag\left(e^{j\theta_1}, e^{j\theta_2}, e^{j\theta_3}, \ldots, e^{j\theta_M}\right) \in \mathbb{C}^{M\times M}$, where each θ_M represents a phase shift within the interval $[0, 2\pi)$. At a particular distance from the edges, indicated by the small coverage within edge nodes, the signal strength starts to fade away, which leads to an increase in power to reach an edge node. Hence, the RIS offers a better communication path to eN. The task requested by vehicles at discrete time $t \in \mathbb{T} = \{1, 2, \ldots, \mathcal{T}\}$ is represented by $k_v^t \in K$, which has data size d_k, CPU cycles for processing C, and deadline T_{k_v}.

Furthermore, we provide the communication scenario of vehicles, where v refers to a single antenna vehicle and e refers to the edge node (eN) equipped with N antennas. In Fig. 1, v_1 and v_2 stochastically encounter tasks k_1 and k_2 respectively, which need to be processed. They upload through a direct link and the RIS link to eN e_1, which provides computational resources. Similarly, v_3 and v_4 upload to eN e_4 using V2I communication. However, considering the impact of distance on channel conditions, which is primarily attributable to the path loss of the vehicle, v_4 is assumed to have poor channel conditions for eN e_4 represented as a weak channel. RIS offers another link to communicate with eN e_4. Both v_3 and v_4 transmit through the same communication bandwidth. Subsequently, e_4 used SIC to decode the signal of each vehicle.

A. TASK OFFLOADING MODEL

Let $\mathbf{h}_v^d \in \mathbb{C}^{N\times 1}$, $\mathbf{h}_v^r \in \mathbb{C}^{M\times 1}$, and $\mathbb{G} \in \mathbb{C}^{N\times M}$ denote the channel gains from vehicle v to eN, from v to RIS, and RIS to eN e, respectively. The channel gains are

$$\mathbf{h}_{v,e}^d = \left[\frac{\eta_1}{d_{v,e}^{\gamma/2}} \quad \frac{\eta_2}{d_{v,e}^{\gamma/2}} \quad \cdots \quad \frac{\eta_N}{d_{v,e}^{\gamma/2}}\right]^T,$$
$$\mathbf{h}_v^r = \left[\frac{\eta_1}{d_{v,r}^{\gamma/2}} \quad \frac{\eta_2}{d_{v,r}^{\gamma/2}} \quad \cdots \quad \frac{\eta_M}{d_{v,r}^{\gamma/2}}\right]^T, \tag{1}$$

where $d_{v,e}$ represents the distance from vehicle v to eN e, $d_{v,r}$ is the distance from vehicle v to RIS r, η_n signifies Rayleigh-distributed small-scale fading, that is, $\eta_n \sim \mathcal{CN}(0, 1)$, for $n \in \{1, 2, \ldots, N\}$ and γ represents the path loss exponent.

The received signal at the edge node from vehicle v at time t can be expressed as [40]

$$\mathbf{y}_e^t = \sum \left(\mathbb{G}\boldsymbol{\Theta}\mathbf{h}_v^r + \mathbf{h}_v^d\right) w_v x_{v,e}^t + \mathbf{n}_e, \tag{2}$$

where w_v is the transmit scalar and $|x_{v,e}^t| = 1$ is the transmit symbol of vehicle v. $\mathbf{n}_e$ represents Additive White Gaussian

noise (AWGN).

$$\hbar_{v,e}^{t} = \mathbb{G}\Theta\mathbf{h}_{v}^{r} + \mathbf{h}_{v,e}^{d} \tag{3}$$

The signal received at eN can subsequently be described in relation to the channel $\hbar_{v,e}^{t}$ and the transmission power of the vehicles $p_{v,e}^{t}$ at time t [41] as

$$\mathbf{y}_{e}^{t} = \sum_{\forall v \in \mathbb{V}_{e}^{t}} \sqrt{p_{v,e}^{t}} x_{v,e}^{t} \hbar_{v,e}^{t} + \sum_{\forall v' \in \mathbb{V}_{e'}^{t}} \sum_{\forall e' \in \mathbb{E}} \sqrt{p_{v',e'}^{t}} x_{v',e'}^{t} \hbar_{v',e}^{t} + \mathbf{n}_{e}, \tag{4}$$

where $x_{v,e}^{t}$ is the vehicle's intended message and $x_{v',e'}^{t}$ represents the interfering message.

B. COMMUNICATION MODEL

We modeled the V2I transmission and intra and inter-edge interference using the NOMA principle for downlink and uplink communication. For an uplink NOMA system, vehicle channel conditions differ at a given time. Both sides are expected to know the channel state information (CSI). To control the power, eN allocates the transmit power based on CSI, including information regarding path loss, fading, and other channel conditions that affect the signal strength. Near vehicles have better channel conditions than far vehicles because of the path loss. According to the decoding of uplink NOMA, the edge node first decodes the near vehicle signals by treating the distant vehicle signals as interference. Then, through SIC, the edge node cancels out the near vehicle signal to decode the signals of distant vehicles [5]. The channel gains of an arbitrary number of vehicles V_e served by eN e are sorted as follows.

$$\begin{aligned} &\|\mathbb{G}\Theta\mathbf{h}_{v_1}^{r} + \mathbf{h}_{v_1,e}^{d}|^2 \geq \cdots \geq \|\mathbb{G}\Theta\mathbf{h}_{V_e}^{r} + \mathbf{h}_{V_e,e}^{d}\|^2, \\ &\therefore \|\hbar_{v_1,e}^{t}\|^2 \geq \|\hbar_{v_1,e}^{t}\|^2 \geq \cdots \geq \|\hbar_{V_e,e}^{t}\|^2. \end{aligned} \tag{5}$$

Following the effective cancellation of signals for vehicles. The interference can be modeled as

$$\chi_{intra} = \sum_{\forall v' \in \mathbb{V}_{\hbar_{v,e}}} \|\hbar_{v',e}^{t}\|^2 p_{v',e}^{t}, \tag{6}$$

$$\chi_{inter} = \sum_{\forall e' \in \mathbb{E}} \sum_{\forall v' \in \mathbb{V}_{e'}^{t}} \|\hbar_{v',e}^{t}\|^2 p_{v',e'}^{t}, \tag{7}$$

where $p_{v',e}^{t}$ and $p_{v',e'}^{t}$ are the transmission powers of the interfering vehicle v' within the same eN e and the interfering eN e', respectively. $\|\boldsymbol{h}_{v',e}^{t}\|^2$ denotes the interfering vehicle's v' channel coefficient.

The signal-to-interference-plus-noise ratio (SINR) at time t is formulated as [25]

$$\varphi_{v,e}^{t} = \frac{\|\hbar_{v,e}^{t}\|^2 p_{v,e}^{t}}{\chi_{intra} + \chi_{inter} + \sigma^2}, \tag{8}$$

where $\|\hbar_{v,e}^{t}\|^2$ and $\|\hbar_{v',e}^{t}\|^2$ are the channel coefficients of the desired and interference vehicles, respectively; σ^2 is the noise power; and $p_{v',e}^{t}$ is the interfering vehicle transmission power. Thus, the transmission rate can be obtained as:

$$\upsilon_{v,e}^{t} = b \log_2(1 + \varphi_{v,e}^{t}), \tag{9}$$

where b in Hz is the bandwidth of the V2I communication.

C. DELAY COST MODEL

1) LOCAL COMPUTATION

Given that vehicles have computing capabilities, certain tasks are computed either partially or fully on a local processor. When vehicle v executes its task locally, the local computing delay is given by [42]:

$$\mathcal{D}_{v}^{t,loc} = \frac{\alpha_{v,l} C d_k}{f_v}, \tag{10}$$

where $\alpha_{v,l} \in [0, 1]$ is the ratio of the task computed at vehicle v, d_k is the size of task k_v^t, C is the CPU cycle required to process a bit of data, and f_v is the computing capability of the vehicle, which is determined by the frequency of the CPU cycles.

2) EDGE COMPUTATION

Given the limited processing power of vehicle v, every vehicle v can partially upload its task to an eN e. The transmission delay of this task is:

$$\mathcal{D}_{v,e}^{t,trans} = \frac{\alpha_{v,e} d_k}{\upsilon_{v,e}^{t}}, \tag{11}$$

where $\alpha_{v,e} = \left(1 - \alpha_{v,l}\right)$ denotes the task ratio to be executed at eN. Following the receipt of the task at eN e, the execution delay at eN e is expressed as:

$$\mathcal{D}_{e}^{t,exec} = \frac{\alpha_{v,e} C d_k}{f_{v,e}^{t}}, \tag{12}$$

where $f_{v,e}^{t}$ is the allocated computational capability of eN for vehicle v at time t. The delay associated with downloading can be ignored because the computed results are relatively small [43]. Hence, the service delay or overall delay for finishing task k_v^t is:

$$\mathcal{D}_{v,e}^{t} = \max\left(\mathcal{D}_{v}^{t,loc}, \mathcal{D}_{v,e}^{t,trans} + \mathcal{D}_{e}^{t,exec}\right). \tag{13}$$

Task k_v^t of the vehicles should be completed before the vehicle moves out of sight of the RIS or eN. Therefore, we denote the task deadline by T_{k_v}. When a task is serviced before its deadline, it is said to be successfully serviced. Hence, the service ratio can be determined as follows:

$$\Gamma_{v}^{t} = \frac{\sum_{\forall k_v^t \in \mathbf{K}_v^t} \mathbb{I}\left(\mathcal{D}_{v,e}^{t} \leq T_{k_v}\right)}{|K_v^t|}, \tag{14}$$

where $|K_v^t|$ is the number of task vehicles requested in an eN and $\mathbb{I}\left(\mathcal{D}_{v,e}^{t} \leq T_{k_v}\right)$ is an indicator function that is 1 if $\mathcal{D}_{v,e}^{t} \leq T_{k_v}$ and 0 otherwise.

D. ENERGY COST MODEL

1) LOCAL COMPUTATION

The energy required by vehicle v to locally compute the task is given by [10]

$$\mathcal{E}_v^{t,loc} = p_v^{t,loc} C \alpha_{v,l}, \tag{15}$$

where $p_v^{t,loc}$ denotes local computing power.

2) COMPUTATION AT THE EDGE

Vehicle v consumes power $p_{v,e}^t$ during the offloading of the task to eN e. The associated energy required by vehicle v to upload a task to eN e is:

$$\mathcal{E}_{v,e}^{t,trans} = p_{v,e}^t \mathcal{D}_v^{trans} = \frac{p_{v,e}^t \alpha_{v,e} d_k}{\upsilon_{v,e}^t}. \tag{16}$$

In addition, the energy consumed during task execution at eN is given as

$$\mathcal{E}_e^{t,exec} = p_e^{t,exec} C \alpha_{v,e}, \tag{17}$$

where $p_e^{t,exec}$ is the power used by eN to process the task. Hence, the total energy cost of vehicle v is

$$\mathcal{E}_{v,e}^t = \mathcal{E}_v^{t,loc} + \mathcal{E}_{v,e}^{t,trans} + \mathcal{E}_e^{t,exec}. \tag{18}$$

IV. PROBLEM FORMULATION

In this section, we formulate the JTORA problem to minimize the overall computational cost of vehicles. To reduce its complexity, we divide it into a TO problem on the vehicle side and an RA problem on the eN side.

A. JTORA PROBLEM

According to (8), (9), (11), and (16), vehicle v can increase its transmission power $p_v^{t,trans}$ to reduce the transmission delay $\mathcal{D}_{v,e}^{t,trans}$. However, this increase leads to an increase in uploading energy consumption $\mathcal{E}_{v,e}^{t,trans}$. Therefore, a trade-off relationship exists between the energy and delay. We determined the cost of each vehicle as the weighted sum of its delay and the energy expenses required for each vehicle device to complete its task at time t.

$$\Phi = \lambda \mathcal{D}_{v,e}^t + (1-\lambda)\,\mathcal{E}_{v,e}^t + \frac{g_t(\mathcal{D}_{v,e}^t - \mathcal{D}_{TH})}{\mathcal{D}_{TH}} + \frac{g_e\left(\mathcal{E}_{v,e}^t - \mathcal{E}_{TH}\right)}{\mathcal{E}_{TH}}, \tag{19}$$

where λ represents the weight, $\mathcal{D}_{TH}$ denotes the delay threshold, $\mathcal{E}_{TH}$ denotes the energy consumption threshold. The variables $g_t \in \{0,1\}$ and $g_e \in \{0,1\}$ are the penalty factors for the delay and energy, respectively. When penalty constraints are not applied,g_t and g_e are both zero [44]. The sum of all the vehicle costs is the overall system cost at time t. The total cost is then calculated over the time window $t_1 < t < T$.

$$\mathcal{P}\mathbf{1}: \min_{\alpha_v,\theta,f_{v,e}^t,q_{v,e}^t,p_{v,e}^t} \sum_{t_1}^{T}\sum_{v_1}^{V_e} \Phi$$
$$\text{s.t.}: \; 1-\alpha_{v,l} = \alpha_{v,e}, \tag{20a}$$
$$\sum_{\forall v \in \mathbb{V}_e^t} p_{v,e}^t \le p_e, \tag{20b}$$
$$\sum_{\forall k_v^t \in \mathbf{K}_{ve}^t} f_{v,e}^t \le F_e, \tag{20c}$$
$$\sum_{\forall e \in \mathbb{E}} q_{v,e}^t = 1, \tag{20d}$$
$$0 \le \theta_n(t) \le 2\pi, \quad 1 \le n \le N. \tag{20e}$$

Constraint (20a) ensures that the total offloading ratios from the vehicle to the edge nodes do not exceed 1. Constraint (20b) ensures that the total transmission power does not exceed the maximum V2I communication power. (20c) stipulates that the amount of computing resources assigned does not surpass the computational capacity of the edge nodes. Constraint (20d) guarantees that each task is limited to associating and offloading to only a single eN when task offloading decision $q_{v,e}^t$ is made. (20e) ensures that the angle of the n-th RIS reflector remains between 0 and 2π.

B. TASK OFFLOADING AT VEHICLE-SIDE

Notably, given each time slot, $\mathcal{P}\mathbf{1}$ can be decomposed into two sub-problems. The first subproblem concerns the decisions to offload and the offloading ratio of vehicles, which can be presented as:

$$\mathcal{P}\mathbf{2}: \min_{q_{v,e}^t,\alpha_v,\theta} \sum_{t_1}^{T}\sum_{v_1}^{V_e} \Phi,$$
$$\text{s.t.}: \; 1-\alpha_{v,l} = \alpha_{v,e}, \tag{21a}$$
$$1\sum_{\forall e \in \mathbb{E}} q_{v,e}^t = 1, \tag{21b}$$
$$0 \le \theta_n(t) \le 2\pi, \quad 1 \le n \le N. \tag{21c}$$

Realistic vehicular environments are dynamic and partially observable. Hence, $\mathcal{P}\mathbf{2}$ on the vehicle side can be regarded as a decentralized Dec-POMDP problem. Owing to the limited communication among vehicles, every vehicle (i.e., agent) can acquire local observations and discover its optimal decentralized policy by continuously engaging with its surroundings. (V, S_v, O_v, A_v, R_v) is used to characterize Dec-POMDP.

1) V is the set of vehicular agents and the number of vehicular agents is represented by V_e.
2) *Global system state* (S_v): At time t, the global system of the RIS-assisted VEC environment is $s^t \in S_v$.
3) O_v is the shared observation of vehicular agents. At time t, each agent v collects the local observations $o_v^t \in O_v$ from the environment. O_{v_n} represents the local observation space of agent v. o_v^t comprises the vehicle's intrinsic state parameters derived from its system configuration. It includes task data size, task deadline, residual battery level, and CPU cycles. The perceived environmental parameters include the distance to eN, computational resources of eN, and channel conditions.

Hence, the local observation can be simplified as

$$o_v^t = \{v, t, \hbar_{v,e}^t, d_k, f_v, T_{k_v}\}, \tag{22}$$

where v is a vehicular index, t is time slot, $\hbar_{v,e}^t$ is the channel conditions of vehicle v to eN, e, d_k represent the data size, and f_v and T_{k_v} are the required commutation resources and deadline of task k_v respectively.

4) *Action space* (A_v): Each agent chooses action $a_v^t \in A_v$ after obtaining its local observation o_v^t to make offloading choices according to the policy. The action a_v^t of vehicular agents is the offloading decision of tasks and the task division ratio. It is donated by

$$a_v^t = \{q_{v,e}, \alpha_{v,l}\}, \tag{23}$$

where $q_{v,e} \in \{0, 1\}$. In the binary offloading procedure, when $q_{v,e}$ is 0 and $\alpha_{v,l}$ is 0, the task is computed locally. If a vehicle does not have idle computing resources in a particular time slot, then the decision $q_{v,e}$ is 1, and $\alpha_{v,l}$ is given 1 (i.e., the entire task is sent to eN). For a partial scenario, when vehicle computing resources are insufficient at a given time, the decision $q_{v,e}$ is 1, and then $0 < \alpha_{v,l} < 1$ where $\alpha_{v,l} + \alpha_{v,e} = 1$ are the offloading ratios.

5) *Reward function*: The goal of each vehicle is to lower the overall communication cost by optimizing the offloading ratio and task offloading decisions.

$$r_v^t(a_v|o_v) = -\sum_{v_1}^{V_e} \Phi. \tag{24}$$

An agent receives fewer rewards when the total communication cost is high.

C. RESOURCE ALLOCATION AT eN-SIDE

After the offloading actions of vehicular agents are executed, eN must allocate the transmission power and computational resources to each agent to simultaneously process the offloaded task, seeking to reduce the service delay and energy consumption. The RA problem is formulated as follows.

$$\mathcal{P}\mathbf{3}: \min_{p_{v,e}^t, f_{v,e}^t} \sum_{t_1}^{T} \sum_{v_1}^{V_e} \Phi$$
$$\text{s.t.}: \sum_{\forall v \in \mathbb{V}_e^t} p_{v,e}^t \le p_e, \tag{25a}$$
$$\sum_{\forall k_v^t \in \mathbf{K}_{qe}^t} f_{v,e}^t \le F_e. \tag{25b}$$

Variables $p_{v,e}^t$ and $f_{v,e}^t$ are independent. Constraints (25a) and (25b) can be separated because the variables do not overlap: Therefore, we can divide $\mathcal{P}\mathbf{3}$ into *transmit power* RA (TPRA) and *computation* RA (CRA), which are formulated as follows [25]:

1) TPRA

It is with respect to variable $p_{v,e}^t$ which concerns the transmission power and is modeled as

$$\mathcal{P}\mathbf{4}: \min f(p_{v,e}^t) = \sum_{\forall e \in \mathbb{E}} \sum_{\forall k_v^t \in \mathbf{K}_{ve}^t} \frac{\alpha_{v,e} d_k}{b \log_2\left(1 + \varphi_{v,e}^t\right)},$$
$$\text{s.t.}: \sum_{\forall v \in \mathbb{V}_e^t} p_{v,e}^t \le p_e. \tag{26a}$$

We can further simplify $f(p_{v,e}^t)$ in problem $\mathcal{P}\mathbf{4}$ to relate to only one eN because the variables associated with each eN are independent. Implies,

$$\min f(p_{v,e}^t) = \sum_{\forall k_v^t \in \mathbf{K}_{ve}^t} b \log_2\left(1 + \varphi_{v,e}^t\right)$$
$$\text{s.t.}: \sum_{\forall v \in \mathbb{V}_e^t} p_{v,e}^t \le p_e. \tag{27a}$$

Nevertheless, owing to the interference in $\varphi_{v,e}^t$, the problem is not convex. Additionally, the objective is not concave in $p_{v,e}^t$. The problem can be relaxed using the lower bound of $f(p_{v,e}^t)$, that is, $\overline{f(p_{v,e}^t)}$ formulated as

$$f(p_{v,e}^t) \ge \overline{f(p_{v,e}^t)}$$
$$\overline{f(p_{v,e}^t)} = \sum_{\forall k_v^t \in \mathbf{K}_{ve}^t} b\left(\xi_{v,e}^t + \delta_{v,e}^t \log_2 \varphi_{v,e}^t\right), \tag{28}$$

where $\xi_{v,e}^t$ and $\delta_{v,e}^t$ are constants defined as [45]

$$\xi_{v,e}^t = \log_2\left(1 + \overline{\varphi_{v,e}^t}\right) - \frac{\overline{\varphi_{v,e}^t}}{1 + \overline{\varphi_{v,e}^t}} \log_2 \overline{\varphi_{v,e}^t},$$
$$\delta_{v,e}^t = \frac{\overline{\varphi_{v,e}^t}}{\left(1 + \overline{\varphi_{v,e}^t}\right)}. \tag{29}$$

Appendix presents the complete derivation of (29). It is tight at the lower bound if $\varphi_{v,e}^t = \overline{\varphi_{v,e}^t}$ and we set $\widetilde{p_{v,e}^t} = \log_2 p_{v,e}^t$. Then, the relaxation is expressed as:

$$\mathcal{P}\mathbf{5}: \min \widetilde{f}(p_{v,e}^t) = \sum_{\forall k_v^t \in \mathbf{K}_{ve}^t} b\left(\xi_{v,e}^t + \delta_{v,e}^t \log_2 \widetilde{\varphi_{v,e}^t}\right)$$
$$\text{s.t.}: \sum_{\forall v \in \mathbb{V}_e^t} 2^{\widetilde{p_{v,e}^t}} \le p_e, \tag{30a}$$

where $\log_2 \widetilde{\varphi_{v,e}^t}$ is given by

$$\log_2 \widetilde{\varphi_{v,e}^t}$$
$$= \widetilde{p_{v,e}^t} + \log_2 \|\hbar_{v,e}^t\|^2 - \log_2$$
$$\times \left(\sum_{\forall v' \in \mathbb{V}_{\hbar_{v,e}}^t} \|\hbar_{v',e}^t\|^2 2^{\widetilde{p_{v',e}^t}} + \sum_{\forall e' \in \mathbb{E}} \sum_{\forall v' \in \mathbb{V}_{e'}^t} \|\hbar_{v',e}^t\|^2 2^{\widetilde{p_{v',e'}^t}} + \mathbf{n}_e \right). \tag{31}$$

Thus, $\mathcal{P}\mathbf{5}$ is a convex optimization problem.

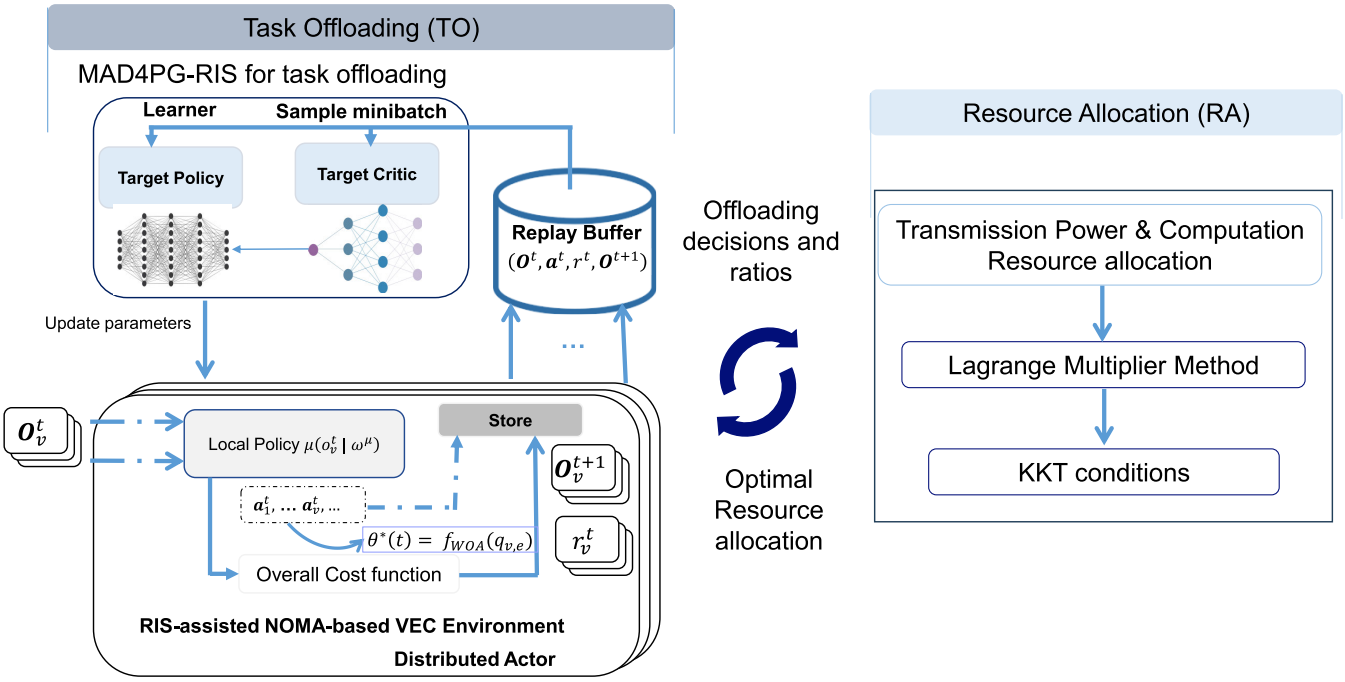


FIGURE 2. MAD4PG-RIS network and resource allocation.

2) CRA

The computational resource allocation problem of every eN is expressed as follows:

$$\mathcal{P}\mathbf{6}: \min g(f_{v,e}^t) = \sum_{\forall k_v^t \in \mathbf{K}_{v_e}^t} \frac{\alpha_{v,1} C d_k}{f_{v,e}^t},$$
$$\text{s.t.}: \sum_{\forall k_v^t \in \mathbf{K}_{v_e}^t} f_{v,e}^t \leq F_e. \tag{32a}$$

Hence, both TPRA and CRA are SCP problems.

V. MULTI-AGENT DEEP REINFORCEMENT LEARNING FOR BINARY AND PARTIAL TASK OFFLOADING

Dec-POMD agents operate in intricate, multidimensional, and diverse action spaces. Hence, the DRL method is used to learn the TO policy. Nevertheless, applying single-agent reinforcement learning to tackle this issue is challenging owing to interference among different vehicles. Also, the TO problem involves a combination of binary and partial offloading within the dynamic RIS-assisted NOMA-based VEC environment, and to address MACA and CDM, we utilized multi-agent distributed distributional reinforcement learning (MAD4PG). The proposed MAD4PG-RIS algorithm learns the decentralized policy of each agent. It jointly optimizes the decisions, offloading ratio, and RIS phase shifts to minimize the overall computational cost.

A. MAD4PG FOR TASK OFFLOADING IN AN RIS-ASSISTED VEC ENVIRONMENT

Fig. 2 illustrates MAD4PG-RIS task offloading and resource allocation in an RIS-assisted VEC environment. MAD4PG is an extension of D4PG [46]. It is a distributed off-policy method that facilitates multi-agent collaboration in learning continuous action-space strategies. It uses a Value Function Distribution (VFD) to improve the critic network and prioritizes the replay experience to speed up learning and facilitate convergence. The illustrated MAD4PG-RIS and resource allocation is similar to the one in [25] which uses MAD4PG. However, our proposed system modifies the MAD4PG to mitigate convergence bias using KL divergence. In addition, our network incorporates RIS and optimizes its phases using WOA.

In the MAD4PG-RIS framework, each agent is designed to maximize its expected return, defined as $R_v(t) = \sum_{j\geq 0} \zeta^j r_v^{t+j}$, where ζ represents the discount factor. Initially, the learner randomly sets the parameters for the local critic $\omega^{\mathcal{Q}}$ and policy network ω^μ. These initial parameters are then synchronized with the target critic network $\omega^{\mathcal{Q}'}$ and the target policy network $\omega^{\mu'}$.

$$\omega^{\mu'} \leftarrow \omega^\mu, \quad \omega^{\mathcal{Q}'} \leftarrow \omega^{\mathcal{Q}}. \tag{33}$$

Furthermore, global replay buffer $\mathcal{B}$ is initialized as part of the setup process. Algorithm 1 provides a full description of the process.

Algorithm 1 MAD4PG-RIS for Task Offloading

Initialize: network parameters ω, replay buffer $\mathcal{B}$;
Start L distributed actors, each with replicated networks $(\omega^{\mu}, \omega^{Q})$;
for *iter* $= 1, 2, \ldots,$ *max iter* **do**
  Initialize: Observation space o_v^t;
  for $t = 1, 2, \ldots, T$ **do**
    for $v = 1, 2, \ldots, V$ **do**
      Draw $\mathcal{M}$ random samples from $\mathcal{B}$ for transitions of length $\mathcal{N}$;
      Create target network distributions;
      Calculate network loss for both policy and critic;
      Update the local critic and policy networks;
    if $t \mod t_{tgt} = 0$ **then**
      Update the target networks;
    if $t \mod t_{act} = 0$ **then**
      Synchronize network weights;

The L distributed actors

while *learner not completed* **do**
  Initialize: random actions for exploration;
  Actor receives initial observation o_1;
  for $t = 1, 2, \ldots, T$ **do**
    for $v = 1, 2, \ldots, V$ **do**
      Retrieve local observation o_v^t;
      Determine action space a_v^t;
    for $a_v^t = (q_v^t, \alpha_{v,l})$ **do**
      Compute optimal RIS phase shift $\theta^*(t) = f_{WOA}\left(q_{v,e}\right)$;
    Obtain reward r_v^t and next observation o_v^{t+1};
    Store $\left(o_v^t, a_v^t, r_v^t, o_v^{t+1}\right)$ in replay buffer $\mathcal{B}$;

The system consists of L distributed actors that interact with the environment concurrently by performing actions and collecting experience. The local policy parameters for the lth actor are derived from the local network of the learner ω_l^{μ}. Each distributed actor makes offloading decisions and determines the offloading ratio for vehicle v at time t based on the locally observed system states, as follows:

$$a_v^t = \mu\left(o_v^t|\omega_j^{\mu}\right) + \epsilon\Delta_t, \tag{34}$$

where Δ_t represents exploration noise, which increases the diversity of vehicular actions, and ϵ is a constant that governs the extent of exploration. The actions a_v^t executed by vehicles at time t occur within the RIS-assisted NOMA-based VEC environment. These actions lead to the generation of an offloading decision that is input into the whale optimization function, producing an optimal phase shift at a specific time t. Rewards r_v^t for vehicles at time t are obtained according to (24). Observations o_v^t, vehicle actions a_v^t, rewards r_v^t, and subsequent observation states o_v^{t+1} are retained in the replay buffer. This procedure is repeated iteratively until the learning process is complete.

Distributed agents interact with each other through learners to improve performance and learning efficiency. This interaction involves sharing experiences and policies among the agents. The experiences collected by each agent are stored in global replay buffer $\mathcal{B}$ which is accessible to all agents. This shared replay buffer enables agents to learn from each other's experiences, thereby enhancing the diversity and quality of training data. In addition, the policy parameters ω^{μ} are periodically updated and shared among all agents. This synchronization ensures that the agents' policies are aligned, promoting coordinated actions and reducing conflicts in the multi-agent system. The exchange of information among agents helps in faster convergence and more robust learning because agents can benefit from the collective knowledge and experience of the entire system.

The learning process involves training learner networks by sampling a minibatch consisting of $\mathcal{M}$ transitions, each of length $\mathcal{N}$, from the replay buffer. The target distribution for vehicle v is calculated as follows:

$$Y_v = \sum_{n=0}^{\mathcal{N}-1}\left(\zeta^n r_v^{j+n}\right) + \zeta^{\mathcal{N}} Q'\left(o_v^{j+\mathcal{N}}, a^{j+\mathcal{N}}|\omega^{Q'}\right). \tag{35}$$

The loss function for the critic network is defined as

$$\begin{aligned}\mathcal{L}(\omega^{Q}) &= \frac{1}{\mathcal{M}}\sum_{j}\frac{1}{V_e}\sum_{v} KL\left[Y_v \| Q\left(o_v^j, a^j|\omega^{Q}\right)\right], \\ &= \frac{1}{\mathcal{M}}\sum_{j}\frac{1}{V_e}\sum_{v} Y_v \log\frac{Y_v}{Q\left(o_v^j, a^j|\omega^{Q}\right)},\end{aligned} \tag{36}$$

where $KL\left[P\|Q\right]$ denotes the Kullback-Leibler divergence between the target and policy distributions. The policy network is updated using the policy gradient given by

$$\nabla_{\omega^{\mu}}\mathcal{J} = \frac{1}{\mathcal{M}}\sum_{j}\frac{1}{V_e}\sum_{v}\nabla_{d_v^j} Q\left(o_v^j, a^j|\omega^{Q}\right)\nabla_{\omega^{\mu}}\mu\left(o_v^j|\omega^{\mu}\right). \tag{37}$$

This comprehensive approach ensures that distributed agents within the MAD4PG-RIS framework can efficiently interact with the environment, optimize their performance, and learn from their experiences to improve the decision-making processes over time. A flowchart of the complete algorithm is shown in Fig. 3.

Similar to [47], the computational complexity per time slot of a distributed MADRL algorithm is

$$\mathcal{O}\left(\sum_{v=1}^{V}\left(\sum_{i=1}^{L_{l,act}} n_{l,act}^{(i)}\left(n_{l,act}^{(i+1)}\right) + \sum_{i=1}^{L_{l,cri}} n_{l,cri}^{(i)}\left(n_{l,cri}^{(i+1)}\right)\right)\right), \tag{38}$$

where V represents the number of agents, $n_{l,cri}^{(i)}$ and $n_{l,act}^{(i)}$ represent the number of neurons in the i-th layer of actor l and critic l, respectively. While $L_{l,act}$ and $L_{l,cri}$ indicate the number of hidden layers in the actor and critic, respectively. The complexity of the WOA per time slot can be analyzed, as demonstrated in [48].

$$\mathcal{O}\left(WD\right), \tag{39}$$

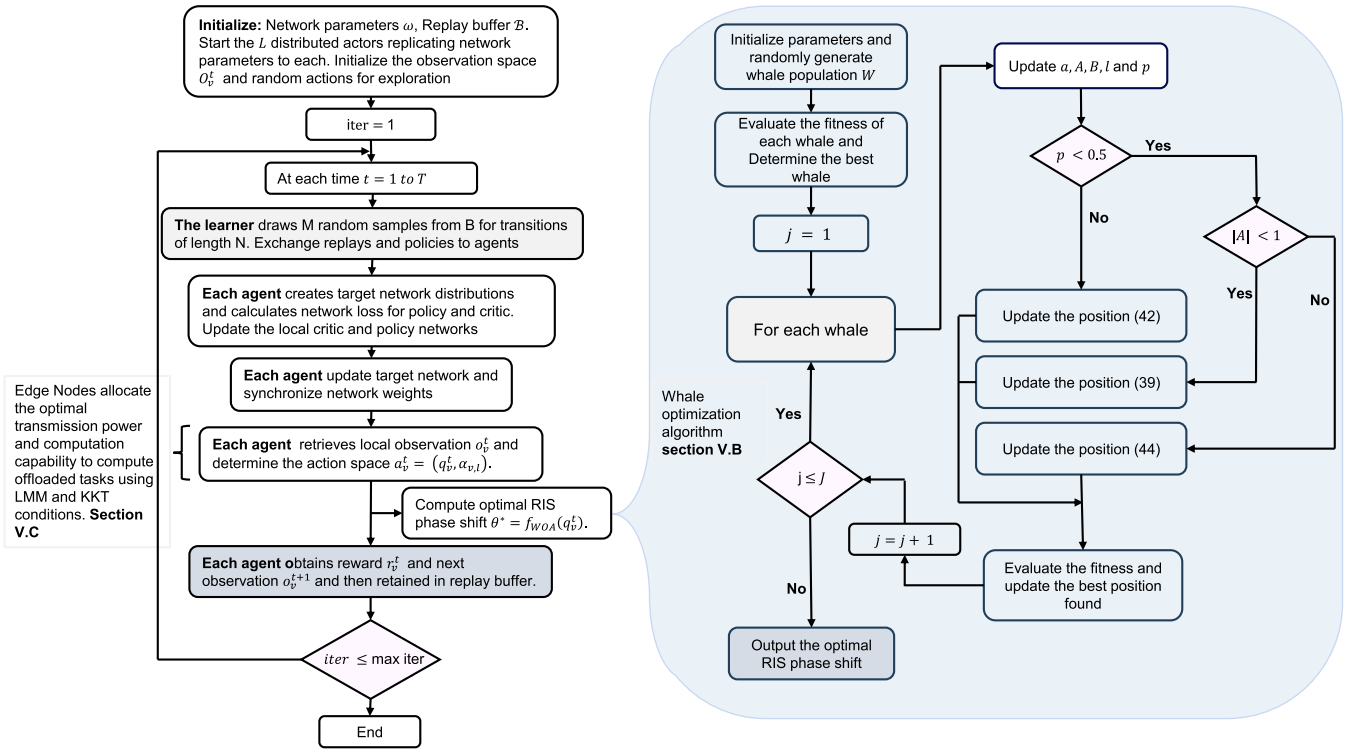


FIGURE 3. A flowchart of the MAD4PG-RIS algorithm.

where W is the number of whales, and D is the size of each search agent. In this regard, the complexity of the MAD4PG-RIS for I_{iter} iterations and time slot T is

$$\mathcal{O}\left(I_{iter}T\sum_{v=1}^{V}\left(\sum_{i=1}^{L_{l,act}} n_{l,act}^{(i)}(n_{l,act}^{(i+1)}+\sum_{i=1}^{L_{l,cri}} n_{l,cri}^{(i)}(n_{l,cri}^{(i+1)}\right)+WD\right). \tag{40}$$

The complexity depends on the number of agents, neural network architectures, and WOA. Local computations such as observation retrieval, action determination, and reward calculation are relatively inexpensive.

Each agent performs local updates asynchronously but periodically synchronizes its parameters with the learner. This synchronization ensures that all the agents' policies converge towards a common optimal policy. The discussion of how a distributed MADRL converges to a locally optimal policy is detailed in [49] and [50].

B. RIS PHASE SHIFT OPTIMIZATION

Given the offloading decision $q_{v,e}^t$ and task offloading ratio α_v, is a nonconvex optimization issue that determines the ideal RIS phase shift θ in $\mathcal{P}\mathbf{2}$. We used the Whale Optimization Algorithm (WOA) [51] described in [52] and [53]. WOA offers flexible methods for striking a balance between the algorithm's exploration and exploitation features. In comparison, this increases the likelihood of avoiding suboptimal solutions compared with other heuristic approaches. Hence, the WOA is applied to numerous optimization issues, including resource allocation in wireless communication [48]. In our design, the WOA algorithm $\theta^*(t) = f_{WOA}(q_{v,e})$ takes an offloading decision $q_{v,e}$ as the input and iteratively finds the optimal RIS phase shift $\theta^*(t)$ through J evolution. We initialized the whale population as $\theta'(0) = \{\theta_1'(0), \theta_2'(0), \ldots, \theta_W'(0)\}$, where W denotes the number of whales in the environment. We randomly generate RIS phase shift, the $i-th$ whale $\theta_i'(0)$. During the $j-th$ iteration, the global RIS phase shift that minimized the overall computational cost is determined. This is formulated as [53]:

$$\theta^*(j) = \underset{\theta' \in \{\theta^*(j-1)\} \cup \theta'(j-1)}{\arg\min} \Phi, \tag{41}$$

where $\theta^*(j)$ represents the global optimal RIS phase shift selected in the j iteration. There are two phases in WOA: exploitation and exploration.

1) **Exploitation phase:** The optimal solution is thought to be fairly close to the best whale in that other whales explore the surroundings of the present good whale for a better root. The ith whale within the whale population is updated using the following equation:

$$\mathbf{D} = \left|B_i(j)\theta^*(j) - \theta_i^*(j-1)\right|, \tag{42}$$

$$\theta_i^*(j) = \left|\theta^*(j) - A_i(j) \cdot \mathbf{D}\right|, \tag{43}$$

where $B_i(j) = 2 \cdot r_i(j)$ and $A_i(j) = a_i(j) \cdot (2r_i(j) - 1)$. The exploitation phase is selected when $|A_i(j)| < 1$. $r_i(j)$ is a real value between 0 and 1 which is generated

at random, and $a_i(j) = 2 \cdot \left(1 - \frac{j}{J}\right)$ is a scalar that decreases linearly over j iterations from 2 to 0. Given p is the probability of choosing spiral updating or shrinking encircling and that $p < 0.5$, the humpback whale's movement (i.e., spiral) is used:

$$\mathbf{D} = \left|\theta^*(j) - \theta_i^*(j-1)\right|, \tag{44}$$

$$\theta_i^*(j) = \left|\left(\mathbf{D} \cdot e^{\flat \cdot l_i(j)} \cdot \cos 2\pi \cdot l_i(j)\right) + \theta^*(j-1)\right|, \tag{45}$$

where $\flat = 1$ is a constant for establishing the spiral shape, and $l_i(j)$ is a value within the interval $[-1, 1]$.

2) **Exploration phase:** When $|A_i(j)| \geq 1$ the i-th whale is updated through exploration. The update rule is given as

$$\mathbf{D} = \left|\theta_i^{rand}(j) - \theta^*(j-1)\right|, \tag{46}$$

$$\theta_i'(j) = \left|\theta_i^{rand}(j) - A_i(j) \cdot \mathbf{D}\right|, \tag{47}$$

where $\theta_i^{rand}(j)$ is an RIS phase shift that is randomly generated. The ultimate output of the WOA is the resulting RIS phase shift, $\theta^*(J+1)$ after all J iterations are completed.

C. RESOURCE ALLOCATION AT EACH eN

In Section IV-C, the RA problem is transformed into an easily solved SCP problem. Hence, we employed the LMM and KKT conditions to solve TPRA and CRA problems. The Lagrange function for the TPRA is constructed as

$$L_1(\widetilde{p_{v,e}^t}, \lambda_1) = \widetilde{f}(p_{v,e}^t) - \lambda_1 \left(\sum_{\forall v \in \mathbb{V}_e^t} 2^{\widetilde{p_{v,e}^t}} - p_e\right). \tag{48}$$

Similarly, the Lagrange function for the CRA is

$$L_2(f_{v,e^t}, \lambda_2) = g(f_{v,e}^t) - \lambda_2 \left(\sum_{\forall k_v^t \in \mathbf{K}_{ve}^t} f_{v,e}^t - F_e\right), \tag{49}$$

where λ_1 and λ_2 are Lagrange multipliers. Equation (48) can then be determined by applying partial derivatives with respect to the choice variable and Lagrange multiplier, whereas (49) is solved following the KKT condition.

Differentiating (48) partially with respect to $\widetilde{p_{v,e}^t}$ gives,

$$\frac{\partial L_1(\widetilde{p_{v,e}^t}, \lambda_1)}{\partial \widetilde{p_{v,e}^t}} = b\delta_{v,e}^t - p_{v,e}^t \times \left(\lambda_1 + \sum_{\forall v \in \mathbb{V}_e^t} b\delta_{v,e}^t \|\hbar_{v,e}^t\|^2 \frac{\varphi_{v',e}^t}{\|\hbar_{v',e}^t\|^2 p_{v',e}^t}\right). \tag{50}$$

The optimal transmission power is determined by setting (50) to 0. Hence, the power for transmission by vehicle v in the jth iteration is obtained as

$$p_{v,e}^{t,(j+1)} = \frac{b\delta_{v,e}^t}{\lambda_1^{(j)} + \sum\limits_{\forall v' \in \mathbb{V}_{\hbar_{v,e}}^t} b\delta_{v,e}^t \|\hbar_{v,e}^t\|^2 \Upsilon_{v',e}^{(j)}}, \tag{51}$$

where $\Upsilon_{v',e}^{(j)}$ is;

$$\Upsilon_{v',e}^{(j)} = \sum_{\forall v' \in \mathbb{V}_{\hbar_{v,e}}^t} \|\hbar_{v',e}^t\|^2 p_{v',e}^{t,(j)} + \sum_{\forall e' \in \mathbb{E}} \sum_{\forall v' \in \mathbb{V}_{e'}^t} \|\hbar_{v',e}^t\|^2 p_{v',e'}^{t,(j)} + \mathbf{n}_e. \tag{52}$$

Following the KKT condition, (49) yields the following formulas [25]:

$$\lambda_2 \nabla f_{v,e}^t \left(\sum_{\forall k_v^t \in \mathbf{K}_{ve}^t} f_{v,e}^t - F_e\right) + \nabla f_{v,e}^t g(f_{v,e}^t) = 0,$$

$$\lambda_2 \left(\sum_{\forall k_v^t \in \mathbf{K}_{ve}^t} f_{v,e}^t - F_e\right) = 0,$$

$$\lambda_2 \geq 0. \tag{53}$$

The optimal CRA for a given task is obtained by solving (53). Hence,

$$f_{v,e}^{t*} = \frac{1/F_e\sqrt{d_k C}}{\sum\limits_{\forall k_v^t \in \mathbf{K}_{ve}^t} 1/F_e\sqrt{d_k C}}. \tag{54}$$

VI. NUMERICAL RESULTS

A. SETTINGS

In this section, we evaluate the performance of the proposed multi-agent DRL-based JTORA algorithm for the RIS-assisted NOMA system. We examined the overall situation in a 300 × 300 m^2 square region. $\mathbb{E} = 9$ eN containing RSUs were evenly dispersed over the road map, and RISs were deployed 300 m away from each edge node. Additionally, realistic vehicle trajectories were extracted from a 300 × 300 m^2 area of Qingyang District, Chengdu, China, on November 16, 2016, and used as traffic inputs obtained from the Didi GAIA open dataset [54], similar to that used in [25]. Specifically, we examined a scenario in the period 8 : 30 to 8 : 35 am, where the number of vehicular traces is 1181, the average dwell time is 233.2 (s), the variance of dwell time is 117.7, the average number of vehicles in each second is 878.8, the average speed of vehicles is 4.91 (m/s), and the variance speed of vehicles is 2.61. Furthermore, the simulation settings are set as follows: the number of M RIS elements is 12, eN antennas N is 8, the number of vehicular agents participating in task offloading V_e is [3, 27], the eN coverage r_e is 500 m, and the computational capabilities of eN and the vehicles are set to 10 GHz and 3 GHz, respectively. The W whale population is set to 12. We selected a small number of elements in the RIS to assess the effect of using the RIS on communication quality.

In addition, we assessed the effect of increasing RIS element size.

TABLE 3. Parameters.

Model parameter	value
Bandwidth b [25]	20 MHz
V2I communication range r_e [25]	500 m
Computation capacity of eN f_e	$[10, 15]$ GHz
Edge node power for task processing p_e	$[15, 20]$ W
Maximum V2I transmission power $p^t_{v,e}$	1 W
Computation cycles C	$[100, 500]$ cycles/bit
Deadline of task T_{k_v} [25]	$[5, 10]$ s
Task size d_k [25]	$[0.01, 5]$ MB
Vehicle local power	10 W
Computation capacity of vehicles f_v	3 GHz
Weighting factor λ [10]	0.5
Delay threshold $\mathcal{D}_{th}$ [44]	200 s
Energy threshold $\mathcal{E}_{th}$ [44]	200 J
Additive white Gaussian noise n_e	-90 dBm
Large scale path loss exponent γ	3
Hyperparameters	**value**
Batch size M	256
Exploration constant ϵ	0.3
Discount	0.996
Replay buffer size $\lvert\mathcal{B}\rvert$	1×10^6
Learning rate	1×10^{-4}
Distributed actors network update period t_{act}	1000
Target network parameter update period t_{tgt}	100

The following sections describe the design of the critic and policy networks used to create the algorithms. We employed ten distributed actors. The local policy network and target policy network are identical and feature three hidden layers, each comprising 256 neurons in a fully connected five-layer neural network. The local critic network and target critic network are also identical and feature three hidden layers of the first two 512 neurons and the other 256 neurons in a fully connected five-layer neural network similar to [25]. The Adam optimizer is used for network weight updates, whereas the Rectifier Linear Unit (ReLU) is used as the activation function [35]. The system model and hyperparameters of the algorithms are listed in Table 3.

The performance was evaluated by comparing it with other algorithms, as described below.

- Local computation *(Local)*: Vehicles prefer to compute tasks locally given their computing capability.
- *D4PG* [46]: It is a single agent that uses the global system status as input and is implemented using a DDPG agent. It jointly makes offloading decisions and offloading ratios and allocates V2I transmission power and computation resources.
- *D4PG-RIS*: An extension of D4PG with a WOA to optimize the RIS phase shift coefficient.
- *MADDPG* [20]: It is designed by using the CTDE approach. To optimize the expected return, each agent updates its actor and critic networks based on their combined observations and actions during the training phase. The trained actor network is then used to obtain offloading decisions and offloading ratios based on local observations.
- *MADDPG-RIS*: The MADDPG is extended with the WOA to optimize the RIS phase-shift parameters.

Furthermore, after training all seven algorithms for 4100 evolutions, we employed three performance metrics to assess these algorithms: the average service delay (ASD), average energy consumption (AEN), and average service rate. If the service time of task k^t_v is less than its deadline T_{k_v}, then it is denoted as a completed task. The service ratio is the percentage of tasks completed of the total task requests received by the vehicle.

B. RESULTS AND ANALYSIS

1) CONVERGENCE ANALYSIS OF ALGORITHMS

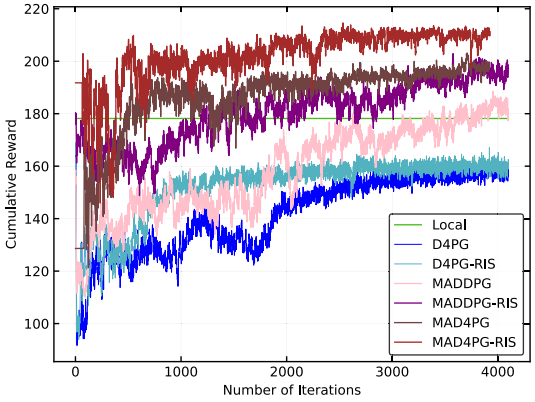
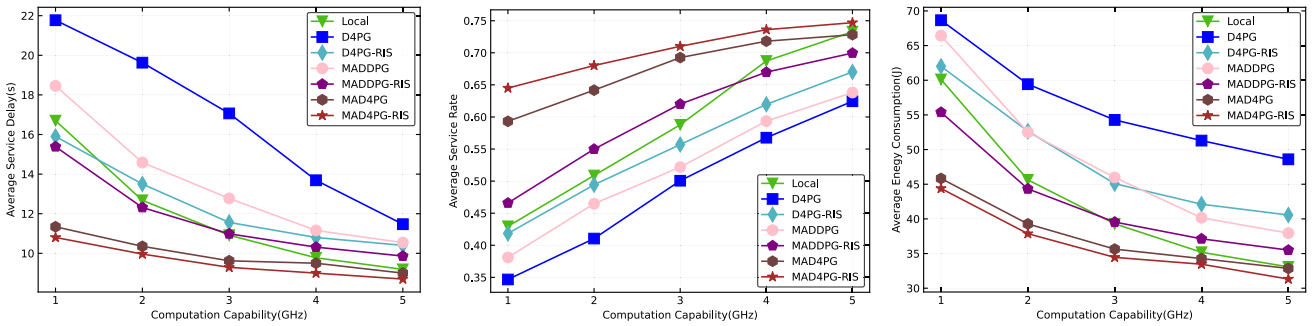


FIGURE 4. Convergence performance of algorithms.

Fig. 4 shows the convergence behavior of the seven algorithms. Notably, the algorithms implemented with RIS converge faster typically around 1500 iterations, and achieve higher cumulative rewards owing to their ability to dynamically adjust the wireless environment. The WOA specifically fine-tunes the RIS phase-shift parameters, which optimizes the signal propagation and reduces the energy consumption or delay. This optimization process allows RIS-based algorithms to achieve better performance than those without an RIS. The MAD4PG-RIS algorithm outperformed the other algorithms because it combined the advantages of the RIS with a multi-agent deep reinforcement learning approach. The distributed actor-critic framework in MAD4PG accelerates the learning process by enabling parallel experience replay, which not only speeds up convergence but also leads to a higher cumulative reward. The distributed nature allows the algorithm to effectively handle complex environments, resulting in superior performance compared to MADDPG and other algorithms. In the baseline model, where the computational capabilities of the vehicles and eNs remained constant, the cumulative reward remained stable. This is because vehicles consistently compute a similar proportion of tasks locally, leading to a predictable and steady performance.



(a) Average Service Delay (ASD) (b) Average Service Rate (ASR) (c) Average Energy Consumption (AEN)

FIGURE 5. **Performance comparison under different computing capabilities of vehicles.**

The lack of dynamic adaptation in this local model limits its ability to further improve performance.

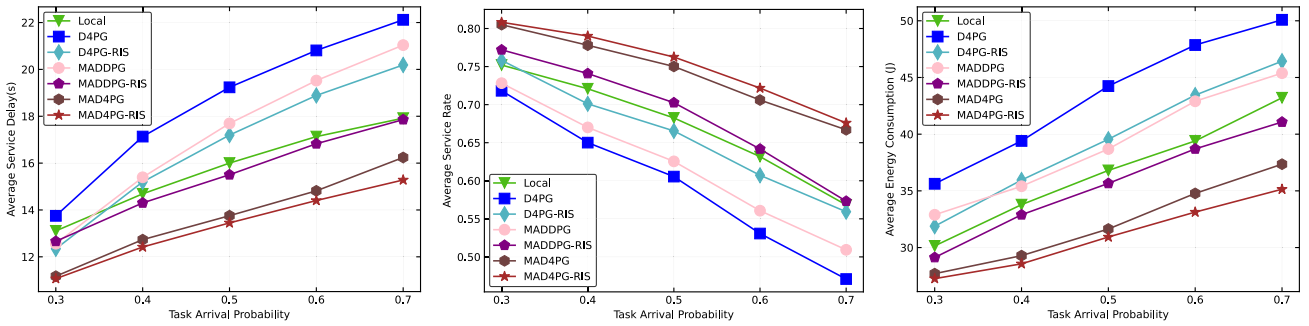
The reason for the performance disparity between MAD4PG and MADDPG lies in their distinct approaches to balancing exploration and exploitation. MAD4PG's use of distributed actors allows for more diverse and efficient sampling of the experience replay buffer, which in turn accelerates learning and leads to faster convergence. MADDPG, while effective, may not sample experiences efficiently, leading to a slower convergence and lower overall rewards. D4PG lacks distributed learning capabilities that allow parallel experience sampling. Without these capabilities, D4PG may require more iterations to explore the state space adequately and determine an optimal policy, particularly in environments with many interacting entities. The single-agent nature of D4PG limits its ability to leverage the full potential of the environment, resulting in lower overall rewards compared to multi-agent strategies that can better coordinate and optimize actions across multiple entities. If the D4PG algorithm is implemented with RIS, it benefits from an optimized signal environment, potentially improving its performance compared to a version without RIS. However, the lack of distributed learning and coordination among agents still places it at a disadvantage relative to multi-agent approaches, such as MAD4PG-RIS, which can more effectively exploit the RIS-enhanced environment.

2) EVALUATION OF ALGORITHMS UNDER VARYING VEHICLE COMPUTING CAPACITIES

Fig. 5 presents a comparative analysis of the seven algorithms for different vehicle-computing capacities. In this set of experiments, we assessed vehicle computing capabilities ranging from $f_v \sim 1$ GHz to $f_v \sim 5$ GHz, with an eN computation capability f_e set at 10 GHz. Increased computing power allows vehicles to execute more tasks. Fig. 5(a) compares the ASD across the seven algorithms, showing a decline in ASD as vehicle computing power increases. This is because the vehicles can execute more tasks locally, reducing the need to offload tasks to eN. This results in a decrease in ASD across all algorithms, as more computing tasks are handled directly by the vehicles, thereby shortening the overall processing time. The MAD4PG-RIS algorithm achieves the lowest ASD because it optimally balances the tasks between local vehicle processing and offloading to the eN, owing to its advanced multi-agent learning strategy and RIS phase adjustments. Fig. 5(b) shows the ASR, indicating that the ASR improved for all algorithms with increased vehicle computing capacity. Fig. 5(c) shows an AEN comparison of the algorithms, demonstrating that the proposed MAD4PG-RIS algorithm also achieves a lower energy consumption. Across all performance indicators, the RIS-optimized algorithms (D4PG-RIS, MADDPG-RIS, and MAD4PG-RIS) outperformed the others (i.e., D4PG, MADDPG, and MAD4PG) because the RIS technology enhanced the communication link quality, reduced packet loss and increased the reliability of task completion. In addition, the local algorithm approaches the performance of both MAD4PG and MAD4PG-RIS when the vehicle computing power is substantial (i.e., 4 GHz to 5 GHz). Under these conditions, the advantages of multi-agent learning and RIS optimization become less pronounced because vehicles can handle their computational demands effectively.

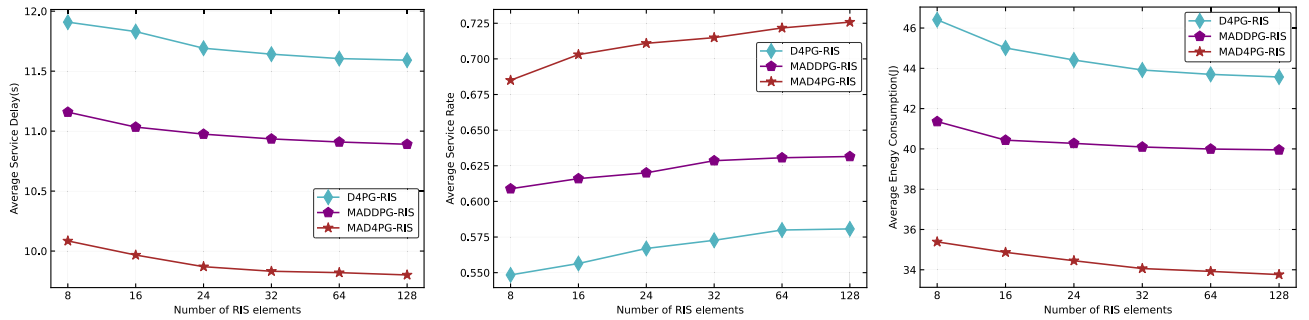
3) EFFECTS OF TASK ARRIVAL PROBABILITY

Fig. 6 presents an analysis of the seven algorithms under varying task arrival probabilities for vehicles. Tasks arrive stochastically in vehicles and require computation. In this evaluation, we considered task arrival probabilities to increase from 0.3 to 0.7 per time slot. As anticipated, all algorithms exhibited worse performance as task arrival probabilities increased. This decline is owing to the higher computational load, which led to increased delays, lower service rates, and higher energy consumption. When tasks arrive less frequently (e.g., at a probability of 0.3), vehicles have more time and resources to compute each task, resulting in improved performance metrics. Between the task arrival probabilities of 0.3 and 0.4, there is a minimal performance difference between the algorithms. This is because the available computational resources (e.g., vehicle and eN computing capacities) are sufficient to efficiently handle



(a) Average Service Delay (ASD) (b) Average Service Rate (ASR) (c) Average Energy Consumption (AEN)

FIGURE 6. Performance comparison under task arrival probabilities at vehicles.



(a) Average Service Delay (ASD) (b) Average Service Rate (ASR) (c) Average Energy Consumption (AEN)

FIGURE 7. Performance comparison under different number of RIS elements.

incoming tasks. In this range, the task scheduling and resource allocation strategies of each algorithm are less critical because there is ample capacity to process incoming tasks, leading to similar performance levels.

When the chance of task arrival exceeds 0.4, the computational load becomes more challenging to manage, leading to a greater divergence in performance among the algorithms. Algorithms that are better at optimizing resource allocation and communication, such as those with an RIS, have demonstrated significant advantages. The MAD4PG-RIS algorithm, for example, achieves the lowest ASD because of its ability to dynamically optimize both the task distribution and the wireless environment, ensuring that tasks are processed more efficiently, even under heavy loads. Fig. 6(a) presents a comparison of the ASD of the seven algorithms, with MAD4PG-RIS achieving the lowest ASD. This is because its advanced distributed learning strategy allows for better coordination between vehicles and eN, effectively managing increased task load. Fig. 6(b) and 6(c) compare the ASR and AEN of the algorithms, respectively.

4) EFFECTS OF RIS ELEMENTS

Fig. 7 shows the effect of varying the quantity of RIS elements M from 8 to 128 on the performance of the RIS-implemented algorithms. The findings clearly indicate that, with an increase in the number of reflecting elements in the RIS, the surface becomes more capable of fine-tuning the signal reflections, leading to better signal quality, reduced interference, and more efficient use of the wireless spectrum. This enhancement directly affects the performance metrics of the algorithms. This enhancement underscores the capability of the WOA to optimize the phase shift of the RIS elements effectively. The ASD values of the algorithms are compared in Fig. 7(a). As observed, the ASD gradually decreased with the increasing number of elements, reflecting the improved efficiency of our proposed method. Notably, the MAD4PG-RIS algorithm achieves the lowest delay, thereby demonstrating its superior performance. Fig. 7(b) shows the ASR across the algorithms. A positive correlation was observed between the performance and the number of RIS elements, with the ASR steadily increasing. The improved signal quality and reduced latency afforded by more RIS elements indicate that a greater number of tasks can be completed within their deadlines, thereby increasing the service rate. The MAD4PG-RIS algorithm benefits the most from this because its advanced learning and decision-making framework can better exploit the enhanced communication environment provided by the RIS. This leads to the highest ASR among the algorithms, reaffirming

its effectiveness in managing complex task offloading scenarios.

Finally, Fig. 7(c) presents a comparison of the AEN among the algorithms. A consistent reduction in energy consumption was observed as the quantity of the RIS passive elements increased. This demonstrates the efficiency of RIS in lowering energy usage during the offloading process. Overall, the results convincingly show that integrating RIS enhances the algorithm performance by reducing delays and lowering energy consumption, with the MAD4PG-RIS algorithm emerging as particularly effective. The MAD4PG-RIS algorithm consistently achieved the best results across all the performance indicators (ASD, ASR, and AEN) as the number of RIS elements increased. Its distributed learning approach allows for more effective coordination between agents, whereas the RIS optimizes the communication environment, ensuring that tasks are processed quickly, successfully, and with the least energy consumption. This combination of advanced learning and communication optimization sets MAD4PG-RIS apart from the other algorithms.

VII. CONCLUSION

In this paper, we introduce an innovative RIS-assisted NOMA-based VEC architecture that leverages RIS to enhance the efficiency of computational task offloading in vehicular networks. The architecture accommodates both binary and partial computation tasks offloading to edge nodes (eN) for processing, thereby offering enhanced flexibility and efficiency in addressing diverse computational demands from vehicles.

We have formulated a Joint Task Offloading and Resource Allocation (JTORA) problem with the dual objective of minimizing overall service delay and energy consumption. We validated the effectiveness of our proposed approach by analyzing performance metrics including service delays, service rates, and energy consumption. Our experimental results demonstrate that the MAD4PG-RIS algorithm significantly reduces both the overall service delay and energy consumption compared with traditional methods. Furthermore, we observed that algorithms employing the Whale Optimization Algorithm (WOA) for RIS phase optimization consistently outperformed their counterparts that did not utilize RIS-assisted communication. Our proposed RIS-assisted NOMA-based VEC architecture and MAD4PG-RIS algorithm offer a robust solution to the challenges of task offloading and resource allocation in vehicular networks. This study paves the way for future research on optimizing VEC systems and highlights the potential of RIS technology to advance the capabilities of intelligent transportation systems, especially in the task offloading domain.

APPENDIX

The constants $\xi_{v,e}^t$ and $\delta_{v,e}^t$ in (29) are derived to ensure that the linear approximation $\left(\xi_{v,e}^t + \delta_{v,e}^t \log_2 \varphi_{v,e}^t\right)$ is tight at $\varphi_{v,e}^t = \overline{\varphi_{v,e}^t}$. This means that the approximation equals the function $\log_2\left(1+\varphi_{v,e}^t\right)$ at a specific point $\varphi_{v,e}^t = \overline{\varphi_{v,e}^t}$ and also has the same slope (derivative) at that point. The constants are derived as follows.

$$\log_2\left(1+\varphi_{v,e}^t\right) \geq \left(\xi_{v,e}^t + \delta_{v,e}^t \log_2 \varphi_{v,e}^t\right), \tag{55}$$

At $\varphi_{v,e}^t = \overline{\varphi_{v,e}^t}$,

$$\log_2\left(1+\overline{\varphi_{v,e}^t}\right) = \left(\xi_{v,e}^t + \delta_{v,e}^t \log_2 \overline{\varphi_{v,e}^t}\right). \tag{56}$$

Taking derivatives with respect to $\overline{\varphi_{v,e}^t}$, implies

$$\frac{1}{\left(1+\overline{\varphi_{v,e}^t}\right)\ln 2} = \delta_{v,e}^t \frac{1}{\overline{\varphi_{v,e}^t}\ln 2} \tag{57}$$

Then making $\delta_{v,e}^t$ the subject, we have:

$$\delta_{v,e}^t = \frac{\overline{\varphi_{v,e}^t}}{1+\overline{\varphi_{v,e}^t}}. \tag{58}$$

Substituting $\delta_{v,e}^t$ in (56) implies:

$$\log_2\left(1+\overline{\varphi_{v,e}^t}\right) = \xi_{v,e}^t + \frac{\overline{\varphi_{v,e}^t}}{\left(1+\overline{\varphi_{v,e}^t}\right)} \log_2 \overline{\varphi_{v,e}^t}. \tag{59}$$

Hence,

$$\xi_{v,e}^t = \log_2\left(1+\overline{\varphi_{v,e}^t}\right) - \frac{\overline{\varphi_{v,e}^t}}{\left(1+\overline{\varphi_{v,e}^t}\right)} \log_2 \overline{\varphi_{v,e}^t}. \tag{60}$$

REFERENCES

[1] Q.-V. Pham, F. Fang, V. N. Ha, Md. J. Piran, M. Le, L. B. Le, W.-J. Hwang, and Z. Ding, "A survey of multi-access edge computing in 5G and beyond: Fundamentals, technology integration, and state-of-the-art," *IEEE Access*, vol. 8, pp. 116974–117017, 2020.

[2] K. Liu, X. Xu, M. Chen, B. Liu, L. Wu, and V. C. S. Lee, "A hierarchical architecture for the future Internet of Vehicles," *IEEE Commun. Mag.*, vol. 57, no. 7, pp. 41–47, Jul. 2019.

[3] X. Xu, K. Liu, K. Xiao, L. Feng, Z. Wu, and S. Guo, "Vehicular fog computing enabled real-time collision warning via trajectory calibration," *Mobile Netw. Appl.*, vol. 25, no. 6, pp. 2482–2494, Dec. 2020.

[4] K. Zhang, J. Cao, S. Maharjan, and Y. Zhang, "Digital twin empowered content caching in social-aware vehicular edge networks," *IEEE Trans. Computat. Social Syst.*, vol. 9, no. 1, pp. 239–251, Feb. 2022.

[5] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *Proc. IEEE*, vol. 105, no. 12, pp. 2347–2381, Dec. 2017.

[6] D. K. Patel, H. Shah, Z. Ding, Y. L. Guan, S. Sun, Y. C. Chang, and J. M.-Y. Lim, "Performance analysis of NOMA in vehicular communications over i.n.i.d Nakagami-*m* fading channels," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6254–6268, Oct. 2021.

[7] M.-M. Zhao, Q. Wu, M.-J. Zhao, and R. Zhang, "Exploiting amplitude control in intelligent reflecting surface aided wireless communication with imperfect CSI," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4216–4231, Jun. 2021.

[8] Z. Xu, J. Liu, J. Zou, and Z. Wen, "Energy-efficient design for IRS-assisted NOMA-based mobile edge computing," *IEEE Commun. Lett.*, vol. 26, no. 7, pp. 1618–1622, Jul. 2022.

[9] J. Liu, S. Guo, Q. Wang, C. Pan, and L. Yang, "Optimal multi-user offloading with resources allocation in mobile edge cloud computing," *Comput. Netw.*, vol. 221, Feb. 2023, Art. no. 109522.

[10] F. Zhang, G. Han, L. Liu, Y. Zhang, Y. Peng, and C. Li, "Cooperative partial task offloading and resource allocation for IIoT based on decentralized multiagent deep reinforcement learning," *IEEE Internet Things J.*, vol. 11, no. 3, pp. 5526–5544, Feb. 2024.

[11] Y. Chen, J. Wu, J. Han, H. Zhao, and S. Deng, "A game-theoretic approach-based task offloading and resource pricing method for idle vehicle devices assisted VEC," *IEEE Internet Things J.*, vol. 11, no. 12, pp. 21954–21969, Jun. 2024.

[12] K. Nguyen, S. Drew, C. Huang, and J. Zhou, "Parked vehicles task offloading in edge computing," *IEEE Access*, vol. 10, pp. 41592–41606, 2022.

[13] W. Zhan, C. Luo, J. Wang, C. Wang, G. Min, H. Duan, and Q. Zhu, "Deep-reinforcement-learning-based offloading scheduling for vehicular edge computing," *IEEE Internet Things J.*, vol. 7, no. 6, pp. 5449–5465, Jun. 2020.

[14] J. Liu, M. Ahmed, M. A. Mirza, W. U. Khan, D. Xu, J. Li, A. Aziz, and Z. Han, "RL/DRL meets vehicular task offloading using edge and vehicular cloudlet: A survey," *IEEE Internet Things J.*, vol. 9, no. 11, pp. 8315–8338, Jun. 2022.

[15] M. A. Mirza, J. Yu, S. Raza, M. Krichen, M. Ahmed, W. U. Khan, K. Rabie, and T. Shongwe, "DRL-assisted delay optimized task offloading in automotive-Industry 5.0 based VECNs," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 35, no. 6, Jun. 2023, Art. no. 101512.

[16] Y. Liu, H. Yu, S. Xie, and Y. Zhang, "Deep reinforcement learning for offloading and resource allocation in vehicle edge computing and networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 11158–11168, Nov. 2019.

[17] A. S. Kumar, L. Zhao, and X. Fernando, "Task offloading and resource allocation in vehicular networks: A Lyapunov-based deep reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 72, pp. 1–14, May 2023.

[18] I. Althamary, C.-W. Huang, and P. Lin, "A survey on multi-agent reinforcement learning methods for vehicular networks," in *Proc. 15th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2019, pp. 1154–1159.

[19] B. Uwizeyimana, O. Muta, A. H. A. El-Malek, M. Abo-Zahhad, and M. Elsabrouty, "A multi-agent multi-armed bandit approach for user pairing in UAV-assisted NOMA-networks," in *Proc. Int. Symp. Netw., Comput. Commun. (ISNCC)*, Oct. 2023, pp. 1–6.

[20] K. Zhang, J. Cao, and Y. Zhang, "Adaptive digital twin and multiagent deep reinforcement learning for vehicular edge computing and networks," *IEEE Trans. Ind. Informat.*, vol. 18, no. 2, pp. 1405–1413, Feb. 2022.

[21] Md. Z. Alam and A. Jamalipour, "Multi-agent DRL-based Hungarian algorithm (MADRLHA) for task offloading in multi-access edge computing Internet of Vehicles (IoVs)," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7641–7652, Sep. 2022.

[22] Y. Ju, Y. Chen, Z. Cao, L. Liu, Q. Pei, M. Xiao, K. Ota, M. Dong, and V. C. M. Leung, "Joint secure offloading and resource allocation for vehicular edge computing network: A multi-agent deep reinforcement learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 5, pp. 5555–5569, May 2023.

[23] P. Sunehag, G. Lever, A. Gruslys, W. M. Czarnecki, V. Zambaldi, M. Jaderberg, M. Lanctot, N. Sonnerat, J. Z. Leibo, and K. Tuyls, "Value-decomposition networks for cooperative multi-agent learning," 2017, *arXiv:1706.05296*.

[24] J. Foerster, G. Farquhar, T. Afouras, N. Nardelli, and S. Whiteson, "Counterfactual multi-agent policy gradients," in *Proc. Conf. AAAI Artif. Intell.*, Apr. 2018, vol. 32, no. 1, pp. 2974–2982.

[25] X. Xu, K. Liu, P. Dai, F. Jin, H. Ren, C. Zhan, and S. Guo, "Joint task offloading and resource optimization in NOMA-based vehicular edge computing: A game-theoretic DRL approach," *J. Syst. Archit.*, vol. 134, Jan. 2023, Art. no. 102780. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S138376212200265X

[26] S. Raza, S. Wang, M. Ahmed, M. R. Anwar, M. A. Mirza, and W. U. Khan, "Task offloading and resource allocation for IoV using 5G NR-V2X communication," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 10397–10410, Jul. 2022.

[27] M. A. Mirza, J. Yu, M. Ahmed, S. Raza, W. U. Khan, F. Xu, and A. Nauman, "DRL-driven zero-RIS assisted energy-efficient task offloading in vehicular edge computing networks," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 35, no. 10, Dec. 2023, Art. no. 101837.

[28] Y. Zhu, B. Mao, and N. Kato, "A dynamic task scheduling strategy for multi-access edge computing in IRS-aided vehicular networks," *IEEE Trans. Emerg. Topics Comput.*, vol. 10, no. 4, pp. 1761–1771, Oct. 2022.

[29] Y. Xie, L. Shi, Z. Li, X. Ding, and F. Liu, "Roadside IRS assisted task offloading in vehicular edge computing network," in *Proc. Int. Conf. Collaborative Comput., Netw., Appl. Worksharing*. Cham, Switzerland: Springer, Oct. 2023, pp. 365–384.

[30] Z. Li, M. Chen, Z. Yang, J. Zhao, Y. Wang, J. Shi, and C. Huang, "Energy efficient reconfigurable intelligent surface enabled mobile edge computing networks with NOMA," *IEEE Trans. Cognit. Commun. Netw.*, vol. 7, no. 2, pp. 427–440, Jun. 2021.

[31] G. Manogaran, S. Mumtaz, C. X. Mavromoustakis, E. Pallis, and G. Mastorakis, "Artificial intelligence and blockchain-assisted offloading approach for data availability maximization in edge nodes," *IEEE Trans. Veh. Technol.*, vol. 70, no. 3, pp. 2404–2412, Mar. 2021.

[32] J. Du, Y. Sun, N. Zhang, Z. Xiong, A. Sun, and Z. Ding, "Cost-effective task offloading in NOMA-enabled vehicular mobile edge computing," *IEEE Syst. J.*, vol. 17, no. 1, pp. 928–939, Mar. 2023.

[33] T. Zhou, Y. Fu, D. Qin, X. Nie, N. Jiang, and C. Li, "Secure and multistep computation offloading and resource allocation in ultradense multitask NOMA-enabled IoT networks," *IEEE Internet Things J.*, vol. 11, no. 3, pp. 5347–5361, Feb. 2024.

[34] T. Zhou, H. Hu, N. Jiang, X. Nie, X. Li, and C. Li, "Joint security service assignment, power control and multi-step computation offloading in ultra-dense multi-task NOMA-enabled networks," *IEEE Trans. Veh. Technol.*, vol. 73, no. 7, pp. 10211–10227, Jul. 2024.

[35] H. Zhu, Q. Wu, X.-J. Wu, Q. Fan, P. Fan, and J. Wang, "Decentralized power allocation for MIMO-NOMA vehicular edge computing based on deep reinforcement learning," *IEEE Internet Things J.*, vol. 9, no. 14, pp. 12770–12782, Jul. 2022.

[36] Y. Liu, H. Zhang, K. Long, A. Nallanathan, and V. C. M. Leung, "Energy-efficient subchannel matching and power allocation in NOMA autonomous driving vehicular networks," *IEEE Wireless Commun.*, vol. 26, no. 4, pp. 88–93, Aug. 2019.

[37] N. Agrawal, A. Bansal, K. Singh, and C.-P. Li, "Performance evaluation of RIS-assisted UAV-enabled vehicular communication system with multiple non-identical interferers," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 9883–9894, Jul. 2022.

[38] M. A. Javed, T. N. Nguyen, J. Mirza, J. Ahmed, and B. Ali, "Reliable communications for cybertwin-driven 6G IoVs using intelligent reflecting surfaces," *IEEE Trans. Ind. Informat.*, vol. 18, no. 11, pp. 7454–7462, Nov. 2022.

[39] H. Liao, Z. Wang, Z. Zhou, Y. Wang, H. Zhang, S. Mumtaz, and M. Guizani, "Blockchain and semi-distributed learning-based secure and low-latency computation offloading in space-air-ground-Integrated power IoT," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 3, pp. 381–394, Apr. 2022.

[40] Z. Wang, J. Qiu, Y. Zhou, Y. Shi, L. Fu, W. Chen, and K. B. Letaief, "Federated learning via intelligent reflecting surface," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 808–822, Feb. 2022.

[41] S. M. R. Islam, N. Avazov, O. A. Dobre, and K.-S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 721–742, 2nd Quart., 2017.

[42] L. Zhang, S. Lai, J. Xia, C. Gao, D. Fan, and J. Ou, "Deep reinforcement learning based IRS-assisted mobile edge computing under physical-layer security," *Phys. Commun.*, vol. 55, Dec. 2022, Art. no. 101896.

[43] J. Chen, S. Chen, Q. Wang, B. Cao, G. Feng, and J. Hu, "IRAF: A deep reinforcement learning approach for collaborative mobile edge computing IoT networks," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 7011–7024, Aug. 2019.

[44] B. Gong and X. Jiang, "Dependent task-offloading strategy based on deep reinforcement learning in mobile edge computing," *Wireless Commun. Mobile Comput.*, vol. 2023, pp. 1–12, Jan. 2023.

[45] J. Papandriopoulos and J. S. Evans, "SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3711–3724, Aug. 2009.

[46] G. Barth-Maron, M. W. Hoffman, D. Budden, W. Dabney, D. Horgan, D. TB, A. Muldal, N. Heess, and T. Lillicrap, "Distributed distributional deterministic policy gradients," 2018, *arXiv:1804.08617*.

[47] G. Zhou, L. Zhao, G. Zheng, S. Song, J. Zhang, and L. Hanzo, "Multiobjective optimization of space–air–ground-integrated network slicing relying on a pair of central and distributed learning algorithms," *IEEE Internet Things J.*, vol. 11, no. 5, pp. 8327–8344, Mar. 2024.

[48] Q.-V. Pham, S. Mirjalili, N. Kumar, M. Alazab, and W.-J. Hwang, "Whale optimization algorithm with applications to resource allocation in wireless networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4285–4297, Apr. 2020.

[49] J. Su, S. Adams, and P. Beling, "Value-decomposition multi-agent actor-critics," in *Proc. Conf. AAAI Artif. Intell.*, 2021, vol. 35, no. 13, pp. 11352–11360.

[50] K. Zhang, Z. Yang, and T. Başar, "Multi-agent reinforcement learning: A selective overview of theories and algorithms," in *Handbook of Reinforcement Learning and Control*. Cham, Switzerland: Springer, Jun. 2021, pp. 321–384.
[51] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, May 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0965997816300163
[52] Y. M. Park, S. S. Hassan, Y. K. Tun, Z. Han, and C. S. Hong, "Joint resources and phase-shift optimization of MEC-enabled UAV in IRS-assisted 6G THz networks," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2022, pp. 1–7.
[53] J. Wu, Z. Yu, J. Guo, Z. Tang, T. Wang, and W. Jia, "A fast task offloading optimization framework for IRS-assisted multi-access edge computing system," 2023, *arXiv:2307.08474*.
[54] *Data Source: DiDi Chuxing GAIA Open Dataset Initiative*. Accessed: Aug. 15, 2024. [Online]. Available: https://outreach.didichuxing.com/research/opendata/en/

ABDUL-BAAKI YAKUBU received the B.Sc. degree (Hons.) in telecommunication engineering from the Kwame Nkrumah University of Science and Technology, Ghana, in 2021. He is currently pursuing the master's degree in electronics and communication engineering with Egypt-Japan University of Science and Technology (E-JUST), Egypt. His research interests include edge computing, 5G, 6G, and beyond wireless technologies, reinforcement learning for wireless communication, computer networks, and digital signal processing.

AHMED H. ABD EL-MALEK (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from Alexandria University, Egypt, in 2007 and 2010, respectively, and the Ph.D. degree from the Electrical Engineering Department, King Fahd University of Petroleum and Minerals (KFUPM), Saudi Arabia, in 2016. He is currently an Associate Professor with the Department of Electronics and Communications Engineering, Egypt-Japan University of Science and Technology. His research interests include cognitive radio, designs and analysis of wireless networks, network coding, physical layer security, and interference cancellation.

MOHAMMED ABO-ZAHHAD (Senior Member, IEEE) received the B.S.E.E. and M.S.E.E. degrees in electrical engineering from Assiut University, Assiut, Egypt, in 1979 and 1983, respectively, and the dual Ph.D. degree from the University of Kent, Canterbury, U.K., and Assiut University (channel system), in 1988. He has been the Director of the AU Management Information System (MIS) Center and the Vice-Dean of Graduate Studies with the Faculty of Engineering, Assiut University, since August 2006. He has also been a Professor in electronics and communication engineering, since January 1999, the Dean of the School of Electronics, Communication and Computer Engineering, and a Professor in communication and electronics engineering with Egypt-Japan University of Science and Technology (E-JUST), Alexandria, Egypt, since January 2017. He is currently the General Director of E-JUST Information and Communication Technology Center, Alexandria. His research interests include switched capacitor, optical and digital filters, biomedical and genomic signal processing, speech processing, data compression, wavelet transforms, genetic algorithms, immune algorithms, wireless sensor networks, and electronic systems. He has published more than 160 papers in national and international journals and conferences in the above fields.

OSAMU MUTA (Member, IEEE) received the Associate B.E. degree from Sasebo Institute of Technology, in 1994, the B.E. degree from Ehime University, in 1996, the M.E. degree from Kyushu Institute of Technology, in 1998, and the Ph.D. degree from Kyushu University, Japan, in 2001. In 2001, he joined the Graduate School of Information Science and Electrical Engineering, Kyushu University, as an Assistant Professor. From 2010 to 2023, he was an Associate Professor with the Center for Japan-Egypt Cooperation in Science and Technology, Kyushu University. Since 2023, he has been a Professor with the Faculty of Information Science and Electrical Engineering, Kyushu University. His research interests include signal processing techniques for wireless communication and power line communication, MIMO techniques, interference coordination techniques, low-power wide-area networks, and nonlinear distortion compensation techniques for high-power amplifiers. He is a Senior Member of the Institute of Electronics, Information, and Communication Engineering (IEICE). He was a recipient of the 2005 Active Research Award from the IEICE Radio Communication Systems Technical Committee, the Chairperson's Award for Excellent Paper from the IEICE Communication Systems Technical Committee, in 2014, 2015, and 2017, the 2020 IEICE Communications Society Best Paper Award, and the International Symposium on Computing and Networking 2022 (CANDAR-22) Best Paper Award.

MAHA M. ELSABROUTY (Senior Member, IEEE) received the B.Sc. degree (Hons.) in electrical, electronics, and communication engineering from Cairo University, Egypt, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Ottawa. She is currently a Professor of wireless communications and signal processing with the Department of Electronics and Communications Engineering, Egypt-Japan University of Science and Technology (E-JUST) University. Her current research interests include massive MIMO techniques, interference management in HetNets, cognitive radio, intelligent techniques for wireless communications, and green communication systems.