Computer Networks 260 (2025) 111099

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

An incentive mechanism for joint sensing and communication Vehicular Crowdsensing by Deep Reinforcement Learning

Gaoyu Luo [a], Shanhao Zhan [a], Chenyi Liang [a], Zhibin Gao [b,*], Yifeng Zhao [a], Lianfen Huang [a]

[a] School of Informatics, Xiamen University, Xiamen, 361005, Fujian, China
[b] Navigation Institute, Jimei University, Xiamen, 361021, Fujian, China

ARTICLE INFO

Keywords:
Vehicular Crowdsensing
Incentive mechanism
Data collection
Small-World networks
Deep Reinforcement Learning

ABSTRACT

Vehicular Crowdsensing (VCS) is a pivotal component in advancing Intelligent Transportation Systems (ITS), facilitating the collection and synthesis of extensive data from distributed vehicular networks. Despite its potential, optimizing participation and data acquisition in VCS is challenged by the self-interested nature of individual participants. In this paper, we propose a novel incentive mechanism for VCS, specifically designed to integrate social benefits within a Vehicle Social Network (VSN). A Small-World (SW) network is introduced to model the VSN, providing a more realistic representation of vehicle interactions and enhancing information propagation. VSN enriches the data utility by sharing data within these networks and acts as a non-monetary incentive that is determined by the strength of connections among participants within the constructed SW networks, sustaining participant engagement even with relatively low monetary rewards. We model the VCS campaign as a Markov Decision Process (MDP) that enables vehicles to independently determine their optimal sensing and communication strategies under the SW networks clustering coefficient K and the rewiring probability p. To maximize individual utility under incomplete information, we introduce a multi-agent Deep Reinforcement Learning (DRL) approach called IM-SJSC that utilizes Variational Autoencoder (VAE) and Proximal Policy Optimization (PPO) for accurate decision-making processes. Simulation results in the T-Drive real-world dataset validate the efficacy of the proposed approach, showing that the average utility outperforms other baseline algorithms by 25.00%, 54.07%, 145.25%, and 181.82% under varying numbers of vehicles. The proposed algorithm also achieves significant performance improvements in other scenarios, such as different numbers of tasks and varying task basic prices.

1. Introduction

Mobile Crowdsensing (MCS) as a promising paradigm [1], utilizes sensors embedded in mobile and connected devices to enhance the quality of smart services [2], revolutionizing traditional information acquisition methods in the Internet of Things (IoT). Compared to stationary sensors, MCS provides greater flexibility and broader coverage.

With the proliferation of connected devices and the rapid development of the Internet of Vehicles (IoV), innovative data collection and analysis methods have paved the way in Intelligent Transportation Systems (ITS). As an application of MCS, Vehicular Crowdsensing (VCS) utilizes IoV as sensing nodes to collect data and has garnered significant attention [3]. Vehicles expand the sensing range and provide timely fine-grained data, gradually becoming a new approach for executing large-scale sensing tasks [4].

Similar to MCS systems, VCS systems adopt a triangular architecture comprising a central cloud platform, requesters, and workers. In this context, IoV devices serve as workers by sensing road and traffic information through embedded sensors and subsequently transmitting the collected data to the central cloud platform via communication modules. After the data are analyzed and processed, the central cloud platform delivers services to requesters, such as optimal route recommendations and traffic congestion warnings [5], thus facilitating efficient and intelligent traffic management and services. Generally, many studies focus on the interaction between the central platform and the workers [6,7], which is also the focus of this paper.

Despite VCS's enormous potential, it still faces substantial challenges, including insufficient participant numbers, lack of enthusiasm among participants, and low-quality data collection. These issues remain significant barriers to the development of VCS. Executing tasks

* Corresponding author.
E-mail addresses: gyluo@stu.xmu.edu.cn (G. Luo), shanhao@stu.xmu.edu.cn (S. Zhan), 23320221154377@stu.xmu.edu.cn (C. Liang), gaozhibin@jmu.edu.cn (Z. Gao), zhaoyf@xmu.edu.cn (Y. Zhao), lfhuang@xmu.edu.cn (L. Huang).

https://doi.org/10.1016/j.comnet.2025.111099
Received 20 October 2024; Received in revised form 11 January 2025; Accepted 2 February 2025
Available online 10 February 2025
1389-1286/© 2025 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

in VCS inevitably consumes resources, especially, frequent access to participating vehicles is required to ensure timely data collection, leading to considerable energy consumption in both data gathering and transmission [8]. Moreover, since some data may compromise user privacy, vehicles are only inclined to participate if sufficient rewards are provided to offset the risks associated with energy consumption and privacy exposure. Without such incentives, vehicles are generally unwilling to participate actively and voluntarily in VCS [9]. Hence, designing effective incentive mechanisms is critical to the successful completion of tasks and to maintaining system performance.

Furthermore, vehicles that frequently travel within the same regions or belong to the same fleet often exhibit strong social ties, leading to more effective information sharing and enhanced efficiency in data collection and transmission. Therefore, incentive mechanisms based on social network have also attracted considerable attention from researchers [10–12]. Yang et al. [10] introduced a social incentive mechanism to recruit more participants, utilizing social connections between participants to offer incentives based on the behaviors of their social friends. This approach motivates participants to influence their friends' behavior. Luo et al. [11] embedded users into a social network of participants, introducing nepotism into participatory crowdsensing. This scheme maximizes the utility for task initiators while ensuring participants' enthusiasm for sensing activities. Yao et al. [12] developed a non-monetary incentive mechanism that capitalizes on the social influence among participants, deeply mining the utility of sensing data while reducing the costs associated with incentives in crowdsensing.

While Vehicular Social Network (VSN) can enhance the performance of IoV, previous incentive works have neglected to study the structure of social networks which may affect information dissemination. The Small-World (SW) networks [13], which lies between regular and random networks, more accurately depicts the realistic social connections among vehicles, offering significant advantages such as shorter path lengths and higher clustering coefficients [14]. These benefits can facilitate the propagation of sensing information within the VSN. The SW networks has been effectively integrated with social networks for the analysis of information diffusion [15]. Besides, the SW networks can serve as a versatile and robust structural paradigm for regulating VSN, thereby optimizing information dissemination and enhancing system resilience within large-scale, dynamic environments. It is demonstrated that the taxi transport network has SW characteristics [16], which also reflects SW relationships between other vehicles. The SW networks enhances the efficient integration of target node data for tasks related to data fusion and processing by reducing path lengths, thus enabling the extraction of more valuable information through fewer nodes and at greater speeds [17]. The Newman-Watts (NW) SW networks model, which combines local connections with random edge additions, is frequently used to simulate real-world network characteristics. Therefore, we explore the NW SW networks model in VCS.

In addition, wireless communication technology enabled the IoV as a bridge that connects vehicle-to-vehicle, vehicle-to-road, vehicle-to-person, and vehicle-to-network, facilitating information sharing and cooperative decision-making. It forms the cornerstone of building ITS. The integration of 6G in IoV aims to optimize and merge communication and sensing resources, breaking away from their current state of isolation [18]. As a promising service paradigm, the fusion of communication and sensing can combine the Integrated Sensing and Communication (ISAC) technology with other technologies [19], further driving the development of ITS. In this framework, deep integration enables the efficient transmission and sharing of sensing information, thus expanding the depth and breadth of sensing capabilities. However, there is limited research on incentive mechanisms that address the joint optimization of sensing and communication.

By considering vehicles' sensing and communication strategies in vehicular incentive mechanisms, the performance of data transmission is expected to improve, thereby enabling vehicles to obtain platform rewards. To our best knowledge, this is the first work to integrate the SW social network with joint sensing and communication strategies within the incentive mechanism for VCS. More specifically, the contributions of this paper can be summarized as follows:

- In VCS, we have devised a social-aware incentive mechanism that accounts for both the reward distribution and the expenses incurred by each vehicle. This novel mechanism facilitates data sharing within a VSN, where connections are modeled by the NW SW networks model. By capturing vehicular connections, the SW network parameters can be dynamically adjusted and affect overall utility.
- The VCS platform dynamically allocates rewards for each time slot based on the deadlines of sensing tasks. We adopt the independent and non-cooperative game to formulate the interactions among vehicles, which do not possess each other's rewards and expenses. By adjusting sensing and communication strategies, vehicles can maximize their own rewards.
- We introduce a multi-agent Deep Reinforcement Learning (DRL) algorithm based on Proximal Policy Optimization (PPO) and Variational Autoencoder (VAE), which derives the sensing and communication strategies for each vehicle under conditions of incomplete information. Simulation results with four baseline methods including DRL-SIM [7], Q-learning [20], Greedy [21], and Random, demonstrate that our proposed algorithm significantly enhances the utility of the vehicles and facilitates the collection of more sensing data.

The remainder of this paper is organized as follows. Section 2 reviews related work on incentive mechanisms in MCS and VCS. We introduce the design of IM-SJSC including its system model and problem formulation in Sections 3 and 4. Subsequently, in Section 5, we present the IM-SJSC algorithm and detail its steps. Then, in Section 6, we conduct several experiments to compare the performance of IM-SJSC with other baselines. Finally, in Section 7, we conclude the paper.

2. Related works

2.1. Traditional incentive mechanism design

In the field of crowdsensing, incentive mechanisms play a crucial role, aiming to stimulate user participation through both monetary and non-monetary rewards. Monetary incentives [22,23], as a direct and common approach, are implemented through payments in cash or financial products, encompassing both platform-centric and participant-centric models. In the former model, the platform sets task prices and selects participants who provide the best value for money. In the latter model, users bid based on their personal costs and expected profits, enhancing both the incentive mechanism's flexibility and user engagement. However, the effective implementation of these mechanisms must consider the diversity and dynamics of tasks, meaning that incentive strategies should be customized for different types of crowdsensing tasks and be adaptable to the changing needs of participants and environmental conditions.

Addressing these issues, Wang et al. [24] proposed an incentive mechanism leveraging blockchain technology to ensure data security and privacy. This mechanism utilizes smart contracts for anonymous authentication and budget-constrained reverse auctions to safeguard vehicle privacy. It further incorporates a fair reward scheme that factors in vehicle reputation and data quality, thereby encouraging the submission of higher-quality data. Yang et al. [25] employed an Edge-assisted Vehicular Crowdsensing (EVCS) framework and a DRL-based incentive mechanism to ensure efficient data collection. Chintakunta et al. [26] developed an incentive mechanism that compensates for the diversity of vehicle locations, encouraging them to deviate from planned routes to maximize the spatial coverage of data collection.

The aforementioned studies overlook the issue in large-scale sensing scenarios, where incentive costs explode with the increase of participants [12]. Although these incentive mechanisms have effectively boosted user participation, their implementation continues to face challenges, including balancing cost-effectiveness, sustainability of incentives, and instability of participation levels.

2.2. Incentive mechanisms based on social networks

The exceptional performance of social networks in the sharing and dissemination of information has been demonstrated by existing studies, showing that participants can gain additional benefits from sensing data through social networks [12]. Zhao et al. [7] proposed a non-cooperative VCS system that implements dynamic pricing tasks and incentive mechanisms based on social network effects. This system effectively enhances participation and maintains long-term incentive effects. Xu et al. [27] developed an incentive mechanism that accounts for data freshness and the social influence among participants. Beyond integrating social benefits into incentive mechanisms as in [7,27], Xu et al. [28] further explored an incentive mechanism in spatial crowdsourcing based on multi-armed bandits and a three-stage Stackelberg game, designing a utility function based on social relationships to achieve systemic optimization based on social benefits. Ji et al. [6] proposed an incentive mechanism for online scenarios targeting social sensing (considering the social relationships among participants) and non-social sensing (not considering social relationships), aiming to maximize platform utility. Li et al. [29] proposed an incentive mechanism based on the Stackelberg game that incorporates social network effects to address user participation, data quality, and communication resource allocation issues in MCS.

Actually, participants engaging in crowdsensing tasks through social networks can gain additional benefits in information sharing, thereby increasing their willingness and improving the efficiency of data collection.

However, existing designs of social-network-based incentive mechanisms often overlook the study of social network structures and the impact of these structures on user benefits and information dissemination.

2.3. Joint sensing and communication in Vehicular Crowdsensing

We further explore the application of integrated sensing and communication in VCS. The joint optimization of communication and sensing offers new solutions for enhancing the real-time availability of sensing data while reducing information processing delays. Zhao et al. [30] established a communication and sensing integration framework that highlights the coordination gains from a joint optimization perspective, laying the foundation for more efficient resource utilization and task execution in sensing. Liu et al. [31] proposed a Lyapunov-guided diffusion-based reinforcement learning framework to jointly optimize DNN partitioning, task offloading, and resource allocation, enabling efficient coordination of communication and computing resources to enhance the real-time processing of sensing data in dynamic vehicular networks. Furthermore, Qi et al. [32] proposed an adaptive method for data transmission and computational instruction offloading. Through collaborative sensing and the distribution of communication and computing resources among devices, their approach reduced transmission resource occupancy and task processing delays.

By aggregating environmental data and vehicle information, safer and more intelligent traffic management can be achieved. Wang et al. [33] integrated sensing and computing to address the challenges associated with collecting and disseminating traffic information in the IoV. They proposed an analysis framework based on complex network theory, along with an information collection architecture and communication model, providing new strategies and theoretical support for data processing and optimized transmission in VCS. Li et al. [34] tackled the joint optimization problem of sensing, processing, and communication in resource-constrained systems, proposing a comprehensive solution framework.

Although these studies offer technical feasibility and strategic guidance for implementing joint communication and sensing in VCS, they do not fully account for the behavior and the connection between participants. Developing effective incentive mechanisms that require substantial resources and spectrum contributions from multiple participants remains an unresolved challenge [35]. The complexity of these tasks demands that incentive mechanisms not only stimulate individual participants' enthusiasm but also promote resource sharing and collaboration to maximize participation and information coverage.

Table 1
Descriptions for variables.

Variables	Descriptions
$\mathcal{N}$	Set of vehicles
$\mathcal{W}$	Set of sensing tasks
m, k	The size of set $\mathcal{N}$ and $\mathcal{W}$
E, e_{ij}	NW models of VSN, social network influence from vehicle j to vehicle i
t, T	Time slot interval and total time
r_0	Basic price of the task
r_n^t	The price of the task at time slot t
t_n^{start}, t_n^{end}	The start time and deadline of the task
x_i^t	Sensing time of vehicle i at time slot t
$\mathbf{x}_{-i}^t$	The set of sensing time of other vehicles except vehicle i at time slot t
p_i^t	Transmission power of vehicle i at time slot t
o_i	Sensing rate of vehicle i
q_i^n	Quality score for vehicle i participating the sensing task n
B_i	Bandwidth of vehicle i
H_i	The channel gain between vehicle i and the MBS
N_0	Background noise
d_i	The distance between vehicle i and the MBS
h	Channel gain factor
θ	Path loss exponent
C_i^t	The total cost of vehicle i at time slot t
α_i	The marginal expense of unit data
c	The cost of unit energy
$T_{i,t}^{trans}$, $T_{i,t}^{total}$	Transmission time and total time at time slot t
Ω_i^t	The utility function of vehicle i at time slot t
U_i^t	The monetary reward for vehicle i at time slot t
ϕ_i^t	The NW social network benefits of vehicle i at time slot t
γ	Discount factor
ω	CLIP factor
ϵ	Random sample from the standard normal distribution
μ_i, σ_i	Distribution parameters

3. System model

As shown in Fig. 1, we consider a VCS consisting of a cloud platform and a group of vehicles. The VCS platform is responsible for publishing sensing tasks essential for targeted applications such as traffic flow monitoring or autonomous driving systems. These tasks are associated with specific durations and rewards, which are adjusted at each time slot $t \in \{1, 2, \dots, T\}$[12]. Vehicles collect data relevant to the assigned sensing tasks by on-board sensors and subsequently transmit the data to the nearest Macro Base Station (MBS) since the coverage of Road Side Unit (RSU) is limited. The MBS is connected with the VCS platform via the fiber optic link, neglecting the wired transmission time compared to the wireless transmission time of the vehicle and the MBS [36]. Upon successful data transmission, vehicles are compensated with rewards for their costs. Meanwhile, vehicles are willing to share and propagate data through VSN to acquire additional benefits [7,37]. Specifically, some key concepts and notations are defined as follows. The variables are organized in Table 1.

Definition 1 (*Sensing Strategy for Vehicles*)**.** The set of vehicles is denoted by $\mathcal{N} = \{1, 2, \dots, i, \dots, m\}$. The sensing strategy of vehicle i refers

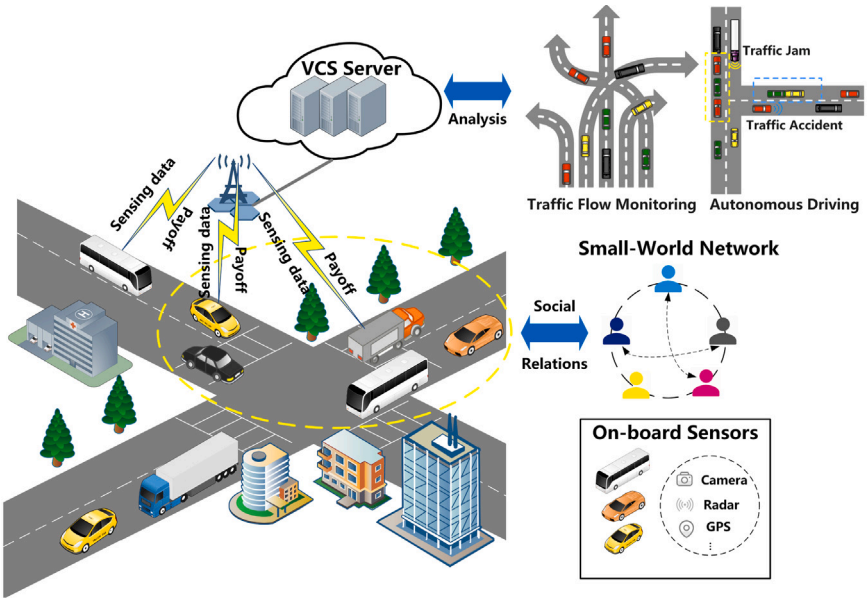


Fig. 1. An illustration of VCS based on social network.

to the sensing time x_i^t that the vehicle contributes in each time slot t. We assume that x_i^t is constrained as $0 \le x_i^t < 1$.

Definition 2 (*Communication Strategy for Vehicles*)**.** The vehicle i determines its transmission power p_i^t in each time slot t, where $0 \le p_i^t \le p_{\max}$. The platform will remunerate the vehicle in each time slot after receiving the sensing data. Therefore, a vehicle must carefully select p_i^t to ensure that the data can arrive in time.

Definition 3 (*Social Networks*)**.** In scenarios where vehicles have strong social connections, the implementation of a VSN model enhances data sharing efficiency and gains social benefits. We utilize the NW SW networks to formulate the intricate social links between vehicles. The NW models of VSN is represented by $E = [e_{ij}]_{m\times m}, e_{ij} \in \{0,1\}$, where e_{ij} indicates the social network influence from vehicle j to vehicle i.

3.1. Task model

The total sensing tasks are represented as $\mathcal{W} = \{1,2,\dots,n,\dots,k\}$. For each task $n \in \mathcal{W}$, $n = \{r_n^t, t_n^{start}, t_n^{end}\}$, where r_n^t is the price of the task at time slot t, t_n^{start} is the start time, and t_n^{end} is the deadline of task w_n. The deadline has a dominant impact on the task requirements. The task will have more significant requirements which can be reflected in r_n^t if the time is closer to the deadline [38]. Therefore, we define the dynamic reward of the task w_n as follows:

$$r_n^t = r_0 \times (1 + F(t, t_n^{end})) \tag{1}$$

where $F(t, t_n^{end}) = 2\arcsin(t/t_n^{end})/\pi$ is the impact function of task w_n's deadline on task requirements, r_0 is the basic price. We define the tasks to be continuous, that is $t_n^{start} = t_{n-1}^{end}$.

3.2. Vehicle model

The vehicle i can be represented by a tuple $n_i = \{E, o_i, x_i^t, p_i^t, q_i^n\} \in \mathcal{N}$, where E is the social network constructed by the NW SW networks model of the vehicles, o_i is the sensing rate of vehicle i, x_i^t and p_i^t is separately the sensing and communication strategy in each time slot t, q_i^n is a quality score for vehicle i participating the sensing task n, which quantifies the quality of sensing data. Data quality is specific to each vehicle, and the platform is aware of the data quality provided by the vehicles. Vehicles transmit sensing data to the MBS, which adopts the orthogonal frequency division multiplexing (OFDM) technique [39]. Thus, the interference is disregarded in the communication model. According to Shannon's theorem, the maximum achievable data rate between vehicle i and the MBS in each time slot t is as follows:

$$\xi_i^t = B_i \log_2(1 + \frac{p_i^t H_i}{N_0}) \tag{2}$$

where B_i is the bandwidth of vehicle i, H_i is the channel gain between vehicle i and the MBS, and N_0 is the background noise. $H_i = d_i^{-\theta}|h|^2$, where d_i is the distance between vehicle i and the MBS, h is the channel gain factor, θ is the path loss exponent [40]. In each time slot t, the total cost of the vehicle i which employ the two strategies x_i^t and p_i^t is as follows:

$$C_i^t = \alpha_i (x_i^t o_i)^2 + \frac{c p_i^t}{\xi_i^t} x_i^t o_i \tag{3}$$

where α_i represents the marginal expense of unit sensing data [41], c is the communication cost of unit energy. After incurring monetary costs, the vehicle i transmits the sensing data to the platform which then compensates the vehicle at each time slot t. Consequently, the duration required for data transmission by the vehicle necessitates careful consideration and is detailed as follows:

$$T_{i,t}^{trans} = \frac{x_i^t o_i}{\xi_i^t} \tag{4}$$

The total time of data sensing and transmission at each time slot t is:

$$T_{i,t}^{total} = \begin{cases} x_i^t + T_{i,t}^{trans} & \text{if } p_i^t \neq 0 \\ \infty & \text{if } p_i^t = 0 \end{cases} \tag{5}$$

3.3. Utility function

We express the utility function of vehicle i at the time slot t as follows:

$$\begin{aligned} \Omega_i^t(x_i^t, \boldsymbol{x}_{-i}^t, p_i^t) &= U_i^t(x_i^t, \boldsymbol{x}_{-i}^t) + \phi_i^t(x_i^t, \boldsymbol{x}_{-i}^t) - C_i^t(x_i^t, p_i^t) \\ &= \frac{q_i^n x_i^t o_i}{\sum_{j=1}^m q_j^n x_j^t o_j} r_n^t + \sum_{j=1}^m e_{ij} x_i^t o_i x_j^t o_j - [\alpha_i (x_i^t o_i)^2 + \frac{c p_i^t}{\xi_i^t} x_i^t o_i] \end{aligned} \tag{6}$$

where the first term $U_i^t(x_i^t, \boldsymbol{x}_{-i}^t) = (q_i^n x_i^t o_i / \sum_{j=1}^m q_j^n x_j^t o_j) r_n^t$ is the monetary reward received by vehicle i at the time slot t which considers the

sensing strategies of other vehicles [7,12], $\boldsymbol{x}_{-i}^t$ is the set which represents the sensing time of other vehicles except vehicle i at the time slot t. Considering the SW characteristic when designing the routing strategies leads to reduced hop counts of sensor data transmission in Wireless Sensor Networks (WSN) [17], we design the second term $\phi_i^t(x_i^t, \boldsymbol{x}_{-i}^t) = \sum_{j=1}^m e_{ij} x_i^t o_i x_j^t o_j$ indicates the social network benefits of NW SW networks model which represents the non-monetary reward. We adopt $e_{ij} x_i^t o_i x_j^t o_j$ to model the social benefits of vehicle i from vehicle j [37,42]. Consequently, the benefits vehicle i received from the entire social network is $\sum_{j=1}^m e_{ij} x_i^t o_i x_j^t o_j$. The last term $C_i^t(x_i^t, p_i^t) = \alpha_i (x_i^t o_i)^2 + (c p_i^t / \xi_i^t) x_i^t o_i$ is the total cost of the vehicle i.

4. Problem formulation

The vehicle determines its sensing and communication strategy in the time slot t as a non-cooperative game, which can be denoted by a tuple $\mathcal{G}^t = \{\mathcal{N}, \varsigma^t, \Omega^t\}$. The game players are the total set of vehicles $\mathcal{N}$. Each vehicle i is rational and maximizes the utility by selecting the optimal strategy group $\varsigma_i^t = \langle x_i^{t^*}, p_i^{t^*} \rangle$. Therefore, the collection of the aggregate strategy group can be represented by $\varsigma^t = \{\varsigma_1^t, \varsigma_2^t, \dots \varsigma_i^t, \dots, \varsigma_m^t\}$. The set of overall utility is $\Omega^t = \{\Omega_1^t, \Omega_2^t, \dots \Omega_i^t, \dots, \Omega_m^t\}$. Concretely, the utility of vehicle i at the time slot t is as follows:

$$\max_{x_i^t, p_i^t} \quad \Omega_i^t(x_i^t, \boldsymbol{x}_{-i}^t, p_i^t), \tag{7a}$$

$$0 \le x_i^t < 1, \forall i \in \mathcal{N}, \tag{7b}$$

$$0 \le p_i^t \le p_{\max}, \forall i \in \mathcal{N}, \tag{7c}$$

$$0 \le T_{i,t}^{total} \le 1, \forall i \in \mathcal{N}. \tag{7d}$$

Constraint (7b) and (7c) is the sensing time and transmission power constraints. (7d) indicates that for a vehicle to receive compensation from the platform, the combined duration of its sensing and communication activities must not exceed the length of a time slot.

5. Proposed solution: IM-SJSC

To address the challenges in our VCS game, several notable obstacles arise. Initially, the NW social network benefit introduces significant complexity to deriving optimal strategies, compounded by the intertwined behaviors of vehicles. Additionally, achieving the optimal sensing and communication strategies becomes difficult with dynamically priced tasks. Furthermore, demanding private information from other vehicles, such as the total cost in Eq. (3).

To overcome these challenges, we adopt the multi-agent DRL methodology that utilizes PPO [43]. In the VCS game, each vehicle becomes an agent to obtain decision sets under incomplete information. Leveraging DRL for optimization offers several advantages, including the ability to efficiently manage high-dimensional state and action spaces, enhance the agents' capacity to learn complex and adaptive strategies, and improve the overall robustness and scalability of the system. Compared with value-based DRL methods, PPO can improve the exploratory of the training process and achieve better convergence performance. We also introduce the VAE [44] as the feature extraction network of the PPO model.

In this section, we conceptualize the VCS game as a multi-agent Markov Decision Process (MDP), where decisions are rendered independently for each vehicle. We also present the construction of NW SW networks model for vehicles in Algorithm 1 and our solution IM-SJSC in Algorithm 2.

5.1. Markov decision process

An MDP is defined by a tuple (S, A, P, R), representing state, action, state transition probability, and reward, respectively. In this framework, each vehicle is an independent agent. At each time slot t, the environment provides a state to the vehicle. The vehicle then selects an action based on its actor network's probability distributions to maximize the total reward, and receives the corresponding reward and next state.

5.1.1. State space

The overall state space can be represented as $S = \{S_i\}$, where S_i is the state of vehicle i. At the time slot t, each vehicle i interacts with the environment to acquire the current state $s_i^t = [\boldsymbol{x}_{-i}^{t-1}, p_i^t, C_i^t, r_n^t] \in S_i$, where $\boldsymbol{x}_{-i}^{t-1}$ is the previous sensing time of other vehicles except vehicle i, p_i^t is the transmission power, C_i^t is the total cost, and r_n^t is the nth task reward.

5.1.2. Action space

The actions of vehicle i in the time slot t consist of the sensing time x_i^t and the transmission power p_i^t. Therefore, the total action space is $A = \{A_i\}$, where $a_i^t = [x_i^t, p_i^t] \in A_i$.

5.1.3. Reward function

The platform dispenses rewards to the vehicles at each time slot t. However, the rewards allocated to the vehicles will be consequently reduced if the platform fails to receive the complete set of sensing data within t. To articulate this, we define the penalty to the utility function as follows:

$$R_i^t = \begin{cases} \Omega_i^t & \text{if } 0 < T_{i,t}^{total} \le 1 \\ 0 & \text{if } p_i^t = 0 \text{ or } T_{i,t}^{total} > 1 \end{cases} \tag{8}$$

where Ω_i^t is the utility function in Eq. (6). In the time slot t, vehicle i will obtain the reward if it not only senses data but also transmits it to the platform within the limits of maximum power $p_{\max}$ and time slot. The vehicle i will not receive any reward if it does not transmit sensing data to the platform or $T_{i,t}^{total}$ is out of a time slot.

5.2. The construction of NW SW model

The NW model uses random edge addition to augment the existing network structure. This distinction not only preserves the SW characteristics, such as high clustering coefficient and short average path length, but also mitigates the risk of generating isolated nodes during the network formation process. This feature is particularly useful for maintaining the robustness and efficiency of information dissemination in social networks.

In VCS environments, the SW model offers a more realistic representation of network connections due to its inherent characteristics, which align with real-world interactions among vehicles [16]. This facilitates efficient and reliable information sharing for crowdsensing applications. In SW networks, each node symbolizes a distinct class. Connections are established between nodes to indicate the presence of a relationship between the classes they represent. This linkage follows the principle that any two classes with a shared attribute or interaction can be interconnected within the network [45]. Therefore, we define the neighboring nodes as vehicles with at least one identical history trajectory on the same road segment. Vehicles are connected to K randomly arranged neighbors in the initialization and subsequently to other vehicles in VSN according to the probability p. The detailed process is shown in Algorithm 1. After the construction process, we obtain the NW social relation matrix of vehicles.

5.3. Proposed IM-SJSC algorithm description

In this section, we first present the overall architecture of our proposed IM-SJSC. Subsequently, we detail our enhancement of the PPO method.

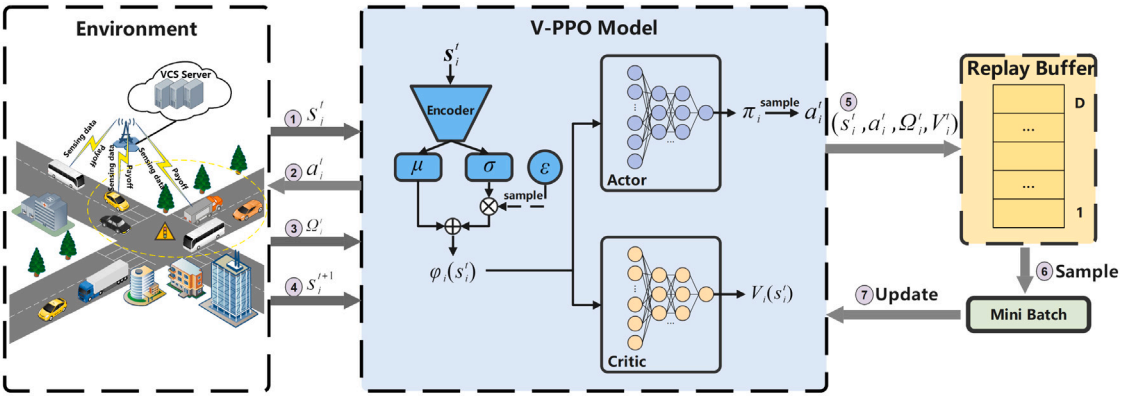


Fig. 2. The architecture of proposed IM-SJSC.

Algorithm 1: The construction of NW SW networks model for Vehicles

Input: Number of vehicle N, even number K, connection probability p
Output: Network $G = (V, E)$

Initialize $G = (V, E)$, $V = \{v_1, v_2, ..., v_N\}$ and $E = \varnothing$;
for *each node* v_i *in* V **do**
  Search neighboring nodes with at least one identical historical trajectory;
  Random order the neighbor nodes;
  for *each neighboring node* v_j *from* 1 *to* K **do**
    Connect v_i to the node v_j with $e_{ij} = 1$;
  end
end
for *each pair of nodes* (v_i, v_j) *in* V **do**
  if $e_{ij} = 0$ *and* $i \neq j$ **then**
    Set $e_{ij} = 1$ with probability p;
  end
end

5.3.1. The overall architecture

As shown in Fig. 2, the actor network of the VAE-PPO (V-PPO) model generates the appropriate actions with probability, and the critic network evaluates and updates the value of the strategy in the training phase. Both the actor and the critic network have the feature extraction network φ_i which calculates features $\varphi_i(s_i^t)$ from state space s_i^t in the time slot t. The actor network π_i derives the probability distribution $\pi_i(\varphi_i(s_i^t))$ of the action by sampling. The vehicle i derives the action with higher probability and will not exceed the maximum sensing time and transmission power with the constraints (7b) and (7c). The critic network V_i computes the value of the action that vehicle i selected, represented as $V_i(s_i^t)$. Meanwhile, V_i updates the actor network π_i based on the value $V_i(s_i^t)$.

5.3.2. The enhancement of the PPO

VAE was originally designed for complex data generation and representation learning, excelling in unsupervised settings where capturing the underlying data distribution is crucial. In contrast, reinforcement learning environments require agents to efficiently interpret high-dimensional state information for decision-making.

When applied within PPO, VAEs provide a structured way to distill state observations into a lower-dimensional latent space, especially as the state space grows with the increasing number of vehicles. VAEs also expand the feature representation of the lower-dimensional state space by increasing the additional possibilities. These characteristics lead to more meaningful state representations than traditional dense layers, which capture superficial patterns without considering the deeper probabilistic structure of the data. Integrating VAEs for state feature extraction network φ_i allows PPO agents to focus on critical features and inherent data uncertainties, potentially enhancing strategy learning and generalization across diverse environmental states. Specifically, We employed a two-layer VAE as the feature extraction network for DRL, with latent space dimensions of 256 and 64, respectively.

5.3.3. The IM-SJSC algorithm description

We introduce the proposed IM-SJSC for each vehicle i in Algorithm 2, which is divided into the action strategy sampling process and the update process.

Initially, the parameters of the feature representation φ_i, the actor network π_i, and the critic network V_i are initialized randomly. At the beginning each episode, the replay buffer D is cleared. The state s_i^0 is set to the current environment state. At each time step t, the state s_i^t is passed through the VAE encoder to obtain the distribution parameters μ_i and σ_i for the latent space. Furthermore, the random sample ε is drawn from a standard normal distribution to ensure the randomness of each sample, as follows:

$$\varphi_i(s_i^t) = \mu_i + \varepsilon\sigma_i \tag{9}$$

This representation is then sampled to form the latent state representation. Subsequently, the actor and critic networks leverage this representation to estimate the value $V_i(s_i^t)$ and determine the action probability distribution $\pi_i(\varphi_i(s_i^t))$, from which the action a_i^t is sampled. Following this, the vehicle participates in sensing tasks with the chosen action, acquires the new state s_i^{t+1} and the associated reward Ω_i^t from the environment, and stores the experience tuple $(s_i^t, a_i^t, \Omega_i^t, v_i^t)$ in the replay buffer D.

The state space is updated, repeating this process for the remaining time slots in the episode. After completing the episode, the algorithm updates the action distribution by sampling mini-batches from the replay buffer D over multiple epochs. Each mini-batch computes the loss for the critic network L_i and the gradients for the actor network J_i. These gradients update both the actor and critic networks. The critic network V_i predicts the cumulative discounted utility of vehicle i from t to T, which is calculated as follows:

$$R_i^t = \Omega_i^t + \gamma\Omega_i^{t+1} + \cdots + \gamma^{T-t+1}\Omega_i^{T-1} + \gamma^{T-t}V_i(s_i^T) \tag{10}$$

where $\gamma \in (0, 1)$ controls the update steps of value estimations. The cumulative discounted utility denotes the expected sum of the reward of vehicle i. Therefore, the loss function of the critic network V_i is as follows:

$$L_i = \hat{E}_t(R_i^t - V_i(s_i^t))^2 \tag{11}$$

where $\hat{E}_t$ denotes the empirical average of finite batches of the sample. The gradient of the critic network V_i is calculated for updating its parameters. To derive the strategy π_i of vehicle i at each time slot t

Algorithm 2: The IM-SJSC for the vehicle i

Input: State space s_i^t
Output: Action space $a_i^t = [x_i^t, p_i^t]$
for *episode in* $1, 2, \ldots$ **do**
  Initialize the environment, obtain current state s_i^0, clear up replay buffer D;
  for t *in* $1, 2, \ldots, T$ **do**
    Encode the current state s_i^t using the VAE encoder φ_i to obtain the distribution parameters μ_i and σ_i for the latent space representation;
    Draw the random sample ϵ from a standard normal distribution;
    Sample latent representation $\varphi_i(s_i^t)$ from distribution by Equation (9);
    Both actor and critic network utilize feature representation $\varphi_i(s_i^t)$ of state s_i^t;
    Get value estimation $V_i(s_i^t)$ from the critic network V_i;
    Get action level probability distribution $\pi_i(\varphi_i(s_i^t))$;
    Sample action a_i^t according to $\pi_i(\varphi_i(s_i^t))$;
    Vehicle i participates in sensing tasks with action a_i^t;
    Acquire the next state s_i^{t+1} and reward Ω_i^t from the environment;
    Adding $(s_i^t, a_i^t, \Omega_i^t, V_i^t)$ into replay buffer D;
    Update the state space;
  end
  Update the action level probability distribution;
  for *epoch in* $1, 2, \ldots$ **do**
    Sample a mini-batch experience from D;
    Calculate the critic network loss L_i by Equation (11);
    Update critic network by gradient ∇L_i;
    Calculate the actor network loss J_i by Equation (12);
    Update actor network by gradient ∇J_i;
  end
end

by probability distributions over a_i^t for each state s_i^t, the actor network utilizes the clip surrogate objective for vehicle i. The calculation of its loss function is as follows:

$$J_i = -\hat{E}_t[\min(\kappa_i^t A_i^t, CLIP(\kappa_i^t, 1-\omega, 1+\omega)A_i^t)] \tag{12}$$

where $\kappa_i^t = \frac{\pi_i(a_i^t|s_i^t)}{\pi_{old}(a_i^t|s_i^t)}$ represents the probability ratio of the update range from the old actor network π_{old} to the new actor network π_i. $A_i^t = R_i^t - V_i(s_i^t)$ is the improvement of the new strategy which estimates the time-step advantage function to quantify the expected reward of adopting particular actions. The CLIP function steadily limits κ_i^t in the range of $[1-\omega, 1+\omega]$, where ω is a hyperparameter. By leveraging the clip function of PPO, our solution can establish a robust strategy through stable and comprehensive updates.

6. Performance evaluation

In this section, we evaluate the performance of our IM-SJSC algorithm. Initially, we describe the dataset used in the experiments, followed by an introduction of the comparison algorithms and evaluation metrics. Subsequently, we train the IM-SJSC, DRL-SIM, and Q-learning methods for 1000 episodes with 100 tests. The Greedy and Random algorithms are also tested for 1000 episodes. Finally, we conduct extensive comparative experiments on a real-world dataset and discuss the evaluation results. Simulation parameter settings are detailed in Table 2.

Table 2
Simulation parameters.

Parameters	Values
Number of vehicles: N	5 – 25
Sensing data rate of vehicle: o_i	10^5 – 10^6 [34]
Sensing cost parameter: α_i	5 – 10 [37]
Bandwidth (MHz): B_i	1 – 5 [37]
Monetary cost per unit energy: c	10 [37]
Noise power spectral density (dBm/Hz): N_0	−95 [34]
Number of tasks: k	5 – 25
Task base price: r_0	45 – 105
Task duration time: $(t_n^{end} - t_n^{start})$	5 – 125
Number of mini-batch: M	6
Mini-batch size:	125 – 300
Discount factor: γ	0.99
CLIP factor: ω	0.1

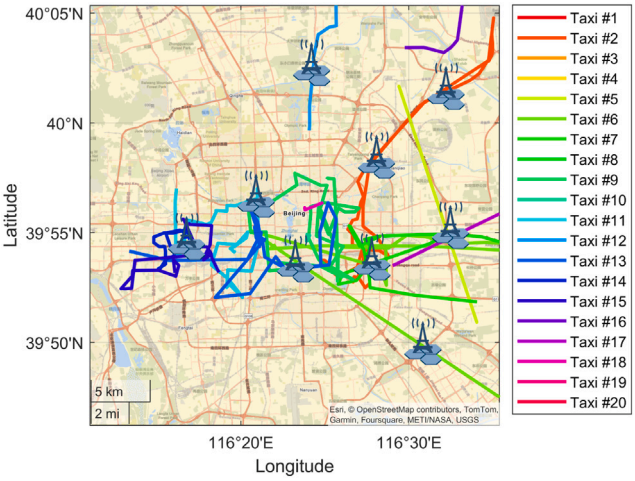


Fig. 3. A illustrated map of trajectory coverage of the Beijing taxi dataset.

6.1. Dataset

In our experiments, we use the T-Drive Beijing vehicle trajectory dataset [46], which contains nearly 15 million trajectory points for 10,357 taxis. To simplify visualization, we show the trajectories of 20 taxis in Fig. 3. We have randomly deployed a number of MBS based on the trajectory information, all connected to the VCS platform. Each vehicle is programmed to select the nearest MBS for data transmission.

6.2. Baselines

6.2.1. DRL-SIM [7]

DRL-SIM utilizes PPO which is an advanced reinforcement learning algorithm gaining considerable attention due to its stability and effectiveness in a wide range of applications. PPO's stability is beneficial in adapting to dynamic and unpredictable environments in VCS. Besides, we introduce PPO as a baseline to show the improvement of applying VAE as the feature extraction network.

6.2.2. Q-learning [20]

This method is widely employed in solving optimization problems in MCS and VCS. It derives an optimal policy through iterative updates of value-action pairs.

6.2.3. Greedy [21]

At each time slot t, vehicle i selects the actions of sensing and communication that maximizes its utility from historical experience.

6.2.4. Random

At each time slot t, vehicle i randomly selects the actions of sensing and communication.

6.3. Evaluation metrics

We introduce seven metrics to evaluate the performance of different algorithms. Firstly, the average utility for each vehicle participating in the sensing tasks is calculated as follows:

$$\Omega_T = \frac{\sum_{t=1}^{T} [\frac{1}{m} \sum_{i=1}^{m} \Omega_i^t(x_i^t, \boldsymbol{x}_{-i}^t, p_i^t)]}{T} \tag{13}$$

To analyze the effect of the components of average utility, we define average monetary and social utility separately [12], as follows:

$$U_M = \frac{\sum_{t=1}^{T} [\frac{1}{m} \sum_{i=1}^{m} U_i^t(x_i^t, \boldsymbol{x}_{-i}^t) - C_i^t(x_i^t, p_i^t)]}{T} \tag{14}$$

$$\phi_S = \frac{\sum_{t=1}^{T} [\frac{1}{m} \sum_{i=1}^{m} \phi_i^t(x_i^t, \boldsymbol{x}_{-i}^t)]}{T} \tag{15}$$

Vehicles receive rewards from the VCS platform and derive utility from the social network after completing tasks that incur costs. Concurrently, the platform anticipates receiving higher-quality data. Consequently, for each vehicle i, we define the average size of sensing data as follows:

$$D_T = \frac{\sum_{t=1}^{T} [\frac{1}{m} \sum_{i=1}^{m} (x_i^t o_i)]}{T} \tag{16}$$

The average completion of sensing tasks is defined as follows:

$$\lambda^n = \begin{cases} 1, \text{if} \sum_{t=t_n^{start}}^{t_n^{end}} \sum_{i=1}^{m} x_i^t o_i \geq \epsilon^n \\ 0, \text{if} \sum_{t=t_n^{start}}^{t_n^{end}} \sum_{i=1}^{m} x_i^t o_i < \epsilon^n \end{cases} \tag{17}$$

where λ^n is a function to measure whether the sum of the sensing data size of all vehicles reaches the minimum requirement ϵ^n of the nth task. If the required data size is satisfied, the task can be completed, otherwise it cannot [7,12]. Therefore, we define the average task completion rate as follows:

$$\eta = \frac{1}{k} \sum_{n=1}^{k} \lambda^n \tag{18}$$

The total average cost incurred after each vehicle makes sensing decision x_i^t and communication decision p_i^t is defined as follows:

$$C_T = \frac{\sum_{t=1}^{T} [\frac{1}{m} \sum_{i=1}^{m} C_i^t(x_i^t, p_i^t)]}{T} \tag{19}$$

The average latency is defined as the time required for vehicles transmitting sensing data to the platform, as follows:

$$L_T = \frac{\sum_{t=1}^{T} [\frac{1}{m} \sum_{i=1}^{m} T_{i,t}^{trans}]}{T} \tag{20}$$

6.4. Utility convergence evaluation

Fig. 4 shows the convergence of the utilities of IM-SJSC, DRL-SIM, Q-learning, Greedy, and Random algorithms with $m = 25$ vehicles, the base price of the tasks $r_0 = 60$, the number of tasks $k = 10$, and the SW VSN characterized by $K = 12$ and $p = 0.3$. We observe that the utility of IM-SJSC converges to around 69 after 210 training rounds and stabilizes. Our IM-SJSC achieves the highest average utility, and the converged average utility for DRL-SIM and Q-learning is higher than that of the Greedy algorithm. This is because Greedy algorithm only considers the optimal solution in the current situation, which often leads to local optimality. In contrast, the Random algorithm generates random actions for each vehicle, yielding the lowest utility. Moreover, our approach outperforms DRL-SIM and Q-learning in average vehicle utility, enhanced by the integration of VAE that effectively extracts features from the state space.

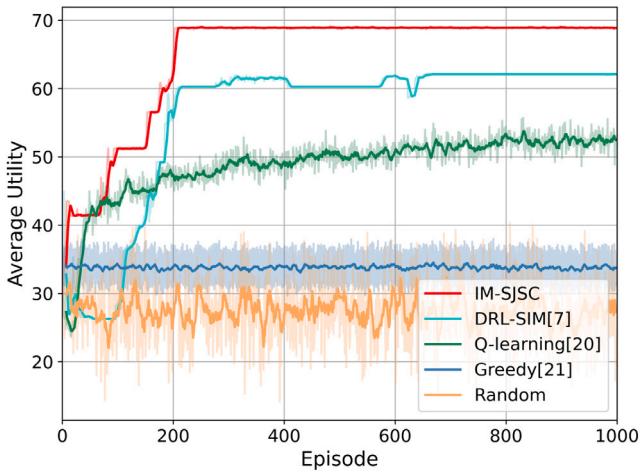
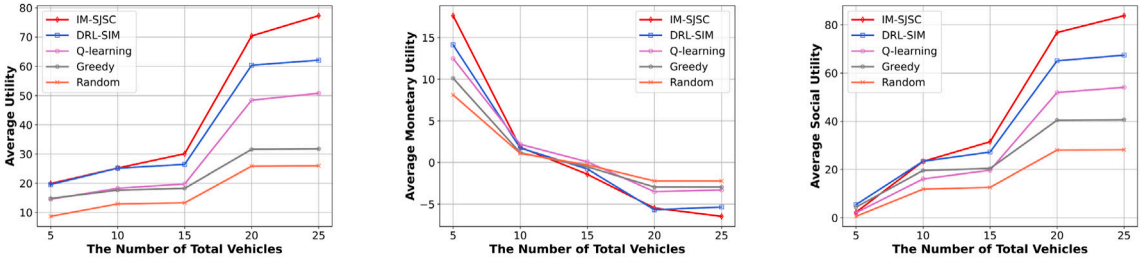


Fig. 4. The convergence of utility.

6.5. Effect of number of vehicles

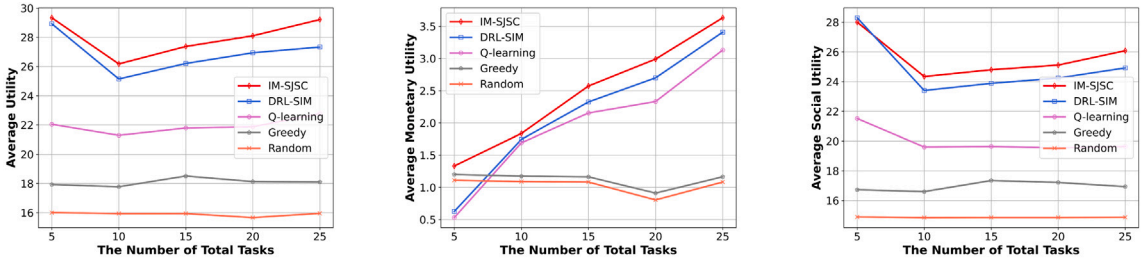
Figs. 5(a), 5(b), and 5(c) demonstrate the impact of varying vehicle numbers on average utility, average monetary utility, and average social utility respectively. We set the task base price $r_0 = 60$, the number of tasks $k = 10$, and the SW VSN with $p = 0.3$. As shown in Fig. 5(a), our IM-SJSC consistently achieves the highest vehicle average utility. For example, when the number of vehicles $m = 25$, IM-SJSC obtains 25.00%, 54.07%, 145.25%, and 181.82% higher average utility than DRL-SIM, Q-learning, Greedy, and Random respectively. This is because the participating vehicles can dynamically adjust sensing time and transmission power, considering factors such as task rewards, sensing and communication costs, and the size of the social network. As the vehicle number varies from 5 to 25, the average vehicle utility increases. Figs. 5(b) and 5(c) reveal that as the number of sensing vehicles rises, the average monetary utility decreases while the social utility increases. For example, in Fig. 5(b), for IM-SJSC, the average monetary utility at $m = 15$ is −1.41, 19.0565 lower than when $m = 5$. In Fig. 5(c), for IM-SJSC, the social utility at $m = 15$ is 31.53, an increase of 29.25 from when $m = 5$. This happens because the platform's rewards are dynamically adjusted based on the deadlines of sensing tasks, but they are not affected by the number of participating vehicles. Moreover, vehicles must also make reasonable decisions about sensing time and transmission power since delayed data transmission leads to discounted rewards. The higher operational costs to ensure timely data transmission reduce the average monetary utility per vehicle. However, as the number of sensing vehicles increases, the size of the VSN expands, improving each vehicle's average social utility and offsetting the reduction in average monetary utility, ensuring an overall increase in the vehicles' average utility.

In some cases, the Q-learning and DRL-based algorithms (IM-SJSC, DRL-SIM) yield lower monetary utilities compared to Greedy and Random algorithms. This is because they aim to maximize the overall average utility of the vehicles, rather than focusing solely on maximizing monetary or social utilities. For instance, at $m = 20$, social utility becomes more beneficial, prompting Q-learning and DRL-based algorithms to design sensing and communication strategies that maximize sensing time and increase transmission power, ensuring data is transmitted on time to the platform. Consequently, vehicles receive lower monetary utility but higher social utility, thereby enhancing the average utility of the vehicles.



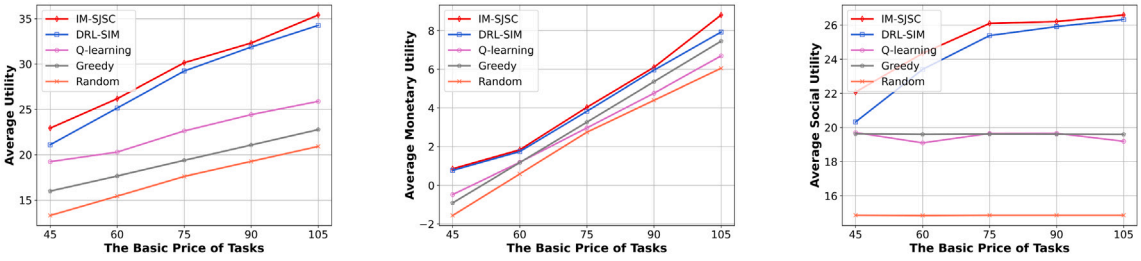
(a) Average Utility **(b)** Average Monetary Utility **(c)** Average Social Utility

Fig. 5. Different utility under variances of m (the number of vehicles). (a) Average Utility; (b) Average monetary utility; (c) Average social utility.



(a) Average Utility **(b)** Average Monetary Utility **(c)** Average Social Utility

Fig. 6. Different utility under variances of k (the number of tasks). (a) Average Utility; (b) Average monetary utility; (c) Average social utility.



(a) Average Utility **(b)** Average Monetary Utility **(c)** Average Social Utility

Fig. 7. Different utility under variances of r_0 (task basic price). (a) Average Utility; (b) Average monetary utility; (c) Average social utility.

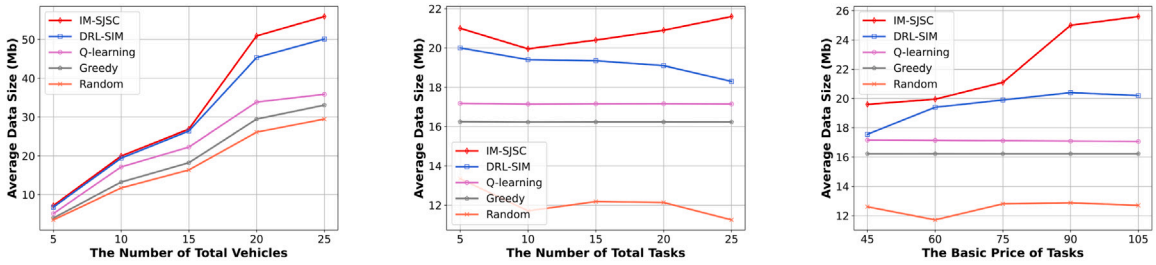
6.6. Effect of number of tasks

Figs. 6(a), 6(b), and 6(c) illustrate the impact of varying task numbers on different utilities. In these experiments, we set the task base price $r_0 = 60$, the number of tasks $m = 10$, and the SW VSN with $K = 4$, $p = 0.3$. As demonstrated in Fig. 6(a), IM-SJSC achieves the highest average utility among the algorithms, demonstrating a more effective sensing and communication strategy. For instance, when $m = 10$, IM-SJSC attains an overall average utility of 0.88, which is 4.11%, 43.13%, 48.26%, and 64.26% higher than DRL-SIM, Q-learning, Greedy, and Random respectively. As the number of tasks increases from 5 to 25, the vehicles' utility first decreases and then increases. In Fig. 6(b), it is observed that the monetary utility generated by Q-learning and DRL-based methods consistently rises, with IM-SJSC generating higher monetary benefits than the other algorithms. This is because of the constant base price of tasks provided by the platform, where an initial rise in task numbers extends the total task duration, increasing overall vehicle costs. Additionally, the platform's dynamic rewards, which were initially insufficiently attractive, result in reduced sensing times by vehicles, as reflected by the lower average social utility in Fig. 6(c). However, as task numbers and deadlines further increase, the dynamic rewards offered by the platform are enhanced, encouraging vehicles to extend their sensing periods. Consequently, an increase in average social utility is observed in Fig. 6(c). Furthermore, the platform's increased pricing compensates adequately for vehicles' expenses, boosting both average monetary utility and overall utility.

As shown in Fig. 6(c), when $k = 5$, IM-SJSC exhibits slightly lower average social utility compared to the DRL-SIM algorithm. This is because vehicles using the IM-SJSC adopt more conservative strategies, such as reducing sensing times and potentially utilizing lower transmission power to meet slot requirements. While this reduces the average social utility, it also lowers vehicle expenses, thereby maintaining a higher average utility.

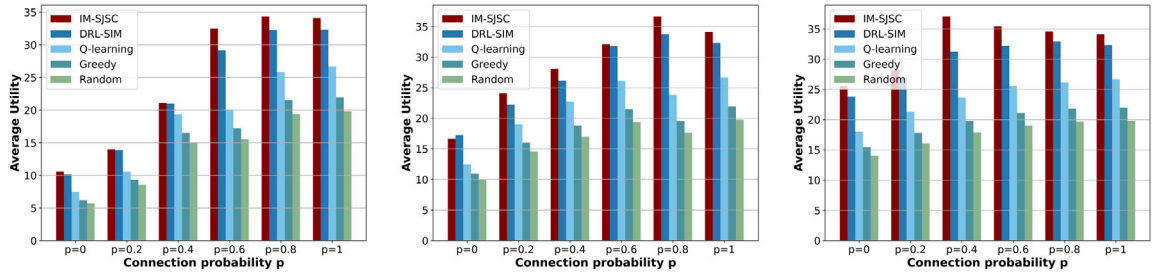
6.7. Effect of basic price of tasks

In our VCS, the rewards provided to vehicles by the platform are dynamically adjusted based on the task slots. Therefore, in Figs. 7(a), 7(b), and 7(c), we examine the impact of varying task base prices on different utility metrics. For these experiments, we set the number of tasks to $m = 10$, the number of tasks to $k = 10$, and the small world vehicular social network with $K = 4$, $p = 0.3$. As shown in Fig. 7(a), the average utility of vehicles increases proportionally with the task base price. For instance, with a task base price of 105, IM-SJSC yields an average utility of 34.40, which is an improvement of 11.47 compared to when the price is 45. Furthermore, IM-SJSC consistently outperforms other baseline algorithms in terms of average utility. For example, at a task base price of 75, the average utility of IM-SJSC



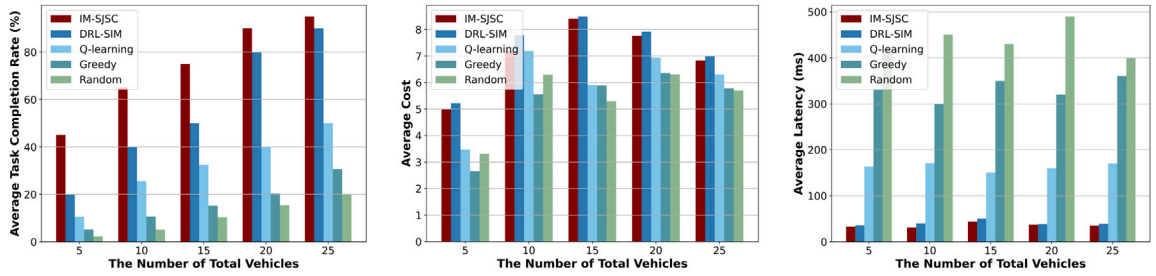
(a) Average Data Size (Mb) **(b)** Average Data Size (Mb) **(c)** Average Data Size (Mb)

Fig. 8. Different average data size under variances of (a) number of vehicles; (b) number of tasks; (c) task basic price.



(a) Average Utility **(b)** Average Utility **(c)** Average Utility

Fig. 9. Different average utility under variances of NW SW networks parameters. (a) $K = 2$; (b) $K = 4$; (c) $K = 6$.



(a) Average Task Completion Rate **(b)** Average Cost **(c)** Average Latency

Fig. 10. Different metrics under variances of vehicles. (a) Average Task Completion Rate; (b) Average Cost; (c) Average Latency.

is 31.54, representing improvements of 7.94%, 39.41%, 62.77%, and 79.09% over DRL-SIM, Q-learning, Greedy, and Random respectively. As shown in Fig. 7(b), the average monetary utility also increases as the task base price rises. Meanwhile, in Fig. 7(c), it is evident that the average social utility for vehicles using the IM-SJSC and DRL-SIM algorithms continuous to rise, while the average social utility for Q-learning, Greedy, and Random remains relatively stable. Overall, IM-SJSC consistently achieves the highest average monetary and social utility for vehicles.

6.8. Variation in sensing data size under different conditions

We have analyzed the variations in the sensing data size with the number of vehicles, the number of tasks, and the task base price. As shown in Fig. 8(a), as the number of vehicles increases, the average size of data transmitted to the platform by the vehicles consistently increases across all methods. Notably, when there are more vehicles, IM-SJSC algorithm motivates the highest average sensing data size among vehicles. For instance, with 25 vehicles, the average data volume for IM-SJSC is 55.90, which is 5.80, 20.05, 22.83, and 26.44 higher than DRL-SIM, Q-learning, Greedy, and Random respectively. In Fig. 8(b), as the number of tasks increases, the data size generated by vehicles using IM-SJSC initially decreases and then increases, which aligns with the observed changes in average vehicle utility in Fig. 6(a) and the average social utility in Fig. 6(c). Moreover, at a certain task number (e.g., $k = 20$), the average data size for IM-SJSC is 20.90, surpassing DRL-SIM, Q-learning, Greedy, and Random by 1.80, 3.74, 4.67, and 8.77 respectively. In Fig. 8(c), with the increase in the task base price, the average data size induced by IM-SJSC continues to rise, with IM-SJSC encouraging vehicles to sense more data compared to other algorithms. For example, at a base price of 105, the average data size for IM-SJSC is 25.6, which is 5.40, 8.53, 9.38, and 12.89 higher than DRL-SIM, Q-learning, Greedy, and Random respectively. Additionally, by comparing Figs. 8(a), 8(b), and 8(c), it is evident that increasing the number of vehicles in the VCS has a more substantial

effect on motivating vehicles to sense more data compared to increasing the number of tasks or the task base price.

6.9. Effect of K and p in NW SW networks

In our VCS, the VSN is constructed using the NW SW networks model. The properties of the SW networks indicate that the social network reflects a random topology where each node is connected to K adjacent nodes when $p = 0$. As p increases, the network's randomness decrease, and at $p = 1$, every node in the social network is interconnected. As shown in Figs. 9(a), 9(b), 9(c), we analyze the variation in the average utility of vehicles under different NW SW networks configurations. We set the number of vehicles to $m = 10$, the number of tasks to $k = 10$, and the base price of tasks to $r_0 = 60$. It is evident that, for the same connection probability $p \in [0, 1)$, as the number of connected nodes K increases, the average utility of the vehicles generally increases, with the IM-SJSC method achieving the highest average utility. Specifically, for IM-SJSC, the peak utility occurs at $p = 0.8$ for both $K = 2$ and $K = 4$, and at $p = 0.4$ for $K = 6$. The DRL-SIM algorithm reaches its maximum utility at $p = 0.8$ for both $K = 4$ and $K = 6$, and at $p = 1$ for $K = 2$. In contrast, the Q-learning, Greedy, and Random algorithms consistently achieve the highest utility at $p = 1$. Given that in real VSNs, it is unlikely that every vehicle has a social connection with all others. Therefore, the highest vehicle utility derived from PPO-based algorithms (IM-SJSC, DRL-SIM) does not depend on establishing social relationships with all vehicles, which aligns better with dynamic and realistic scenarios. Moreover, the VSN constructed through the NW SW networks in conjunction with the IM-SJSC algorithm offers higher utility compared to random or regular network structures. By configuring K and p appropriately, vehicles can maintain the highest average utility in SW VSNs.

6.10. Task completion rate, cost, and latency under variances of vehicles

We analyzed the variations in average task completion rate, cost, and latency under different numbers of vehicles. As shown in Fig. 10(a), the average task completion rate steadily increases as the number of vehicles rises. This enhancement is due to the platform receiving a larger size of sensing data with more vehicles, as illustrated in Fig. 8(a), which consequently boosts the probability of task completion. When there are 20 vehicles, IM-SJSC achieves task completion rates that are 10%, 50%, 70%, and 75% higher than those of DRL-SIM, Q-learning, Greedy, and Random algorithms, respectively. Thus, IM-SJSC effectively meets the data size requirements of sensing tasks, resulting in the highest average task completion rate. Fig. 10(b) shows the variation in average cost. To achieve higher utility, vehicles must allocate more sensing time and increase transmission power to ensure timely data delivery to the platform. Due to the marginal cost per unit of data, cost grows quadratically with the sensing strategy x_i^t, as defined in Eq. (3). Policy-based reinforcement learning methods (IM-SJSC and DRL-SIM) provide higher vehicle utilities compared to other algorithms, leading to increased cost. As the number of vehicles increases, the average cost initially rises and then decreases. This pattern arises because the benefits from the social network grow with more vehicles, as depicted in Fig. 5. The improved utility from the social network allows vehicles to maintain high utility without a proportional increase in sensing time, thereby reducing cost. Our approach ensures maximum vehicle utility while keeping cost at a reasonable level. Fig. 10(c) presents the changes in average delay. It is evident that the average delay across all algorithms remains relatively stable with minor fluctuations. IM-SJSC achieves the lowest average delay. With 20 vehicles, the average delay for IM-SJSC is 36.8 ms, which is 4.1%, 77%, 88.5%, and 92.4% lower than those of DRL-SIM, Q-learning, Greedy, and Random methods, respectively. Other methods experience higher delays because they cannot effectively optimize both sensing and communication strategies.

7. Conclusion

In this paper, we introduce a social-aware incentive mechanism for VCS that enhances the average utility of vehicles. This mechanism motivates vehicles to share data through the VSN, thereby reducing the platform's monetary rewards. Specifically, this paper leverages the social network effect among vehicles as a non-monetary incentive, constructs the VSN as an NW SW networks, and increases the amount of data collected by the vehicles. Furthermore, the IM-SJSC enhances the feature extraction capabilities of the original PPO algorithm by employing the VAE to extract key features from vehicle states, making PPO more suitable for social-aware incentive mechanism. This novel approach allows the derivation of accurate sensing and communication strategies for each vehicle. Extensive experiments demonstrate the superiority of IM-SJSC over four baseline methods. As the number of vehicles increases and reaches 25, the average utility achieved by IM-SJSC surpasses those of DRL-SIM, Q-learning, Greedy, and Random by 25.00%, 54.07%, 145.25%, and 181.82%, respectively. In addition, IM-SJSC yields 11.5%, 55.92%, 69.04%, and 89.75% more data on average than other methods, while also attaining the highest task completion rate. Notably, the superior utility of IM-SJSC does not depend on establishing social relationships with every vehicle, making it more applicable in real-world scenarios.

Since vehicles often operate under limited communication resources, such as restricted bandwidth and power, future research could focus on developing data selection in incentive mechanisms that upload only the most critical and time-sensitive sensing data, rather than transmitting all collected data. Meanwhile, more dynamic channel conditions could be introduced to examine how the potential loss of sensing data during uploads might influence the utility of individual vehicles. Furthermore, the motivation for applying SW networks to construct social networks can also be extended to the flexible regulation of large-scale networks beyond vehicular contexts, thereby enhancing user utility.

CRediT authorship contribution statement

Gaoyu Luo: Writing – review & editing, Writing – original draft, Software, Methodology, Conceptualization. **Shanhao Zhan:** Writing – review & editing, Writing – original draft, Investigation. **Chenyi Liang:** Writing – review & editing. **Zhibin Gao:** Writing – review & editing, Supervision, Resources, Funding acquisition. **Yifeng Zhao:** Supervision, Funding acquisition. **Lianfen Huang:** Supervision, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Gaoyu Luo, Shanhao Zhan, Chenyi Liang, Zhibin Gao, Yifeng Zhao, Lianfen Huang reports financial support was provided by National Natural Science Foundation of China. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant (62371406 and 62171392), in part by the Natural Science Foundation of Xiamen, China (Grant number 3502Z202473053), in part by the Science Technology Project of Fujian under Grant 2024H6030, and in part by the Key Laboratory of Digital Fujian on IoT Communication, Architecture and Security Technology under Grant 2010499.

Data availability

The open source dataset used in this paper has been shared in [46].

References

[1] Y. Li, J. Sun, W. Huang, X. Tian, Detecting anomaly in large-scale network using mobile crowdsourcing, in: IEEE INFOCOM 2019 - IEEE Conference on Computer Communications, 2019, pp. 2179–2187, http://dx.doi.org/10.1109/INFOCOM.2019.8737541.

[2] L. XU, S. YU, Z. FENG, F. YIN, Trustworthy and efficient data trading in decentralized mobile crowd sensing systems, J. Donghua Univ.(Engl. Ed.) 41 (01) (2024) 89–101, http://dx.doi.org/10.19884/j.1672-5220.202301004.

[3] S. Liu, Z. Zheng, F. Wu, S. Tang, G. Chen, Context-aware data quality estimation in mobile crowdsensing, in: IEEE INFOCOM 2017 - IEEE Conference on Computer Communications, 2017, pp. 1–9, http://dx.doi.org/10.1109/INFOCOM.2017.8057033.

[4] S. Guo, X. Qian, Optimal drive-by sensing in urban road networks with large-scale ridesourcing vehicles, IEEE Trans. Intell. Transp. Syst. 25 (10) (2024) 14389–14400, http://dx.doi.org/10.1109/TITS.2024.3394748.

[5] S. Abdul Rahman, A. Mourad, M. El Barachi, An infrastructure-assisted crowdsensing approach for on-demand traffic condition estimation, IEEE Access 7 (2019) 163323–163340, http://dx.doi.org/10.1109/ACCESS.2019.2953002.

[6] G. Ji, B. Zhang, G. Zhang, C. Li, Online incentive mechanisms for socially-aware and socially-unaware mobile crowdsensing, IEEE Trans. Mob. Comput. 23 (5) (2024) 6227–6242, http://dx.doi.org/10.1109/TMC.2023.3321701.

[7] Y. Zhao, C.H. Liu, Social-aware incentive mechanism for vehicular crowdsensing by deep reinforcement learning, IEEE Trans. Intell. Transp. Syst. 22 (4) (2021) 2314–2325, http://dx.doi.org/10.1109/TITS.2020.3014263.

[8] J. Ma, Y. Yu, G. Liu, T. Huang, Freshness aware vehicular crowdsensing with multi-agent reinforcement learning, Comput. Netw. 257 (2025) 110978, http://dx.doi.org/10.1016/j.comnet.2024.110978.

[9] C. Hu, M. Xiao, L. Huang, G. Gao, Truthful incentive mechanism for vehicle-based nondeterministic crowdsensing, in: 2016 IEEE/ACM 24th International Symposium on Quality of Service, IWQoS, 2016, pp. 1–10, http://dx.doi.org/10.1109/IWQoS.2016.7590452.

[10] G. Yang, S. He, Z. Shi, J. Chen, Promoting cooperation by the social incentive mechanism in mobile crowdsensing, IEEE Commun. Mag. 55 (3) (2017) 86–92, http://dx.doi.org/10.1109/MCOM.2017.1600690CM.

[11] T. Luo, S.S. Kanhere, H.-P. Tan, SEW-ing a simple endorsement web to incentivize trustworthy participatory sensing, in: 2014 Eleventh Annual IEEE International Conference on Sensing, Communication, and Networking, SECON, 2014, pp. 636–644, http://dx.doi.org/10.1109/SAHCN.2014.6990404.

[12] X.-W. Yao, X.-T. Yang, Q. Li, C.-F. Qi, X.-J. Kong, X.-Y. Li, UMIM: Utility-maximization incentive mechanism for mobile crowd sensing, IEEE Trans. Mob. Comput. 23 (5) (2024) 6334–6346, http://dx.doi.org/10.1109/TMC.2023.3320106.

[13] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world'networks, Nature 393 (6684) (1998) 440–442, http://dx.doi.org/10.1038/30918.

[14] A. Helmy, Small worlds in wireless networks, IEEE Commun. Lett. 7 (10) (2003) 490–492, http://dx.doi.org/10.1109/LCOMM.2003.818887.

[15] Y. Hu, Z. Wang, X. Li, Impact of policies on electric vehicle diffusion: An evolutionary game of small world network analysis, J. Clean. Prod. 265 (2020) 121703, http://dx.doi.org/10.1016/j.jclepro.2020.121703.

[16] Y. Yang, Z. He, Z. Song, X. Fu, J. Wang, Investigation on structural and spatial characteristics of taxi trip trajectory network in Xi'an, China, Phys. A 506 (2018) 755–766, http://dx.doi.org/10.1016/j.physa.2018.04.096.

[17] O.J. Pandey, A. Mahajan, R.M. Hegde, Joint localization and data gathering over a small-world WSN with optimal data MULE allocation, IEEE Trans. Veh. Technol. 67 (7) (2018) 6518–6532, http://dx.doi.org/10.1109/TVT.2018.2805921.

[18] Q. Qi, X. Chen, A. Khalili, C. Zhong, Z. Zhang, D.W.K. Ng, Integrating sensing, computing, and communication in 6G wireless networks: Design and optimization, IEEE Trans. Commun. 70 (9) (2022) 6212–6227, http://dx.doi.org/10.1109/TCOMM.2022.3190363.

[19] Y. Cui, F. Liu, X. Jing, J. Mu, Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges, IEEE Netw. 35 (5) (2021) 158–167, http://dx.doi.org/10.1109/MNET.010.2100152.

[20] L. Xiao, T. Chen, C. Xie, H. Dai, H.V. Poor, Mobile crowdsensing games in vehicular networks, IEEE Trans. Veh. Technol. 67 (2) (2018) 1535–1545, http://dx.doi.org/10.1109/TVT.2016.2647624.

[21] H. Zhao, M. Xiao, J. Wu, Y. Xu, H. Huang, S. Zhang, Differentially private unknown worker recruitment for mobile crowdsensing using multi-armed bandits, IEEE Trans. Mob. Comput. 20 (9) (2021) 2779–2794, http://dx.doi.org/10.1109/TMC.2020.2990221.

[22] H. Zhang, B. Liu, H. Susanto, G. Xue, T. Sun, Incentive mechanism for proximity-based mobile crowd service systems, in: IEEE INFOCOM 2016 - the 35th Annual IEEE International Conference on Computer Communications, 2016, pp. 1–9, http://dx.doi.org/10.1109/INFOCOM.2016.7524549.

[23] W. Wu, W. Wang, M. Li, J. Wang, X. Fang, Y. Jiang, J. Luo, Incentive mechanism design to meet task criteria in crowdsourcing: How to determine your budget, IEEE J. Sel. Areas Commun. 35 (2) (2017) 502–516, http://dx.doi.org/10.1109/JSAC.2017.2659278.

[24] L. Wang, Z. Cao, P. Zhou, X. Zhao, Towards a smart privacy-preserving incentive mechanism for vehicular crowd sensing, Secur. Commun. Netw. 2021 (2021) 1–16, http://dx.doi.org/10.1155/2021/5580089.

[25] X. Yang, B. Gu, Toward incentive-compatible vehicular crowdsensing: A reinforcement learning-based approach, in: Deep Learning and Its Applications for Vehicle Networks, CRC Press, 2023, pp. 259–272, http://dx.doi.org/10.1109/MNET.104.2000773.

[26] H. Chintakunta, X. Wang, L.G. Jaimes, Improving sensing coverage in vehicular crowdsensing using location diversity, in: 2022 International Conference on Connected Vehicle and Expo, ICCVE, 2022, pp. 1–6, http://dx.doi.org/10.1109/ICCVE52871.2022.9742961.

[27] Y. Xu, M. Xiao, Y. Zhu, J. Wu, S. Zhang, J. Zhou, AoI-Guaranteed incentive mechanism for mobile crowdsensing with freshness concerns, IEEE Trans. Mob. Comput. 23 (5) (2024) 4107–4125, http://dx.doi.org/10.1109/TMC.2023.3285779.

[28] Y. Xu, M. Xiao, J. Wu, S. Zhang, G. Gao, Incentive mechanism for spatial crowdsourcing with unknown social-aware workers: A three-stage stackelberg game approach, IEEE Trans. Mob. Comput. 22 (8) (2023) 4698–4713, http://dx.doi.org/10.1109/TMC.2022.3157687.

[29] M. Li, M. Ma, L. Wang, B. Yang, Quality-improved and delay-aware incentive mechanism for mobile crowdsensing with social concerns: A stackelberg game approach, IEEE Trans. Comput. Soc. Syst. 11 (6) (2024) 7618–7633, http://dx.doi.org/10.1109/TCSS.2024.3430396.

[30] L. Zhao, D. Wu, L. Zhou, Y. Qian, Radio resource allocation for integrated sensing, communication, and computation networks, IEEE Trans. Wirel. Commun. 21 (10) (2022) 8675–8687, http://dx.doi.org/10.1109/TWC.2022.3168348.

[31] Z. Liu, H. Du, J. Lin, Z. Gao, L. Huang, S. Hosseinalipour, D. Niyato, DNN partitioning, task offloading, and resource allocation in dynamic vehicular networks: A Lyapunov-guided diffusion-based reinforcement learning approach, IEEE Trans. Mob. Comput. (2024) 1–17, http://dx.doi.org/10.1109/TMC.2024.3486728.

[32] Y. Qi, Y. Zhou, Y.-F. Liu, L. Liu, Z. Pan, Traffic-aware task offloading based on convergence of communication and sensing in vehicular edge computing, IEEE Internet Things J. 8 (24) (2021) 17762–17777, http://dx.doi.org/10.1109/JIOT.2021.3083065.

[33] J. Wang, C. Jiang, Z. Han, Y. Ren, L. Hanzo, Internet of vehicles: Sensing-aided transportation information collection and diffusion, IEEE Trans. Veh. Technol. 67 (5) (2018) 3813–3825, http://dx.doi.org/10.1109/TVT.2018.2796443.

[34] X. Li, G. Feng, Y. Liu, S. Qin, Z. Zhang, Joint sensing, communication, and computation in mobile crowdsensing enabled edge networks, IEEE Trans. Wirel. Commun. 22 (4) (2023) 2818–2832, http://dx.doi.org/10.1109/TWC.2022.3214535.

[35] P. Singh, B. Hazarika, K. Singh, W.-J. Huang, C.-P. Li, Augmented multiagent DRL for multi-incentive task prioritization in vehicular crowdsensing, IEEE Internet Things J. 11 (22) (2024) 36140–36156, http://dx.doi.org/10.1109/JIOT.2024.3446954.

[36] J. Huang, J. Wan, B. Lv, Q. Ye, Y. Chen, Joint computation offloading and resource allocation for edge-cloud collaboration in internet of vehicles via deep reinforcement learning, IEEE Syst. J. 17 (2) (2023) 2500–2511, http://dx.doi.org/10.1109/JSYST.2023.3249217.

[37] B. Li, K. Xie, X. Huang, Y. Wu, S. Xie, Deep reinforcement learning based incentive mechanism design for platoon autonomous driving with social effect, IEEE Trans. Veh. Technol. 71 (7) (2022) 7719–7729, http://dx.doi.org/10.1109/TVT.2022.3164656.

[38] X. Ma, Y. Huang, Q. Li, B. Huang, X. Liu, J. Liu, Dynamic incentive for reliable MCS participant selection, IEEE Internet Things J. 10 (18) (2023) 15912–15922, http://dx.doi.org/10.1109/JIOT.2023.3266022.

[39] H. Zhu, J. Wang, Chunk-based resource allocation in OFDMA systems - part I: chunk allocation, IEEE Trans. Commun. 57 (9) (2009) 2734–2744, http://dx.doi.org/10.1109/TCOMM.2009.09.080067.

[40] Y. Wang, P. Lang, D. Tian, J. Zhou, X. Duan, Y. Cao, D. Zhao, A game-based computation offloading method in vehicular multiaccess edge computing networks, IEEE Internet Things J. 7 (6) (2020) 4987–4996, http://dx.doi.org/10.1109/JIOT.2020.2972061.

[41] M.H. Cheung, F. Hou, J. Huang, Make a difference: Diversity-driven social mobile crowdsensing, in: IEEE INFOCOM 2017 - IEEE Conference on Computer Communications, 2017, pp. 1–9, http://dx.doi.org/10.1109/INFOCOM.2017.8057035.

[42] Z. Xiong, D. Niyato, P. Wang, Z. Han, Y. Zhang, Dynamic pricing for revenue maximization in mobile social data market with network effects, IEEE Trans. Wirel. Commun. 19 (3) (2020) 1722–1737, http://dx.doi.org/10.1109/TWC.2019.2957092.

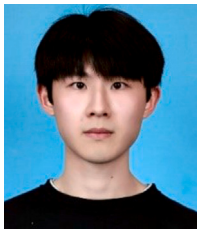
[43] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, Proximal policy optimization algorithms, 2017, http://dx.doi.org/10.48550/arXiv.1707.06347, arXiv preprint arXiv:1707.06347.

[44] D.P. Kingma, M. Welling, Auto-encoding variational bayes, 2013, http://dx.doi.org/10.48550/arXiv.1312.6114, arXiv preprint arXiv:1312.6114.

[45] X.F. Wang, G. Chen, Complex networks: small-world, scale-free and beyond, IEEE Circuits Syst. Mag. 3 (1) (2003) 6–20, http://dx.doi.org/10.1109/MCAS.2003.1228503.

[46] J. Yuan, Y. Zheng, C. Zhang, W. Xie, X. Xie, G. Sun, Y. Huang, T-drive: driving directions based on taxi trajectories, in: Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2010, pp. 99–108, http://dx.doi.org/10.1145/1869790.1869807.

Gaoyu Luo received the B.S. degree in electronic and information engineering from Fuzhou University, Fuzhou, China in 2022. He is currently working towards the M.S. degree in signal and information processing with Xiamen University, Xiamen, China. His research interests include Internet of Vehicles, wireless communication, wireless network resource management, mobile crowdsensing, deep reinforcement learning, and deep learning.



Shanhao Zhan received the B.S. degree in Communication Engineering from the University of Electronic Science and Technology of China in 2022. He is currently pursuing a M.S. degree in Communication Engineering at Xiamen University in China. His research interests include the Internet of Vehicles, ISAC, resource management in 6G networks, graph representation learning, graph reinforcement learning, and multi-agent reinforcement learning.

Chenyi Liang received the B.S. degree in Mechanical Engineering and Automation from School of Mechatronics Engineering, Harbin Institute of Technology in 2022. He is currently pursuing his M.S. degree in Signal and Information Processing at School of Informatics, Xiamen University. His research interests include wireless communications, mobile edge computing, crowdsensing and Internet of Vehicle (IoV).

Zhibin Gao received the B.S. degree in communication engineering, the M.S. degree in radio physics, and the Ph.D. degree in communication engineering from Xiamen University, Xiamen, China, in 2003, 2006, and 2011, respectively. He is a professor of Navigation Institute, Jimei University, Xiamen, China. Previously, he worked with Xiamen University, Xiamen, China, as a senior engineer of Communication Engineering and he was a visiting scholar with University of Washington from 2016 to 2017. His research interests include Internet of Vehicles, marine communications, wireless network resource management, and signal processing.

Yifeng Zhao received his B.S. degree in Communication Engineering in 2002, M.S degree in electronic circuit system in 2005 and Ph.D. degree in Communication Engineering in 2014 from Xiamen University. He is an assistant professor of Communication Engineering, Xiamen University, Xiamen, Fujian, China. His current research interests include Mmwave communication, Massive MIMO and Machine learning applied in wireless communication.

Lianfen Huang received the B.S. degree in radio physics and the Ph.D. degree in communication engineering from Xiamen University, Xiamen, China, in 1984 and 2008, respectively. She was a Visiting Scholar with Tsinghua University, Beijing, China, in 1997. She is currently a Professor with the Department of Communication Engineering, Xiamen University. Her current research interests include wireless communication, wireless networks, and signal process.