

# Mobile Crowdsensing Games in Vehicular Networks

Liang Xiao\*, Tianhua Chen\*, Caixia Xie\*, Huaiyu Dai<sup>†</sup>, H. Vincent Poor<sup>‡</sup>

\*Dept. CE, Xiamen University, Xiamen, China. Email: lxiao@xmu.edu.cn

<sup>†</sup>Dept. ECE, North Carolina State University, Raleigh, USA. Email: Huaiyu\_Dai@ncsu.edu

<sup>‡</sup>Dept. EE, Princeton University, Princeton, USA. Email: poor@princeton.edu

**Abstract**—Vehicular crowdsensing takes advantage of the mobility of vehicles to provide location-based services in large-scale areas. In this paper, we analyze the mobile crowdsensing (MCS) in vehicular networks and formulate the interactions between a crowdsensing server and vehicles equipped with sensors in the area of interest as a vehicular crowdsensing game. Each participating vehicle chooses its sensing strategy based on the sensing cost, radio channel state, and the expected payment. The MCS server evaluates the accuracy of each sensing report and pays the vehicle accordingly. The Nash equilibrium of the static vehicular crowdsensing game is derived for both accumulative sensing tasks and best-quality sensing tasks, showing the tradeoff between the sensing accuracy and the overall payment by the MCS server. We propose a Q-learning based MCS payment strategy and sensing strategy for the dynamic vehicular crowdsensing game, and apply the post decision state learning technique to exploit the known radio channel model to accelerate the learning speed of each vehicle. Simulations based on the Markov-chain channel model are performed to verify the efficiency of the proposed mobile crowdsensing system, showing that it exceeds the benchmark MCS system in terms of the average utility, the sensing accuracy and the energy consumption of the vehicles.

**Index Terms**—Vehicular networks, mobile crowdsensing, game theory, reinforcement learning.

## I. INTRODUCTION

With the booming developments of vehicles and smartphones equipped with sensors, such as cameras, microphones, GPS devices and digital compasses, mobile crowdsensing (MCS) over vehicular networks can provide location-based services, such as urban monitoring [1], road and traffic condition monitoring [2], [3], pollution levels measurements, wildlife habitat monitoring [4], and cross-space public information sharing [5]. Mobile crowdsensing systems consist of servers and participant vehicles that use their embedded sensors to gather the information requested by the servers. Due to the mobility of vehicles, vehicular crowdsensing can improve the sensing coverage of location-based services [6].

As shown in [7], the MCS server pays each vehicle according to its sensing quality to stimulate vehicles to provide full sensing efforts. More specifically, a vehicle is motivated to send an accurate sensing report, if its expected payment from the MCS server exceeds its sensing cost due to the energy and time spent to sense and transmit the report. Game theory has been used to investigate the interactions between participants and auction-based servers in mobile crowdsensing [8]. Auctions and pricing strategies such as [9], [10] stimulate mobile users to participate in crowdsensing applications. The discriminating payment strategies proposed in [11] can address faked sensing attacks in mobile crowdsensing.

However, to our best knowledge, most existing works ignore the impact of the type of the sensing tasks. For the accumulative sensing tasks, such as the metropolitan traffic mapping applications as investigated in [2], the MCS server benefits from receiving a large number of sensing reports, even if some reports are not very accurate. On the other hand, for the best-quality sensing tasks, such as the photo-based applications in a specific location as in [5], the MCS server does not gain from receiving more inaccurate sensing reports. Moreover, in contrast to the static network assumptions in most game theoretic studies of MCS, the vehicular mobility makes it challenging for a vehicle to accurately estimate the sensing environment and the responses of the other vehicles in time.

In this paper, we investigate the mobile crowdsensing in vehicular networks for both the accumulative sensing tasks and best-quality sensing tasks, and formulate a vehicular crowdsensing game, in which each vehicle chooses its sensing effort such as the sensing time and energy. The server applies classification techniques such as [5] and [12] to evaluate the quality of each received sensing report, and pays the corresponding vehicle according to its sensing effort. We derive the Nash equilibria (NEs) of the static vehicular crowdsensing games and provide the condition that the NE exists, in terms of the number of vehicles in the target area, sensing contributions, sensing costs, radio channel gains, and the type of the sensing tasks. The NE of the game provides a tradeoff between the sensing accuracy and the overall payment by the server.

Dynamic vehicular crowdsensing games are also investigated, in which the MCS server can not accurately observe the sensing costs of the vehicles and their radio channel conditions in time. Reinforcement learning techniques, such as Q-learning and post decision state (PDS) learning [13], are applied to achieve the optimal payment and sensing strategies via trial-and-error. The sensing strategies with PDS-learning exploit the partially known radio channel model to accelerate the convergence rate, and thus further improves the learning speed and increases the utility of the vehicle.

Compared with existing works, our major contributions include:

- (1) We formulate the static vehicular MCS game and derive its NE for both the accumulative and best-quality sensing tasks for vehicular networks **with up to two nonzero sensing levels**.
- (2) We investigate the dynamic MCS game, propose the Q-learning based payment and sensing strategies for the MCS system and apply PDS-learning technique to accelerate the convergence of vehicles in MCS.
- (3) **Simulations are performed** to evaluate the performance

of the dynamic vehicular MCS game for up to 10 nonzero sensing levels based on Markov-chain channel model, in which the state transition probability increases with the vehicle speed.

The rest of the paper is organized as follows. We review related work in Section II, and present the system model in Section III. We formulate the vehicular crowdsensing game in Section IV, investigate the NE of the static game in Section V. We study the dynamic MCS game and present the learning-based crowdsensing strategies in Section VI. Simulation results are presented in Section VII, and conclusions are drawn in Section VIII.

## II. RELATED WORK

The vehicular information transfer protocol in [14] provides location-based, traffic-oriented services in distributed ad-hoc networks. In [2], vehicles equipped with sensors monitor road surface conditions and exploit vehicular mobility to opportunistically gather vibration data. The proactive urban monitoring system in [1] applies vehicles to opportunistically diffuse the summaries of sensed data. The interactive music recommendation system developed in [15] applies mobile crowdsensing to suggest music to drivers and can relieve their fatigue and negative emotion. Vehicular mobile surveillance examined in [16] motivates vehicles to collect data and share bandwidth. The vehicle platooning system described in [17] controls its topology based on the network status. Participant recruitment based on the predicted trajectory was developed in [18] for vehicle-based crowdsourcing. TV white space networks are utilized in [19] to provide robust and long range connectivity for vehicles against power asymmetry.

Mobile crowdsensing systems based on smartphones can support a wide range of large-scale monitoring applications. Multidimensional context-aware social networks discussed in [20] enable context awareness to develop mobile crowdsourcing applications. The available individual reputation information was exploited in [21] to improve the performance of crowdsourcing. The cooperation and competition among the service providers were studied in [22] to develop an incentive mechanism for MCS. The participatory application in [23] is resilient to the newcomer, on-off and collusion attacks. The crowdsourcing system developed in [24] can enhance disaster management. An efficient and privacy-aware data aggregation was proposed in [25] to aggregate time-series data without leaking mobile users' privacy. In the cloud-assisted image sharing system in [26], smartphones evaluate sensed images based on their significance and redundancy, and only upload valuable and unique images under disaster environments.

The Stackelberg equilibrium was derived in [10] for the MCS game, in which the utility of each participant is known by the MCS platform. A Bayesian MCS game was formulated with unknown participation efforts in [27]. The interactions between the service user and MCS provider were formulated as in [28] an online double auction with two dynamic users. The online auction with users arriving in sequence analyzed in [29] maximizes the service value from the selected users under a budget constraint. The Nash bargaining solution of MCS is investigated in [30], in which the MCS platform

bargains with each mobile user independently. The auction-based indoor localization MCS system in [9] pays each mobile user based on the quality of its sensed data.

In [7], we formulated a vehicular crowdsensing game consisting of an MCS server and some participating vehicles in static networks. In this paper, we extend the study to dynamic environments, based on the Markov-chain channel model and discuss the types of MCS applications. We also propose a sensing strategy based on PDS-learning to improve the Q-learning based system in [7], and perform simulations in dynamic games to evaluate their performance.

## III. SYSTEM MODEL

### A. Network Model

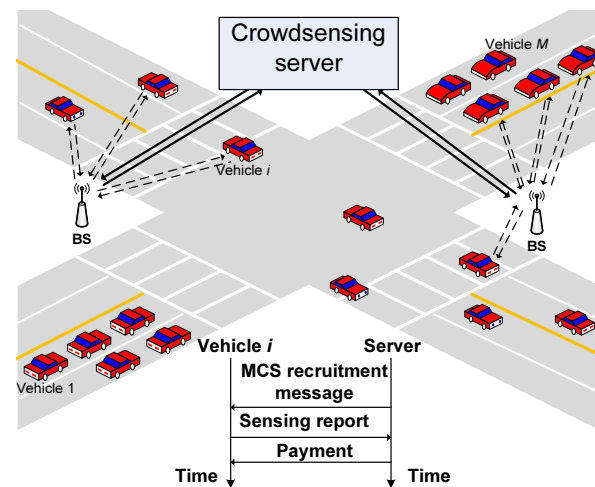


Fig. 1. Illustration of a vehicular crowdsensing system consisting of  $M$  vehicles that connect to the crowdsensing server via base stations.

We consider a vehicular network consisting of  $M$  vehicles equipped with sensors, such as cameras and GPSs as shown in Fig. 1. The MCS server aims to collect certain location-based information and broadcasts its sensing tasks to the vehicles, including its payment range for each participating vehicle. The vehicles use their sensors to gather the requested local information and transmit their sensing data to the server via the serving base stations (BSs). For simplicity, the MCS payment is quantized into  $N_b + 1$  levels, and the action set of the MCS server is denoted by  $\mathbf{B} = \{b_j\}_{0 \leq j \leq N_b}$ , where  $b_j$  is the payment at level  $j$ , with  $b_m < b_n, \forall 0 \leq m < n \leq N_b$ .

Upon receiving the recruiting message from the server, vehicle  $i$  determines whether or not to participate in the service and its effort to perform the sensing task, denoted by  $x_i$ . For simplicity, the sensing effort of the vehicle is quantized into  $N_a + 1$  levels, and  $x_i \in \mathbf{A} = \{0, 1, 2, \dots, N_a\}$ . For example, the vehicle does not participate in the MCS service if  $x_i = 0$ , while it applies its full sensing effort if  $x_i = N_a$ .

The server applies the evaluation algorithm such as [9] to evaluate the accuracy of the sensing report and thus estimate the sensing effort of the corresponding mobile device. Based on an accurate evaluation algorithm, the server is assumed to know action  $x_i$ . A vehicle offered more sensing effort deserves

a higher payment from the MCS server. The payment to a vehicle with sensing effort  $j$ , with  $0 \leq j \leq N_a$ , is denoted by  $y_j \in \mathbf{B}$ . The discriminant payment strategy of the server is given by  $\mathbf{y} = [y_j]_{0 \leq j \leq N_a}$ . This crowdsensing system provides a tradeoff between the cost of the server and the sensing quality, instead of optimizing one of them. The payment of the server depends on the expected number of sensing reports, the channel states, contributions and sensing costs of the vehicles. If a server has a higher budget, it pays more to stimulate more vehicles to take the MCS task.

### B. Channel Model

The radio channel state between a vehicle and its serving BS depends on its mobility. For simplicity, the radio channel state of vehicle  $i$  that determines the signal-to-noise ratio (SNR) of the signal received by the BS is quantized into  $N_h + 1$  levels, and is denoted by  $h_i^k$ , where  $k$  is the time index. We assume that  $h_i^k \in \{H_j\}_{0 \leq j \leq N_h}$ , with  $H_m < H_n, \forall 0 \leq m < n \leq N_h$ . The channel state is constant during the transmission of a sensing report, and changes afterwards.

As shown in Fig. 2, we consider a Markov-chain channel model, in which the probability that the channel state of vehicle  $i$  will change from  $H_m$  to  $H_n$  during a time slot is denoted by  $P_{m,n}^i = \Pr(h_i^{k+1} = H_n | h_i^k = H_m)$ . For simplicity, the channel state transition occurs only between neighboring states.

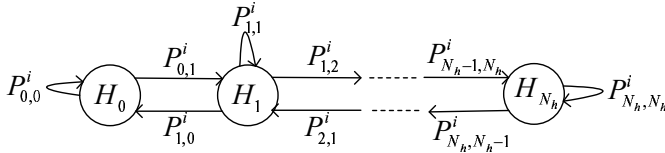


Fig. 2. The Markov-chain channel model of vehicle  $i$ .

The maximum speed of the vehicles is denoted by  $V$ , and the speed of vehicle  $i$  is denoted by  $v_i$ . The transition probability between the neighboring states of the channel model increases linearly with the vehicle speed, and is set for simplicity to be

$$P_{m,n}^i = \begin{cases} 1 - \gamma \frac{v_i}{V}, & 0 \leq m = n \leq N_h \\ \gamma \frac{v_i}{2V}, & 1 \leq m \leq N_h - 1 \text{ and } n = m \pm 1 \\ \gamma \frac{v_i}{V}, & (m, n) = (0, 1) \text{ or } (N_h, N_h - 1) \\ 0, & o.w., \end{cases} \quad (1)$$

where  $\gamma$  indicates the impact of other factors such as environmental changes. Note that our proposed crowdsensing system is not restricted to (1) and can be applied to other channel models straightforwardly. For ease of reference, our commonly used notation is summarized in TABLE 1.

## IV. VEHICULAR CROWDSENSING GAME

The interactions between the MCS server and  $M$  vehicles are formulated as a vehicular crowdsensing game. Each vehicle is assumed to be rational and autonomous to determine its sensing effort. The actions of the  $M$  vehicles, denoted by

TABLE I  
SUMMARY OF SYMBOLS AND NOTATIONS

$M$	Number of vehicles
$x_i \in \{0, 1, \dots, N_a\}$	Sensing effort of vehicle $i$
$\mathbf{y} = [y_j]_{0 \leq j \leq N_a}$	Payment strategy
$y_j \in \{b_i\}_{0 \leq i \leq N_b}$	Payment set
$\varrho = \{H_j\}_{0 \leq j \leq N_h}$	Channel state set
$v_i \in [0, V]$	Speed of vehicle $i$
$h_i$	Radio channel state of vehicle $i$
$c_i$	Sensing cost of vehicle $i$
$u_i^{a/b}$	Utility of vehicle $i$ for accumulative/best-quality tasks
$u_s^{a/b}$	Utility of the server for accumulative/best-quality tasks
$\beta_i$	Contribution of vehicle $i$ in MCS
$s/s_i$	State of the server/vehicle $i$
$\mathbf{Q}/\mathbf{Q}_i$	Quality function of the server/vehicle $i$
$\mathbf{V}/\mathbf{V}_i$	Value function of the server/vehicle $i$
$\alpha$	Learning rate
$\delta$	Discount factor in the learning process
$\tilde{s}_i$	Immediate state of vehicle $i$
$p^{w/u}$	Known/unknown state transition probability
$r^{w/u}$	Known/unknown reward
$N_a$	Number of nonzero sensing levels

$\mathbf{x} = [x_i]_{1 \leq i \leq M}$ , correspond to the accuracies of their sensing reports, with  $0 \leq x_i \leq N_a$ . The sensing cost of vehicle  $i$  depends on the performance of its sensors and the battery levels, with the unit sensing cost denoted by  $c_i$ . In addition, the transmission cost of a sensing report is determined by the channel condition from the vehicle to the serving BS.

The MCS server evaluates each sensing report based on the type of the MCS application, and determines the payment to the corresponding mobile device corresponding to the sensing effort. The type of the application depends on the location resolution of the service compared with the vehicle size. For example, if the vehicles on the same road provide different information for an MCS application, such as a 3-D street view application, the server benefits from receiving more sensing reports, and we call it an accumulative sensing task. Otherwise, if the MCS application has a coarse location resolution, e.g., all the vehicles on the same road provide similar sensing results, we call it best-quality sensing and the server does not gain from the sensing reports with similar or lower qualities.

1) Accumulative sensing tasks: The MCS server that performs an accumulative sensing task has to pay all the participating vehicles to stimulate more vehicles to take the sensing task. Take the pollution monitoring application in a city in [4] as an example, the MCS server enables the mapping of large-scale environmental phenomena by recruiting vehicles with built-in air quality sensing devices to measure air pollutants such as  $\text{CO}_2$ , and derives better measurements by receiving more sensing reports in the area. The contribution factor  $\beta_i > 0$  is introduced to weigh the contribution of vehicles  $i$  to the MCS task, such as the importance of its location, which can be estimated by the server according to the arrival

time of the sensing report and the distance of the vehicle from the area of priority in the MCS service using algorithms such as that described in [31]. More specifically, if vehicle  $i$  is in a well known location or several vehicles nearby have already sent sensing reports,  $\beta_i$  is low.

In this case, the gain of the server depends on the total contribution of all the sensing reports, given by  $\sum_{i=1}^M \beta_i x_i$ . Vehicle  $i$  receives a payment  $y_{x_i}$  according to the accuracy of its sensing data. Therefore, the immediate utility of the MCS server carrying on the accumulative sensing tasks, denoted by  $u_s^a$ , is given by

$$u_s^a(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^M (\beta_i x_i - y_{x_i}). \quad (2)$$

For simplicity, we assume that the sensing cost of vehicle  $i$  is proportional to its sensing effort  $x_i$ , and the inverse of the radio channel condition  $h_i$ . The immediate utility of vehicle  $i$  in this case, denoted by  $u_i^a$ , depends on the payment from the server, its sensing efforts, and its transmission costs, given by

$$u_i^a(x_i, \mathbf{y}) = y_{x_i} - \frac{x_i c_i}{h_i}, \quad 1 \leq i \leq M. \quad (3)$$

2) Best-quality sensing tasks: The utility of the server that performs a best-quality application depends on the best sensing report in a given area, and thus the server aims to stimulate a signal vehicle with the best sensing condition to participate in the task, and suppress the sensing incentive of the other vehicles to avoid data collisions. The best-quality sensing application has a coarse location resolution, i.e., all the vehicles on the same road are assumed to provide similar sensing reports, and  $\beta_i$  weigh the contribution of the vehicle in the sensing task. Therefore, the immediate utility of the server for a best-quality application, denoted by  $u_s^b$  and depended on the vehicle with the best sensing quality is given by

$$u_s^b(\mathbf{x}, \mathbf{y}) = \max_{1 \leq i \leq M} \beta_i x_i - y_{\max_{1 \leq i \leq M} x_i}. \quad (4)$$

The utility of vehicle  $i$ , denoted by  $u_i^b$ , is given by

$$u_i^b(x_i, \mathbf{y}, \mathbf{x}_{-i}) = y_{x_i} I(x_i = \max_{1 \leq j \leq M} x_j) - \frac{x_i c_i}{h_i}, \quad 1 \leq i \leq M, \quad (5)$$

where  $\mathbf{x}_{-i} = [x_j]_{1 \leq j \leq M, j \neq i}$ , and  $I(\cdot)$  is the indicator function, which takes the value 1 if its argument is true, and 0 otherwise.

In summary, we consider the vehicular crowdsensing game between the MCS server and  $M$  vehicles, denoted by  $\mathbf{G} = \langle \{s, 1, 2, \dots, M\}, \{\mathbf{y}, \mathbf{x}\}, \{u_s, u_{1 \leq i \leq M}\} \rangle$ . The MCS server (or vehicle  $i$ ) chooses its payment policy  $\mathbf{y}$  (or sensing strategy  $x_i$ ) to maximize its utility  $u_s$  (or  $u_i$ ). The total cost of the server, depends on the type of the sensing task and the total number of participant vehicles. The server makes a tradeoff between the sensing accuracy and the total payment by the server. For example, the server assigns a higher payment if it aims to obtain more accurate sensing in the target area. For best-quality sensing tasks, the server only pays the vehicle with the best sensing quality and its total payment is limited by the maximum payment to a vehicle, i.e.,  $y_{N_a}$ .

## V. NASH EQUILIBRIUM OF STATIC VEHICULAR CROWDSENSING GAME

The interactions between an MCS server and  $M$  vehicles in a time slot is formulated as a static vehicular crowdsensing game. The Nash equilibrium (NE) of the game is denoted by  $[\mathbf{x}^*, \mathbf{y}^*]$ , where  $\mathbf{x}^* = [x_i^*]_{0 \leq i \leq M}$  and  $\mathbf{y}^* = [y_j^*]_{0 \leq j \leq N_a}$ . By definition, an NE of the game  $\mathbf{G}$  (if one exists) is given by

$$x_i^* = \arg \max_{x_i \in \mathbf{A}} u_i(x_i, \mathbf{y}^*), \quad 1 \leq i \leq M \quad (6)$$

$$y_j^* = \arg \max_{y_j \in \mathbf{B}} u_s(\mathbf{x}^*, y_j), \quad 0 \leq j \leq N_a. \quad (7)$$

### A. Accumulative Sensing Tasks

We consider the NEs of the static MCS game  $\mathbf{G}$  for the accumulative sensing tasks, starting from a special case with two-level sensing effort, i.e.,  $N_a = 1$ . In this case, vehicle  $i$  either sends an accurate report (i.e.,  $x_i = 1$ ) or keeps silence (i.e.,  $x_i = 0$ ). By (7) and (2), we have  $y_0^* = 0$ . Thus the payment strategy of the server at the NE is given by  $\mathbf{y}^* = [0, y_1^*]$ .

**Theorem 1.** *If  $\frac{1}{M} \sum_{i=1}^M \beta_i > \max_{1 \leq i \leq M} \frac{c_i}{h_i}$ , the Pareto optimal NE of the static vehicular crowdsensing game  $\mathbf{G}$  for the accumulative sensing tasks with  $N_a = 1$  is given by*

$$\begin{aligned} x_i^* &= 1, \quad 1 \leq i \leq M \\ \mathbf{y}^* &= \left[ 0, \max_{1 \leq i \leq M} \frac{c_i}{h_i} \right]. \end{aligned} \quad (8)$$

*Proof.* By (3), if  $y_1^* \geq \frac{c_i}{h_i}$ , we have  $u_i^a(1, \mathbf{y}^*) = y_1^* - \frac{c_i}{h_i} \geq 0 = y_0^* = u_i^a(0, \mathbf{y}^*)$ . Thus, if  $y_1^* \geq \max_{1 \leq i \leq M} \frac{c_i}{h_i}$ , (6) holds for  $x_i^* = 1, \forall 1 \leq i \leq M$ . By (2),  $u_s^a$  monotonically decreases with  $y_1$ , yielding  $u_s^a([1, \dots, 1], [0, y_1]) = \sum_{i=1}^M \beta_i - M y_1 < \sum_{i=1}^M \beta_i - M y_1^* = u_s^a([1, \dots, 1], [0, y_1^*])$ ,  $\forall y_1 > y_1^*$ , and by (7), we have  $y_1^* = \max_{1 \leq i \leq M} \frac{c_i}{h_i}$ . If  $\frac{1}{M} \sum_{i=1}^M \beta_i > \max_{1 \leq i \leq M} \frac{c_i}{h_i}$ , we have  $u_s^a([1, \dots, 1], [0, y_1^*]) > 0$ . Thus (7) holds for (8), which is an NE of the game.

Now we prove that this NE is Pareto optimal. Suppose  $(\mathbf{x}', \mathbf{y}')$  is another NE of the game, with  $(\mathbf{x}^*, \mathbf{y}^*) \neq (\mathbf{x}', \mathbf{y}')$ . Without loss of generality, assume  $x'_i = 0$  for vehicle  $i'$ . As shown in (3),  $u_{i'}^a(x'_i, \mathbf{y}') = 0 < u_{i'}^a(x_i^*, \mathbf{y}^*)$ . Thus  $(\mathbf{x}^*, \mathbf{y}^*)$  is Pareto optimal.  $\square$

**Remark:** In this case, all the  $M$  vehicles participate in the sensing tasks, if the payment exceeds the sensing cost of all the vehicles, the total contribution of the vehicles is greater than the payment and  $\frac{1}{M} \sum_{i=1}^M \beta_i > \max_{1 \leq i \leq M} \frac{c_i}{h_i}$ .

We now derive the NE for the case with  $N_a = 2$ , in which vehicle  $i$  chooses to send a high-quality sensing report (i.e.,  $x_i = 2$ ), a low-quality sensing report (i.e.,  $x_i = 1$ ), or keep silent (i.e.,  $x_i = 0$ ). As  $y_0^* = 0$ , the corresponding payment strategy at the NE is given by  $\mathbf{y}^* = [0, y_1^*, y_2^*]$ .

**Proposition 1.** *In the static vehicular crowdsensing game  $\mathbf{G}$  with  $N_a = 2$ , vehicle  $i$  is motivated to send a high-quality sensing report, i.e.,  $x_i^* = 2$ , if*

$$y_2^* = \max \left( \frac{2c_i}{h_i}, y_1^* + \frac{c_i}{h_i} \right). \quad (9)$$

*Proof.* If  $y_2^* \geq \frac{2c_i}{h_i}$ , by (6) we have

$$u_i^a(2, \mathbf{y}^*) = y_2^* - \frac{2c_i}{h_i} \geq 0 = y_0^* = u_i^a(0, \mathbf{y}^*). \quad (10)$$

If  $y_2^* \geq y_1^* + \frac{c_i}{h_i}$ , we have

$$u_i^a(2, \mathbf{y}^*) = y_2^* - \frac{2c_i}{h_i} \geq y_1^* - \frac{c_i}{h_i} = u_i^a(1, \mathbf{y}^*). \quad (11)$$

Combing (10) and (11), we have  $x_i^* = 2$ , if  $y_2^* \geq \max\left(\frac{2c_i}{h_i}, y_1^* + \frac{c_i}{h_i}\right)$ . As  $u_s^a$  decreases with the payment, if  $0 \leq y_1^* \leq \frac{c_i}{h_i}$ , we have  $y_2^* = \frac{2c_i}{h_i}$ ; otherwise, if  $y_1^* > \frac{c_i}{h_i}$ , we have  $y_2^* = y_1^* + \frac{c_i}{h_i}$ .  $\square$

**Remark:** If the server offers a sufficiently high payment for the high-quality sensing reports, and a low payment for the low-quality sensing reports, as in (9), a vehicle is motivated to send a high-quality sensing report.

**Proposition 2.** *In the static vehicular crowdsensing game  $\mathbf{G}$  with  $N_a = 2$ , vehicle  $i$  is motivated to send a low-quality sensing report, i.e.,  $x_i^* = 1$ , if*

$$y_1^* = \max\left(\frac{c_i}{h_i}, y_2^* - \frac{c_i}{h_i}\right). \quad (12)$$

*Proof.* Similar to that of Proposition 1.  $\square$

**Remark:** A vehicle is motivated to send a low-quality sensing report if the payment for low-quality sensing reports is high enough to cover the sensing and transmission cost of the vehicle, and the payment gap regarding the sensing quality is small as shown in (12).

**Proposition 3.** *In the static vehicular crowdsensing game  $\mathbf{G}$  with  $N_a = 2$ , vehicle  $i$  has no motivation to participate in the crowdsensing, i.e.,  $x_i^* = 0$ , if*

$$y_1^* < \frac{c_i}{h_i} \quad \text{and} \quad y_2^* < \frac{2c_i}{h_i}. \quad (13)$$

*Proof.* Similar to that of Proposition 1.  $\square$

**Remark:** If the sensing and transmission cost of a vehicle is high compared with the expected payment offered by the server, the vehicle has no motivation to respond.

**Theorem 2.** *If  $\frac{1}{M} \sum_{i=1}^M \beta_i > \max_{1 \leq i \leq M} \frac{c_i}{h_i}$ , the static vehicular crowdsensing game  $\mathbf{G}$  with accumulative sensing tasks and  $N_a = 2$  has an NE given by*

$$\begin{aligned} x_i^* &= 2, \quad 1 \leq i \leq M \\ \mathbf{y}^* &= \left[0, 0, 2 \max_{1 \leq i \leq M} \frac{c_i}{h_i}\right]. \end{aligned} \quad (14)$$

*Proof.* According to Proposition 1, if  $y_2^* \geq \max\left(\frac{2c_i}{h_i}, y_1^* + \frac{c_i}{h_i}\right)$ , we have  $x_i^* = 2, \forall 1 \leq i \leq M$ . By (2),  $u_s^a$  monotonically decreases with  $y_2$ , yielding  $u_s^a([2, \dots, 2], [0, y_1, y_2]) = \sum_{i=1}^M 2\beta_i - My_2 < \sum_{i=1}^M 2\beta_i - My_2^* = u_s^a([2, \dots, 2], [0, y_1, y_2^*]), \forall y_2 > y_2^*$ . Therefore, by (7), we have  $y_2^* = \max_{1 \leq i \leq M} \frac{2c_i}{h_i}$  and  $y_1^* = 0$  in this case. If  $\frac{1}{M} \sum_{i=1}^M \beta_i > \max_{1 \leq i \leq M} \frac{c_i}{h_i}$ , we have  $u_s^a([2, \dots, 2], \mathbf{y}^*) > 0$ . The remainder of the proof is similar to that of Theorem 1.  $\square$

**Remark:** As shown in (14), the vehicle with the best sensing condition receives double payment, and the other vehicles also have motivation to send accurate sensing reports. Otherwise, if the server has a tight budget, the payment is reduced and some vehicles stay away from the sensing task.

**Faked Sensing Attacks:** Now we consider a static vehicular crowdsensing game  $\mathbf{G}'$  with faked sensing attacks. For simplicity, we assume  $\mathbf{A} = \{-1, 0, 1\}$ . Vehicle  $i$  can choose to send an accurate report (with  $x_i = 1$ ), keep silence ( $x_i = 0$ ), or upload a faked report ( $x_i = -1$ ). It is clear that  $\mathbf{y}^* = [0, 0, y_1^*]$ , where  $y_1^*$  is the payment to an accurate report.

**Theorem 3.** *If  $\frac{1}{M} \sum_{i=1}^M \beta_i > \max_{1 \leq i \leq M} \frac{2c_i}{h_i}$ , the static secure vehicular crowdsensing game  $\mathbf{G}'$  with accumulative sensing tasks and  $\mathbf{A} = \{-1, 0, 1\}$  has an NE given by*

$$\begin{aligned} x_i^* &= 1, \quad 1 \leq i \leq M \\ \mathbf{y}^* &= \left[0, 0, 2 \max_{1 \leq i \leq M} \frac{c_i}{h_i}\right]. \end{aligned} \quad (15)$$

*Proof.* Similar to the proof to Theorem 1, if  $y_1^* \geq \frac{2c_i}{h_i}$ , we have  $u_i^a(1, \mathbf{y}^*) = y_1^* - \frac{c_i}{h_i} \geq 0 = y_0^* = u_i^a(0, \mathbf{y}^*)$  and  $u_i^a(1, \mathbf{y}^*) = y_1^* - \frac{c_i}{h_i} \geq y_{-1}^* + \frac{c_i}{h_i} = u_i^a(-1, \mathbf{y}^*)$ . Thus, if  $y_1^* \geq \max_{1 \leq i \leq M} \frac{2c_i}{h_i}$ , (5) holds for  $x_i^* = 1, \forall 1 \leq i \leq M$ . By (2),  $u_s^a$  monotonically decreases with  $y_1$ , yielding  $u_s^a([1, \dots, 1], [0, 0, y_1]) = \sum_{i=1}^M \beta_i - My_1 < \sum_{i=1}^M \beta_i - My_1^* = u_s^a([1, \dots, 1], [0, 0, y_1^*]), \forall y_1 > y_1^*$ . Therefore, by (7), we have  $y_1^* = \max_{1 \leq i \leq M} \frac{2c_i}{h_i}$ . If  $\frac{1}{M} \sum_{i=1}^M \beta_i > \max_{1 \leq i \leq M} \frac{2c_i}{h_i}$ , we have  $u_s^a([1, \dots, 1], [0, 0, y_1^*]) > 0$ . Thus (6) holds for (15), which is an NE of the game.  $\square$

**Remark:** When the participants report untruthfully, the server stimulates accurate sensing by offering a higher payment to the vehicles with accurate reports and suppresses untruthful behavior by not paying the cheating participants. In other words, if a participant reports untruthfully, the honest participants receive higher payments and the normalized loss due to cheating is large.

As shown in Eqs. (14) and (15), the server only pays the most accurate sensing reports with a high payment to stimulate accurate sensing. Both theorems provide the NE of the game with three sensing levels, i.e.,  $\mathbf{A} = \{0, 1, 2\}$  in Theorem 2 and  $\mathbf{A} = \{-1, 0, 1\}$  in Theorem 3. In both case, the server pays the same to the vehicles that provide accurate sensing report, i.e.,  $x_i = 2$  in Theorem 2 and  $x_i = 1$  in Theorem 3. The inaccurate reports (i.e.,  $x_i = 1$  in Theorem 2) or faked reports (i.e.,  $x_i = -1$  in Theorem 3) are not paid to stimulate all the  $M$  vehicles to send accurate sensing reports.

## B. Best-Quality Sensing Tasks

We consider the NEs of the static MCS game with the best-quality sensing tasks in the following.

**Theorem 4.** If  $\beta_{i^*} > \frac{c_{i^*}}{h_{i^*}}$ ,  $i^* = \arg \min_{1 \leq i \leq M} \frac{c_i}{h_i}$ , the Pareto optimal NE of the static vehicular crowdsensing game  $\mathbf{G}$  with the best-quality sensing tasks and  $N_a = 1$  is given by

$$x_i^* = \begin{cases} 1, & i = i^* \\ 0, & o.w. \end{cases} \quad (16)$$

$$\mathbf{y}^* = \left[ 0, \frac{c_{i^*}}{h_{i^*}} \right]. \quad (17)$$

*Proof.* By (5), if  $i \neq i^*$ ,  $u_i^b(\mathbf{x}, \mathbf{y}) = -\frac{x_i c_i}{h_i}$ , indicating that  $x_i^* = 0$ . Otherwise, if  $i = i^*$ , by (5), we have  $u_{i^*}^b(\mathbf{x}, \mathbf{y}) = y x_{i^*} - \frac{x_{i^*} c_{i^*}}{h_{i^*}}$ , with  $u_{i^*}^b(\mathbf{0}, \mathbf{y}) = 0$ .

If  $y_1^* \geq \frac{c_{i^*}}{h_{i^*}}$ , we have  $u_{i^*}^b(1, \mathbf{y}^*) \geq 0 = u_{i^*}^b(0, \mathbf{y}^*)$ . By (6), we see that  $x_{i^*}^* = 1$ . If  $\beta_{i^*} > \frac{c_{i^*}}{h_{i^*}}$ , by (4) and (7), we have  $u_s^b(\mathbf{x}^*, \mathbf{y}^*) = \beta_{i^*} - y_1^* > \beta_{i^*} - y_1 = u_s^b(\mathbf{x}^*, \mathbf{y})$ , yielding  $y_1^* \leq y_1$ . Thus  $y_1^* = \frac{c_{i^*}}{h_{i^*}}$ , and  $[\mathbf{x}^*, \mathbf{y}^*]$  in (16) and (17) provides an NE of the game in this case.

If  $y_1^* < \frac{c_{i^*}}{h_{i^*}}$ , we have  $u_{i^*}^b(1, y_1^*) < 0 = u_{i^*}^b(0, y_1^*)$ , yielding  $x_{i^*}^* = 0$ . By (4), we have  $u_s^b(\mathbf{x}^*, \mathbf{y}^*) = 0 = u_s^b(\mathbf{x}^*, \mathbf{y}^*)$ . Combining both cases, we see that (16) and (17) provide the Pareto optimal NE of the game.  $\square$

**Remark:** As shown in Theorem 4, the vehicle with the best sensing condition applies its full sensing effort, while the others do not respond to save energy, and avoid sensing redundancy and transmission collision. In addition, the sensing task in this case is more likely to be taken than the accumulative sensing tasks, as  $\Pr(\beta_i > \min_{1 \leq i \leq M} \frac{c_i}{h_i}) \geq \Pr(\beta_i > \max_{1 \leq i \leq M} \frac{c_i}{h_i})$ .

**Theorem 5.** If  $\beta_{i^*} > \frac{c_{i^*}}{h_{i^*}}$ ,  $i^* = \arg \min_{1 \leq i \leq M} \frac{c_i}{h_i}$ , an NE of the static vehicular crowdsensing game  $\mathbf{G}$  with best-quality sensing tasks and  $N_a = 2$  is given by

$$x_i^* = \begin{cases} 2, & i = i^* \\ 0, & o.w. \end{cases} \quad (18)$$

$$\mathbf{y}^* = \left[ 0, 0, 2 \frac{c_{i^*}}{h_{i^*}} \right].$$

*Proof.* Similar to that of Theorem 3.  $\square$

**Remark:** In this case, the vehicle with the lowest sensing cost and transmission consumption (i.e.,  $\min_{1 \leq i \leq M} \frac{2c_i}{h_i}$ ) is motivated to send a high-quality sensing report.

In summary, we have considered the cases with  $N_a = 1$  sensing level, or  $N_a = 2$  (i.e., a vehicle chooses to send an accurate sensing report with  $x_i = 2$ , a coarse sensing report with  $x_i = 1$ , or no report with  $x_i = 0$ ). In addition, we have also discussed faked sensing attacks (i.e.,  $x_i = -1$ ) in Theorem 3. The NE of the MCS game in these scenarios show the impacts of the radio channel conditions, sensing costs, and the type of the MCS tasks on mobile crowdsensing. We can see similar trends for the case with more sensing effort levels.

## VI. DYNAMIC VEHICULAR CROWDSENSING WITH LEARNING

In this section, we formulate the repeated interactions between the MCS server and  $M$  independent vehicles in a dynamic network as a dynamic vehicular crowdsensing

game. As it is impractical for a vehicle or the server to accurately estimate all the system parameters in time, they can apply reinforcement learning techniques, such as Q-learning, to derive their optimal strategy via trials without explicitly knowing the system model. We also apply the PDS-learning technique [32] to accelerate the learning process of a vehicle by exploiting the known radio channel model.

### A. Payment based on Q-learning

The payment decision process can be formulated as a Markov decision process with finite states. Thus the MCS server can pay the participating vehicles based on Q-learning technique. In each time slot, the server builds the system state that consists of the previous sensing experience to determine its payment.

Take the accumulative sensing tasks as an example. The MCS server calculates the number of received sensing reports with quality level  $j$ , in the last time slot, which is denoted by  $\zeta_j^{k-1}$ , and given by

$$\zeta_j^{k-1} = \sum_{i=1}^M I(x_i^{k-1} = j), \quad 0 \leq j \leq N_a, \quad (19)$$

where  $I(\cdot)$  is the indicator function. The system state at time  $k$  consists of the number of sensing reports of each type in the last time slot, i.e.,  $\mathbf{s}^k = [\zeta_j^{k-1}]_{1 \leq j \leq N_a} \in \mathbf{S}^a = \{0, 1, \dots, M\}^{N_a}$ . Similarly, for best-quality sensing tasks, the system state at time  $k$  is given by  $\mathbf{s}^k = \max_{1 \leq i \leq M} x_i^{k-1} \in \mathbf{S}^b = \{0, 1, \dots, N_a\}$ . Fig. 3 presents the state transition of the MCS server in this case, in which the state depends on the sensing efforts of the  $M$  vehicles in the last time slot.

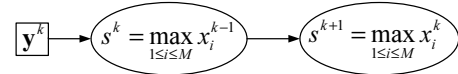


Fig. 3. State transition of a Q-learning based MCS server for best-quality sensing tasks.

For both types of MCS applications, the sensing efforts of all vehicles in the last time slot determine the system state in the next time slot, i.e.,  $(s/s)^{k+1} \in \mathbf{S}^{a/b}$ . Let  $Q(\mathbf{s}, \mathbf{y})$  denote the Q function of the server at state  $\mathbf{s}$  with action  $\mathbf{y}$ . According to the  $\epsilon$ -greedy policy, with  $0 < \epsilon \leq 1$ , the server chooses the payment with the highest expected utility or Q-function with a high probability  $1 - \epsilon$ , and randomly selects one of the other actions with a small probability. The payment policy  $\mathbf{y}^k$  is given by

$$\Pr(\mathbf{y}^k = \mathbf{y}^*) = \begin{cases} 1 - \epsilon, & \mathbf{y}^* = \arg \max_{\mathbf{y}} Q(\mathbf{s}^k, \mathbf{y}) \\ \frac{\epsilon}{|\mathbf{B}|^{N_a - 1}}, & o.w., \end{cases} \quad (20)$$

where  $|\cdot|$  denotes the size of the set. The server observes the qualities of the received sensing reports,  $\mathbf{x}^k$ , and obtains the next system state  $\mathbf{s}^{k+1}$ . The value function of the state  $\mathbf{s}$

denoted by  $V(\mathbf{s})$ , indicates the highest Q-function at state  $\mathbf{s}$ . The MCS server updates its Q function by the following:

$$Q(\mathbf{s}^k, \mathbf{y}^k) \leftarrow (1 - \alpha)Q(\mathbf{s}^k, \mathbf{y}^k) + \alpha(u_s^a(\mathbf{s}^k, \mathbf{y}^k) + \delta V(\mathbf{s}^{k+1})) \quad (21)$$

$$V(\mathbf{s}^k) = \max_{\mathbf{y}} Q(\mathbf{s}^k, \mathbf{y}), \quad (22)$$

where  $\alpha \in (0, 1]$  is the learning rate of the payment strategy, and  $\delta \in [0, 1]$  is the discount factor that indicates the weight of a future payoff over the current one. The payment strategy is summarized in Algorithm 1.

---

**Algorithm 1** Payment strategy with Q-learning.

---

- 1: **Initialize**  $\alpha, \delta, \mathbf{s}^1 = \mathbf{0}, Q(\mathbf{s}, \mathbf{y}) = 0, V(\mathbf{s}) = 0, \forall \mathbf{s}, \mathbf{y}$
  - 2: **for**  $k = 1, 2, 3, \dots$  **do**
  - 3:   Choose  $\mathbf{y}^k$  via (20)
  - 4:   Receive  $u_s^a$
  - 5:   Observe  $\mathbf{s}^{k+1}$
  - 6:   Update  $Q(\mathbf{s}^k, \mathbf{y}^k)$  via (21)
  - 7:   Update  $V(\mathbf{s}^k)$  via (22)
  - 8: **end for**
- 

### B. Sensing based on Q-learning

The sensing decision process of a vehicle can be modelled as an MDP. Thus, a vehicle can apply Q-learning technique to derive its sensing strategy. The system state observed by vehicle  $i$  at time  $k$  is given by  $\mathbf{s}_i^k = [h_i^{k-1}, \mathbf{y}^{k-1}] \in \mathbf{S}_i$ , and consists of the channel condition and the server's payment in the last time slot, where  $\mathbf{S}_i$  is its state set. Fig. 4 presents the state transition for vehicle  $i$  in the sensing process.

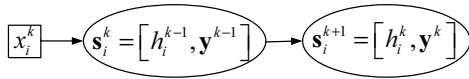


Fig. 4. State transition of vehicle  $i$  with Q based sensing strategy in the dynamic mobile crowdsensing.

Let  $Q_i(\mathbf{s}_i, x_i)$  denote the quality function of vehicle  $i$  at state  $\mathbf{s}_i$  and sensing effort  $x_i$ , and  $V_i(\mathbf{s}_i)$  be the value function of state  $\mathbf{s}_i$  for vehicle  $i$ . The Q-function is updated by

$$Q_i(\mathbf{s}_i^k, x_i^k) \leftarrow (1 - \alpha)Q_i(\mathbf{s}_i^k, x_i^k) + \alpha(u_i^a(\mathbf{s}_i^k, x_i^k) + \delta V_i(\mathbf{s}_i^{k+1})) \quad (23)$$

$$V_i(\mathbf{s}_i^k) = \max_{x_i} Q_i(\mathbf{s}_i^k, x_i). \quad (24)$$

Based on the  $\epsilon$ -greedy policy, vehicle  $i$  chooses its sensing effort as

$$\Pr(x_i^k = x^*) = \begin{cases} 1 - \epsilon, & x^* = \arg \max_{x_i} Q_i(\mathbf{s}_i^k, x_i) \\ \frac{\epsilon}{N_a}, & o.w. \end{cases} \quad (25)$$

The sensing strategy with Q-learning is summarized in Algorithm 2.

---

**Algorithm 2** Sensing strategy of vehicle  $i$  with Q-learning.

---

- 1: **Initialize**  $\alpha, \delta, \mathbf{s}_i^1 = \mathbf{0}, Q_i(\mathbf{s}_i, x_i) = 0, V(\mathbf{s}_i) = 0, \forall \mathbf{s}_i, x_i$
  - 2: **for**  $k = 1, 2, 3, \dots$  **do**
  - 3:   Choose  $x_i^k$  via (25)
  - 4:   Receive  $u_i^a$
  - 5:   Observe  $\mathbf{s}_i^{k+1}$
  - 6:   Update  $Q_i(\mathbf{s}_i^k, x_i^k)$  via (23)
  - 7:   Update  $V_i(\mathbf{s}_i^k)$  via (24)
  - 8: **end for**
- 

### C. Sensing Strategy based on PDS-learning

The radio channel variation speed depends on the moving speed of the vehicle, as shown in Eq. (1). By exploiting the partially known channel model, a vehicle can apply the PDS-learning technique [32] to improve its learning speed, because less information has to be learned.

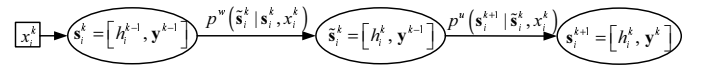


Fig. 5. State transition of vehicle  $i$  with PDS based sensing strategy in the dynamic mobile crowdsensing game.

As shown in Fig. 5, vehicle  $i$  chooses its sensing effort  $x_i^k \in \{0, 1, \dots, N_a\}$  based on the system state  $\mathbf{s}_i^k = [h_i^{k-1}, \mathbf{y}^{k-1}]$ . According to PDS-learning, we define the post-decision state as an immediate state that is reached after the known dynamics (channel variances) takes place but before the unknown dynamics (payment by the server) occurs. The vehicle sees an immediate state  $\tilde{\mathbf{s}}_i^k = [h_i^k, \mathbf{y}^{k-1}]$  with a known probability  $p^w(\tilde{\mathbf{s}}_i^k | \mathbf{s}_i^k, x_i^k) = \Pr(h_i^k | h_i^{k-1})$ . Then the next state  $\mathbf{s}_i^{k+1} = [h_i^k, \mathbf{y}^k]$  follows the unknown distribution  $p^u(\mathbf{s}_i^{k+1} | \tilde{\mathbf{s}}_i^k, x_i^k)$ . The immediate reward to the vehicle consists of both a known and an unknown part, in this setting defined by  $r^w(\mathbf{s}_i, x_i) = u_i^a(\mathbf{s}_i, x_i)$  and  $r^u = 0$ , respectively.

The quality function of vehicle  $i$  in the post-decision state, denoted by  $\tilde{Q}_i(\tilde{\mathbf{s}}_i, x_i)$ , is updated iteratively by

$$\tilde{Q}_i(\tilde{\mathbf{s}}_i^k, x_i^k) \leftarrow (1 - \alpha)\tilde{Q}_i(\tilde{\mathbf{s}}_i^k, x_i^k) + \alpha(r^w(\tilde{\mathbf{s}}_i^k, x_i^k) + \delta V_i(\mathbf{s}_i^{k+1})). \quad (26)$$

The sensing strategy is chosen according to the  $\epsilon$ -greedy policy in (25). Then the quality function is updated for all the state-action pairs as follows:

$$Q_i(\mathbf{s}_i, x_i) \leftarrow r^u(\mathbf{s}_i, x_i) + \sum_{\tilde{\mathbf{s}}_i} p^w(\tilde{\mathbf{s}}_i | \mathbf{s}_i, x_i) \tilde{Q}_i(\tilde{\mathbf{s}}_i, x_i), \forall \mathbf{s}_i, x_i. \quad (27)$$

The sensing strategy based on PDS-learning is summarized in Algorithm 3.

## VII. SIMULATION RESULTS

Simulations have been performed to evaluate the performance of the vehicular crowdsensing game. To represent the quality of the MCS application, we introduced the sensing

**Algorithm 3** Sensing strategy of vehicle  $i$  with PDS-learning.

- 1: **Initialize**  $\alpha, \delta, \mathbf{s}_i^1 = \mathbf{0}, Q_i(\mathbf{s}, x) = 0, V(\mathbf{s}) = 0, \forall \mathbf{s}, x$
- 2: **for**  $k = 1, 2, 3, \dots$  **do**
- 3:   Choose  $x_i^k$  via (25)
- 4:   Receive  $r^w$
- 5:   Obtain  $\tilde{\mathbf{s}}_i^k$
- 6:   Observe  $\mathbf{s}_i^{k+1}$
- 7:   Update  $\tilde{Q}_i(\tilde{\mathbf{s}}_i^k, x_i^k)$  via (26)
- 8:   Update  $Q_i(\mathbf{s}, x)$  via (27)
- 9:   Update  $V_i(\mathbf{s})$  via (24)
- 10: **end for**

quality index of the server, which is denoted by  $\pi$  and given by

$$\pi = \begin{cases} \frac{1}{TN_a} \sum_{i=1}^M \beta_i x_i, & \text{accumulative sensing} \\ \frac{1}{N_a} \max_{1 \leq i \leq M} \beta_i x_i, & \text{best-quality sensing,} \end{cases} \quad (28)$$

where  $T$  is the maximum number of sensing results expected by the MCS server for this application.

Unless specified otherwise, we set in the simulations  $\alpha = 0.7, \delta = 0.8, \epsilon = 0.1, V = 5, N_a = 10, N_b = 25, N_h = 1, \mathbf{A} = \{0 \leq i \leq 10\}, \mathbf{B} = \{0 \leq k \leq 25\}, \varrho = \{1, 10\}$ , and  $0 \leq M \leq 80$ . The contribution factors of the  $M$  vehicles were chosen from  $[1, 20]$  at random, i.e.,  $\beta_i \in [1, 20], \forall 1 \leq i \leq M$ . We evaluate the  $\epsilon$ -greedy MCS strategy as a benchmark, in which the server selects the payment that maximizes its immediate utility with a high probability, and chooses the other payments randomly each with a low probability.

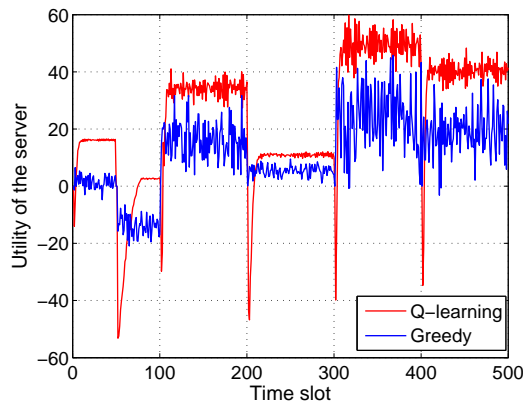
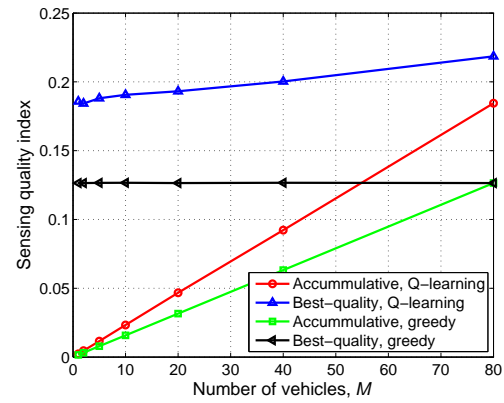
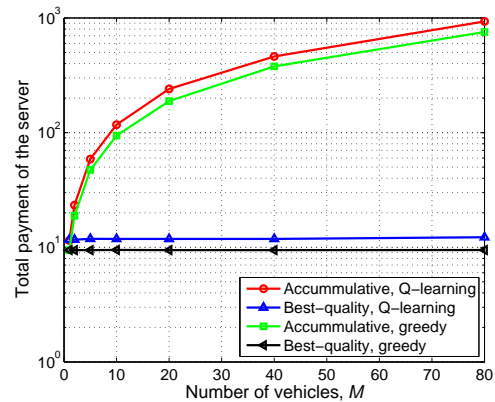


Fig. 6. Utility of the server with accumulative sensing tasks with  $c = 10$  and  $\beta_i \in [1, 20], \forall 1 \leq i \leq 10$  in which the server chooses the greedy payment or the Q-learning based payment, while each vehicle chooses the sensing effort to maximize its immediate utility.

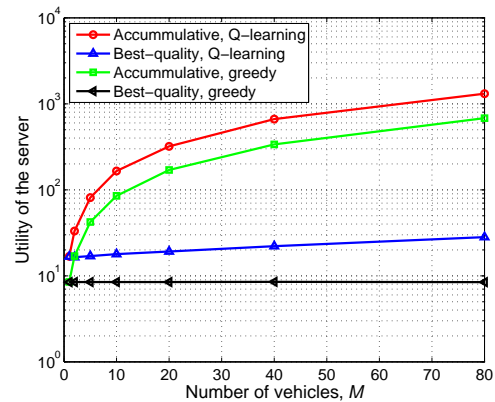
We first evaluate the vehicular crowdsensing game, in which the server applies the Q-learning based payment strategy and each of  $M = 10$  vehicles that joins or leaves the sensing task over time chooses its sensing effort for accumulative sensing tasks to maximize its immediate utility with  $c = 10$ . As shown in Fig. 6, the utility of the server drops as the vehicles join or leave the sensing task, but increases rapidly to a high value. The Q-learning based payment strategy exceeds the greedy payment strategy with a higher utility.



(a) Average sensing quality index



(b) Average total payment of the server



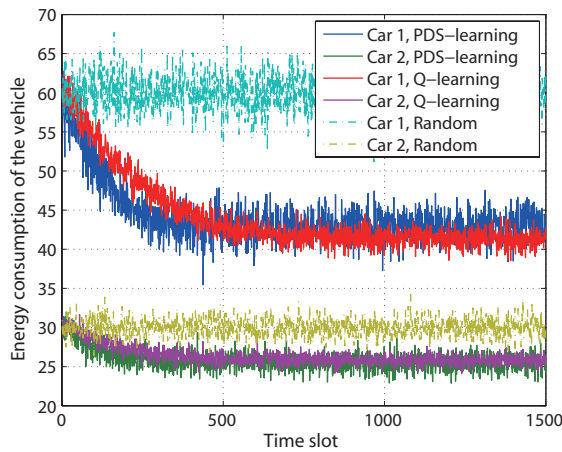
(c) Average utility of the server

Fig. 7. Performance of the vehicular crowdsensing game with  $c = 10$  and  $\beta_i = 40, \forall 1 \leq i \leq M$ , in which each vehicle chooses the sensing effort to maximize its immediate utility.

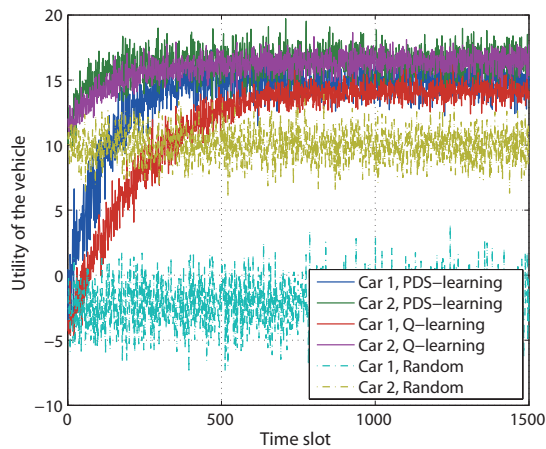
The performance of the vehicular crowdsensing game with  $\beta_i = 40, \forall 1 \leq i \leq M$  is presented in Fig. 7, in which all the  $M$  vehicles choose their sensing efforts to maximize their immediate utilities. As shown in Fig. 7 (a), the sensing quality index of the server with the Q-learning based payment for accumulative sensing tasks increases with the number of vehicles, e.g., it increases from 0.002 to 0.185 as the number of vehicles changes from 1 to 80, because the server benefits from receiving more sensing reports. The performance gain of



the sensing quality due to more vehicles is smaller for best-quality sensing tasks, e.g., the sensing quality increases by 18% as  $M$  increases from 1 to 80. In addition, our proposed payment can improve the sensing quality index compared with the greedy payment by 72% with  $M = 80$  vehicles in a best-quality sensing task. Similar observations can be made for an accumulative sensing task. Fig. 7 (b) shows that the total payment of the server with Q-learning for an accumulative sensing task increases with the number of vehicles, e.g., it increases from 11.48 to 932.33, as  $M$  increases from 1 to 80. Fig. 7 (c) indicates that the utility of the server with the proposed payment strategy for an accumulative sensing task increases with the number of vehicles, it increases from 16.8 to 1310.3 as  $M$  increases from 1 to 80. As shown in Fig. 7 (b) and (c), e.g., the utility of the server increases by 92%, and the total payment is only 23% higher, compared with the greedy payment with  $M = 80$ .



(a) Energy consumption of the vehicle

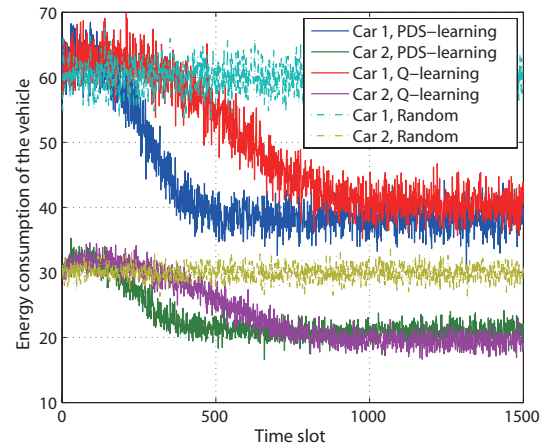


(b) Utility of the vehicle

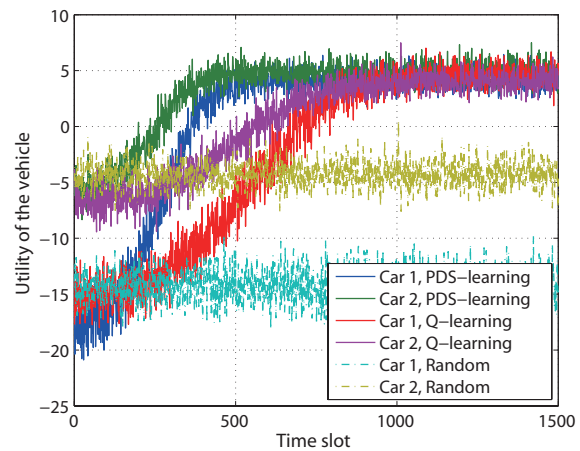
Fig. 8. Performance of vehicles in a dynamic vehicular crowdsensing game for an accumulative sensing task, with  $M = 3$  vehicles,  $c_1 = 10$ ,  $c_2 = 5$ ,  $c_3 = 1$  and  $\beta_i \in [1, 20]$ ,  $\forall 1 \leq i \leq 3$ .

In Fig. 8, the MCS server for accumulative sensing tasks with  $M = 3$ ,  $c_1 = 10$ ,  $c_2 = 5$ , and  $c_3 = 1$  chooses the random payment showing that the mobile crowdsensing with the

proposed sensing strategy reduces the energy consumption of the vehicles, e.g., the energy consumption of Car 1 decreases by 30% after 400 iterations. The vehicle with a better sensing condition (i.e., Car 2 with  $c_2 = 5$ ) consumes less energy for accumulative sensing tasks, e.g., energy consumption of Car 2 is 40% of that of Car 1. The PDS-learning based sensing strategy increases the utility of Car 2 by 70% compared with random sensing strategy. The PDS-learning based sensing strategy enables a vehicle to learn faster than Q-learning, e.g., the convergence time of Car 1 with PDS-learning is 4/7 of that with the Q-learning.



(a) Energy consumption of the vehicle



(b) Utility of the vehicle

Fig. 9. Performance of vehicles in a dynamic vehicular crowdsensing game for a best-quality sensing task, with  $M = 3$  vehicles,  $c_1 = 10$ ,  $c_2 = 5$ ,  $c_3 = 1$  and  $\beta_i \in [1, 20]$ ,  $\forall 1 \leq i \leq 3$ .

As shown in Fig. 9 (a), the proposed sensing strategies reduce the energy consumption of vehicles that take the best-quality sensing task, e.g., the energy consumption of Car 1 of the Q-learning based sensing strategy is reduced by 42%. The Q-learning based sensing strategy improves the utility of the vehicle compared with the benchmark strategy, and the convergence rate is further improved by PDS-learning. For example, the convergence time of Car 1 with the PDS-learning based sensing is only 50% of the Q-learning based strategy.



- [15] X. Hu, J. Deng, J. Zhao, *et al.*, "SafeDJ: A crowd-cloud codesign approach to situation-aware music delivery for drivers," *ACM Trans. Multimedia Comput., Commun., and Appl. (TOMM)*, vol. 12, no. 1, Oct., 2015.
- [16] K. Lan, C. Chou, and H. Wang, "An incentive-based framework for vehicle-based mobile sensing," *Procedia Computer Science*, vol. 10, no. 2, pp. 1152–1157, Aug., 2012.
- [17] S. Santini, A. Salvi, A. Valente, *et al.*, "A consensus-based approach for platooning with inter-vehicular communications," in *Proc. IEEE Int'l Conf. Computer Commun. (INFOCOM)*, pp. 1158–1166, Apr., 2015.
- [18] Z. He, J. Cao, and X. Liu, "High quality participant recruitment in vehicle-based crowdsourcing using predictable mobility," in *Proc. IEEE Int'l Conf. Computer Commun. (INFOCOM)*, pp. 2542–2550, Apr., 2015.
- [19] S. Sur and X. Zhang, "Bridging link power asymmetry in mobile whitespace networks," in *Proc. IEEE Int'l Conf. Computer Commun. (INFOCOM)*, pp. 1176–1184, Apr., 2015.
- [20] X. Hu, X. Li, E. Ngai, V. Leung, and P. Kruchten, "Multidimensional context-aware social network architecture for mobile crowdsensing," *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 78–87, Jun., 2014.
- [21] A. Tarable, A. Nordio, E. Leonardi, and M. Marsan, "The importance of being earnest in crowdsourcing systems," in *Proc. IEEE Int'l Conf. Computer Commun. (INFOCOM)*, pp. 2821–2829, Apr., 2015.
- [22] X. Zhang, G. Xue, R. Yu, D. Yang, and J. Tang, "Truthful incentive mechanisms for crowdsourcing," in *Proc. IEEE Int'l Conf. Computer Commun. (INFOCOM)*, pp. 2830–2838, Apr., 2015.
- [23] X. Wang, W. Cheng, P. Mohapatra, and T. Abdelzaher, "Artsense: Anonymous reputation and trust in participatory sensing," in *Proc. IEEE Int'l Conf. Computer Commun. (INFOCOM)*, pp. 2517–2525, Apr., 2013.
- [24] A. C. Weaver, J. P. Boyle, L. Besaleva, *et al.*, "Applications and trust issues when crowdsourcing a crisis," in *Proc. IEEE Computer Commun. and Networks (ICCCN)*, pp. 1–5, Jul., 2012.
- [25] Q. Li, G. Cao, and T. F. Porta, "Efficient and privacy-aware data aggregation in mobile sensing," *IEEE Trans. Dependable and Secure Computing*, vol. 11, no. 2, pp. 115–129, Apr., 2014.
- [26] Y. Hua, W. He, X. Liu, and D. Feng, "SmartEye: Real-time and efficient cloud image sharing for disaster environments," in *Proc. IEEE Int'l Conf. Computer Commun. (INFOCOM)*, pp. 1616–1624, Apr., 2015.
- [27] I. Koutsopoulos, "Optimal incentive-driven design of participatory sensing systems," in *Proc. IEEE Int'l Conf. Computer Commun. (INFOCOM)*, pp. 1402–1410, Apr., 2013.
- [28] Y. Wei, Y. Zhu, H. Zhu, Q. Zhang, and G. Xue, "Truthful online double auction for dynamic mobile crowdsourcing," in *Proc. IEEE Int'l Conf. Computer Commun. (INFOCOM)*, pp. 2074–2082, Apr., 2015.
- [29] D. Zhao, X.-Y. Li, and H. Ma, "How to crowdsourcing tasks truthfully without sacrificing utility: Online incentive mechanisms with budget constraint," in *Proc. IEEE Int'l Conf. Computer Commun. (INFOCOM)*, pp. 1213–1221, Apr., 2014.
- [30] S. He, D.-H. Shin, J. Zhang, and J. Chen, "Toward optimal allocation of location dependent tasks in crowdsensing," in *Proc. IEEE Int'l Conf. Computer Commun. (INFOCOM)*, pp. 745–753, Apr., 2014.
- [31] D. Moltchanov, "Distance distributions in random networks," *Ad Hoc Networks*, vol. 10, no. 6, pp. 1146–1166, Aug. 2012.
- [32] N. Mastrorarde and M. Van der Schaar, "Fast reinforcement learning for energy-efficient wireless communication," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6262–6266, Dec., 2011.