Mobile Crowdsensing Games in Vehicular Networks

Liang Xiao , *Senior Member, IEEE*, Tianhua Chen, Caixia Xie, Huaiyu Dai, *Fellow, IEEE*,
and H. Vincent Poor , *Fellow, IEEE*

Abstract—Vehicular crowdsensing takes advantage of the mobility of vehicles to provide location-based services in large-scale areas. In this paper, mobile crowdsensing (MCS) in vehicular networks is analyzed and the interactions between a crowdsensing server and vehicles equipped with sensors in an area of interest is formulated 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. A Q-learning-based MCS payment strategy and sensing strategy is proposed for the dynamic vehicular crowdsensing game, and a postdecision state learning technique is applied to exploit the known radio channel model to accelerate the learning speed of each vehicle. Simulations based on a Markov-chain channel model are performed to verify the efficiency of the proposed MCS system, showing that it outperforms the benchmark MCS system in terms of the average utility, the sensing accuracy, and the energy consumption of the vehicles.

Index Terms—Game theory, mobile crowdsensing (MCS), reinforcement learning, vehicular networks.

I. INTRODUCTION

WITH the rapid growth in the number of vehicles and smartphones equipped with sensors, such as cameras, microphones, and GPS devices, mobile crowdsensing (MCS) over vehicular networks can provide location-based services, such as urban monitoring [1], road and traffic condition monitoring [2], [3], pollution level measurements, wildlife habitat monitoring [4], and cross-space public information sharing [5]. MCS 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 described in [7], an 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 MCS [8]. Auctions and pricing strategies, such as those described in [9] and [10], stimulate mobile users to participate in crowdsensing applications. The discriminating payment strategies proposed in [11] can address faked sensing attacks in MCS.

However, to the best of our 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 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 of the game theoretic studies of the MCS, vehicular mobility makes it challenging for a vehicle to accurately estimate the sensing environment and the responses of other vehicles in a timely manner.

In this paper, we investigate the MCS in vehicular networks for both 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 those in [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 conditions under which an 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 cannot accurately observe the sensing costs of the vehicles and their radio channel conditions in time. Reinforcement learning techniques, such as Q-learning and

postdecision 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 improve the learning speed and increase the utility of the vehicle.

Compared with existing works, our major contributions include the following.

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 rate in MCS.

3) Simulations are performed to evaluate the performance of the dynamic vehicular MCS game for up to ten nonzero sensing levels based on a 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, and 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 described 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 the sensed data. The interactive music recommendation system developed in [15] applies MCS 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 predicted trajectories was developed in [18] for vehicle-based crowdsourcing. Television white space networks are utilized in [19] to provide robust and long-range connectivity for vehicles in the presence of power asymmetry. A distributed real-time pricing control technology was proposed for energy management of large-scale electrical vehicles in [20].

MCS systems based on smartphones can support a wide range of large-scale monitoring applications. Multidimensional context-aware social networks discussed in [21] enable context awareness to develop mobile crowdsourcing applications. The available individual reputation information was exploited in [22] to improve the performance of crowdsourcing. Cooperation and competition among service providers were studied in [23] to develop an incentive mechanism for MCS. The participatory application in [24] is resilient to the newcomer, on–off and collusion attacks. The crowdsourcing system developed in [25] can enhance disaster management. An efficient and privacy-aware data aggregation was proposed in [26] to aggregate time-series data without leaking mobile users' privacy. In the cloud-assisted image sharing system in [27], smartphones evaluate sensed images based on their significance and redundancy, and only upload valuable and unique images during disasters.

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 [28]. The interactions between the service user and MCS provider were formulated in [29] as an online double auction with two dynamic users. The online auction with users arriving in sequence analyzed in [30] maximizes the service value from the selected users under a budget constraint. The Nash bargaining solution of MCS is investigated in [31], 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 a Markov-chain cannel 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 described in [7], and perform simulations in dynamic games to evaluate their performance.

III. SYSTEM MODEL

A. Network Model

We consider a vehicular network consisting of M vehicles equipped with sensors, such as cameras and GPS receivers, 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 $x_i \in \mathbf{A} = \{0, 1, 2, \ldots, N_a\}$. 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 an evaluation algorithm such as that described in [9] to evaluate the accuracy of the sensing report and, thus, estimates the sensing effort of the corresponding mobile device. Based on an accurate evaluation algorithm, the server
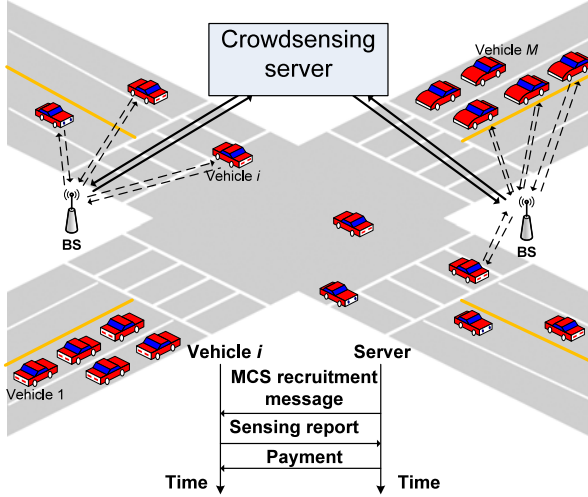
Fig. 1. Illustration of a vehicular crowdsensing system consisting of M vehicles that connect to the crowdsensing server via base stations.
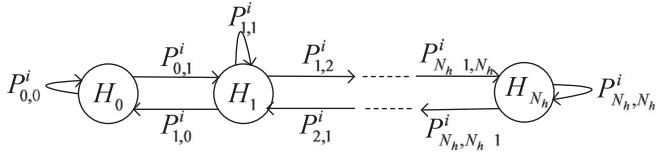

Fig. 2. Markov-chain channel model of vehicle i.

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 accept 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 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 is assumed to occur only between neighboring states.

TABLE I
SUMMARY OF SYMBOLS AND NOTATIONS

M	Number of vehicles
$x_i \in \{0, 1, \ldots, N_a\}$	Sensing effort of vehicle i
$\mathbf{y} = [y_j]_{0 \leq j \leq N_a}$	Payment strategy
$y_j \in \{b_i\}_{0 \leq j \leq N_b}$	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
β_i	Contribution of vehicle i in MCS
$\mathbf{s}/\mathbf{s}_i$	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
α	Learning rate
δ	Discount factor in the learning process
$\bar{\mathbf{s}}_i$	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

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

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

$$(1)$$

where γ 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 $\mathbf{x} = [x_i]_{1 \leq i \leq M}$, 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 in-

formation for an MCS application, such as a three-dimensional 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 described in [4] as an example; here, 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 CO_2, and derives better measurements by receiving more sensing reports in the area. A contribution factor $\beta_i > 0$ is introduced to weigh the contribution of vehicle 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 [32]. More specifically, if vehicle i is in a well-known location or several vehicles nearby have already sent sensing reports, β_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

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

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, which is given by

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

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 single vehicle with the best sensing condition to participate in the task, and suppresses 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 β_i weighs 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, depends on the vehicle with

the best sensing quality, and is given by

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

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

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

where $\mathbf{x}_{-i} = [x_j]_{1 \le j \ne i \le M}$ 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 $\mathbf{G} = \langle \{s, 1, 2, \ldots, M\}, \{\mathbf{y}, \mathbf{x}\}, \{u_s, u_{1 \le i \le M}\} \rangle$. 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_{Na}.

V. NE 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 NE of the game is denoted by $[\mathbf{x}^*, \mathbf{y}^*]$, where $\mathbf{x}^* = [x_i^*]_{0 \le i \le M}$ and $\mathbf{y}^* = [y_j^*]_{0 \le j \le N_a}$. By definition, an NE of the game $\mathbf{G}$ (if one exists) is given by

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

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

A. Accumulative Sensing Tasks

We consider the NEs of the static MCS game $\mathbf{G}$ for 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 silent (i.e., $x_i = 0$). From (2) and (7), 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 $\frac{1}{M} \sum_{i=1}^{M} \beta_i > \max_{1 \le i \le M} \frac{c_i}{h_i}$, then the Pareto optimal NE of the static vehicular crowdsensing game $\mathbf{G}$ for the accumulative sensing tasks with $N_a = 1$ is given by

$$x_i^* = 1, 1 \le i \le M$$
$$\mathbf{y}^* = \left[0, \max_{1 \le i \le M} \frac{c_i}{h_i} \right]. \qquad (8)$$

Proof: From (3), if $y_1^* \ge \frac{c_i}{h_i}$, we have $u_i^a(1, \mathbf{y}^*) = y_1^* - \frac{c_i}{h_i} \ge 0 = y_0^* = u_i^a(0, \mathbf{y}^*)$. Thus, if $y_1^* \ge \max_{1 \le i \le M} \frac{c_i}{h_i}$, (6) holds for $x_i^* = 1$, $\forall 1 \le i \le M$. From (2), u_s^a monotonically decreases with y_1, yielding $u_s^a([1, \ldots, 1], [0, y_1]) = \sum_{i=1}^{M} \beta_i - My_1 < \sum_{i=1}^{M} \beta_i - My_1^* = u_s^a([1, \ldots, 1], [0, y_1^*])$, $\forall y_1 > y_1^*$,

and from (7), we have $y_1^* = \max_{1 \le i \le M} \frac{c_i}{h_i}$. If $\frac{1}{M} \sum_{i=1}^{M} \beta_i > \max_{1 \le i \le M} \frac{c_i}{h_i}$, we have $u_s^a([1, \ldots, 1], [0, y_1^*]) > 0$. 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 $(\mathbf{x}^*, \mathbf{y}^*) \ne (\mathbf{x}', \mathbf{y}')$. 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. ∎

Remark: In this case, all M vehicles participate in the sensing tasks.

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 keeps 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

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

Proof: If $y_2^* \ge \frac{2c_i}{h_i}$, from (6) we have

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

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

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

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

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

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

Proof: The proof is similar to that of Proposition 1. ∎

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 costs of the vehicle, and the payment gap regarding the sensing quality is small, as given 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

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

Proof: The proof is similar to that of Proposition 1. ∎

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

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

$$x_i^* = 2, \quad 1 \le i \le M$$

$$\mathbf{y}^* = \left[0,\ 0,\ 2 \max_{1 \le i \le M} \frac{c_i}{h_i}\right]. \tag{14}$$

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

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 silent ($x_i = 0$), or upload a faked report ($x_i = -1$). It is clear that $\mathbf{y}^* = [0, 0, y_1^*]$, where y_1^* is the payment for an accurate report.

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

$$x_i^* = 1, \quad 1 \le i \le M$$

$$\mathbf{y}^* = \left[0,\ 0,\ 2 \max_{1 \le i \le M} \frac{c_i}{h_i}\right]. \tag{15}$$

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

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 given in (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 reports, 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 in order 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 $\beta_{i^*} > \frac{c_{i^*}}{h_{i^*}}$, $i^* = \arg\min_{1 \leq i \leq M} \frac{c_i}{h_i}$, then 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

$$x_i^* = \begin{cases} 1, & i = i^* \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

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

Proof: From (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 $u_{i^*}^b(1, \mathbf{y}^*) \geq 0 = u_i^b(0, \mathbf{y}^*)$. From (6), we see that $x_{i^*}^* = 1$. If $\beta_{i^*} > \frac{c_{i^*}}{h_{i^*}}$, from (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 $x_{i^*}^* = 0$. From (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. ∎

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

$$x_i^* = \begin{cases} 2, & i = i^* \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbf{y}^* = \left[0, 0, 2\frac{c_{i^*}}{h_{i^*}}\right]. \tag{18}$$

Proof: The proof is similar to that of Theorem 3. ∎

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

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

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 NEs 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 MCS. 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 [33] 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 the 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 ζ_j^{k-1}, and given by

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

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., $\mathbf{s}^k = \left[\zeta_j^{k-1}\right]_{1 \leq j \leq N_a} \in \mathbf{S}^a = \{0, 1, \ldots, M\}^{N_a}$. Similarly, for the best-quality sensing tasks, the system state at time k is given by $s^k = \max_{1 \leq i \leq M} x_i^{k-1} \in \mathbf{S}^b = \{0, 1, \ldots, N_a\}$. Fig. 3 shows 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.

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., $(\mathbf{s}/s)^{k+1} \in \mathbf{S}^{a/b}$. Let $Q(\mathbf{s}, \mathbf{y})$ denote the Q-function of the server at state $\mathbf{s}$ with action $\mathbf{y}$. According to the ϵ-greedy policy, with $0 < \epsilon \leq 1$, the server chooses the

---

Algorithm 1: Payment strategy with Q-learning.

1: **Initialize** α, δ, $\mathbf{s}^1 = \mathbf{0}$, $Q(\mathbf{s}, \mathbf{y}) = 0$, $V(\mathbf{s}) = 0$, $\forall \mathbf{s}$, $\mathbf{y}$
2: **for** $k = 1, 2, 3, \ldots$ **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)
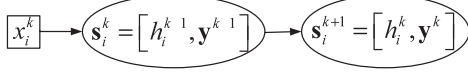8: **end for**

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

1: **Initialize** α, δ, $\mathbf{s}_i^1 = \mathbf{0}$, $Q_i(\mathbf{s}_i, x_i) = 0$, $V(\mathbf{s}_i) = 0$, $\forall \mathbf{s}_i$, x_i
2: **for** $k = 1, 2, 3, \ldots$ **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)
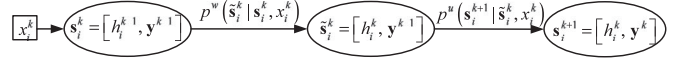7:   Update $V_i(\mathbf{s}_i^k)$ via (24)
8: **end for**

Fig. 4. State transition of vehicle i with a Q-learning-based sensing strategy in the dynamic MCS.

Fig. 5. State transition of vehicle i with a PDS-based sensing strategy in the dynamic MCS game.

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

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

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:

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

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

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

B. Sensing Based on Q-Learning

The sensing decision process of a vehicle can be modeled as an MDP. Thus, a vehicle can apply the Q-learning technique to derive its sensing strategy. The system state observed by vehicle i at time k is given by $\mathbf{s}_i^k = \left[h_i^{k-1}, \mathbf{y}^{k-1}\right] \in \mathbf{S}_i$, 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 shows the state transition for vehicle i in the sensing process.

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

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

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

Based on the ϵ-greedy policy, vehicle i chooses its sensing effort as

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

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

C. Sensing Strategy Based on PDS-Learning

The radio channel variation speed depends on the moving speed of the vehicle, as shown in (1). By exploiting the partially known channel model, a vehicle can apply the PDS-learning technique [33] to improve its learning speed, because less information has to be learned.

As shown in Fig. 5, vehicle i chooses its sensing effort $x_i^k \in \{0, 1, \ldots, N_a\}$ based on the system state $\mathbf{s}_i^k = \left[h_i^{k-1}, \mathbf{y}^{k-1}\right]$. According to PDS-learning, we define the PDS 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 $\tilde{\mathbf{s}}_i^k = \left[h_i^k, \mathbf{y}^{k-1}\right]$ with a known probability $p^w(\tilde{\mathbf{s}}_i^k | \mathbf{s}_i^k, x_i^k) = \Pr\left(h_i^k | h_i^{k-1}\right)$. Then, the next state $\mathbf{s}_i^{k+1} = \left[h_i^k, \mathbf{y}^k\right]$ 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 PDS, denoted by $\tilde{Q}_i(\tilde{\mathbf{s}}_i, x_i)$, is updated iteratively by

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

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

1: **Initialize** α, δ, $\mathbf{s}_i^1 = \mathbf{0}$, $Q_i(\mathbf{s}, x) = 0$, $V(\mathbf{s}) = 0$, $\forall \mathbf{s}$, x
2: **for** $k = 1, 2, 3, \ldots$ **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**

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

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

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 introduce the sensing quality index of the server, which is denoted by π and given by

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

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 $\alpha = 0.7$, $\delta = 0.8$, $\epsilon = 0.1$, $V = 5$, $N_a = 10$, $N_b = 25$, $N_h = 1$, $\mathbf{A} = \{0 \le i \le 10\}$, $\mathbf{B} = \{0 \le k \le 25\}$, $\varrho = \{1, 10\}$, and $0 \le M \le 80$. The contribution factors of the M vehicles were chosen from $[1, 20]$ at random, i.e., $\beta_i \in [1, 20]$, $\forall 1 \le i \le M$. We evaluate the ε-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.

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.

The performance of the vehicular crowdsensing game with $\beta_i = 40$, $\forall 1 \le i \le M$, is shown in Fig. 7, in which all the M vehicles choose their sensing efforts to maximize their
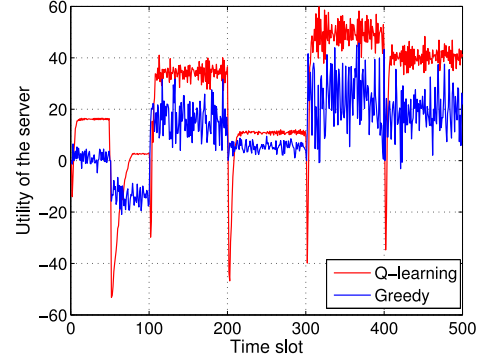

Fig. 6. Utility of the server with accumulative sensing tasks with $c = 10$ and $\beta_i \in [1, 20]$, $\forall 1 \le i \le 10$, in which the server chooses the greedy payment or the Q-learning-based payment, while each vehicle chooses the sensing effort to maximize its immediate utility.
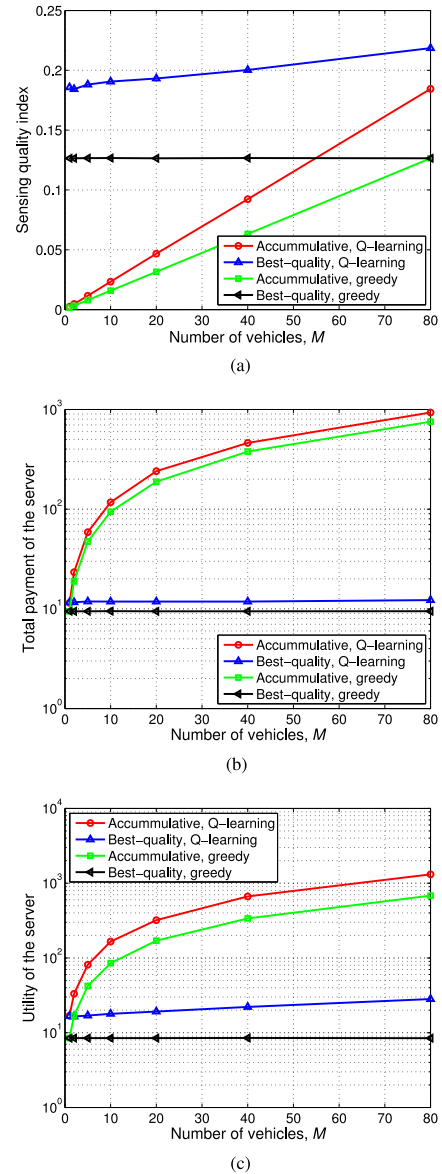

Fig. 7. Performance of the vehicular crowdsensing game with $c = 10$ and $\beta_i = 40$, $\forall 1 \le i \le M$, in which each vehicle chooses the sensing effort to maximize its immediate utility. (a) Average sensing quality index. (b) Average total payment of the server. (c) Average utility of the server.
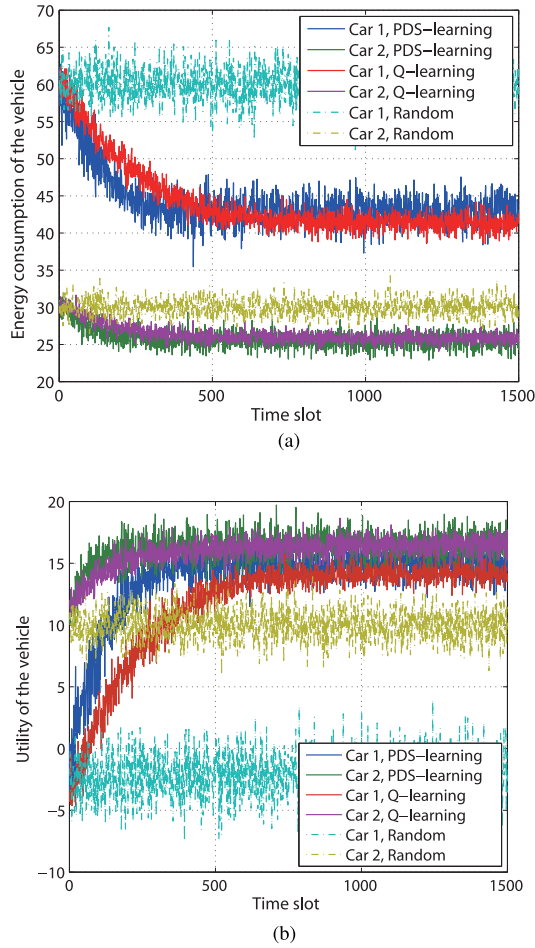
Fig. 8. Performance of vehicles in a dynamic vehicular crowdsensing game for an accumulative sensing task, with $M = 3$ vehicles, $c_1 = 10$, $c_2 = 5$, $c_3 = 1$, and $\beta_i \in [1, 20]$, $\forall 1 \leq i \leq 3$. (a) Energy consumption of the vehicle. (b) Utility of the vehicle.
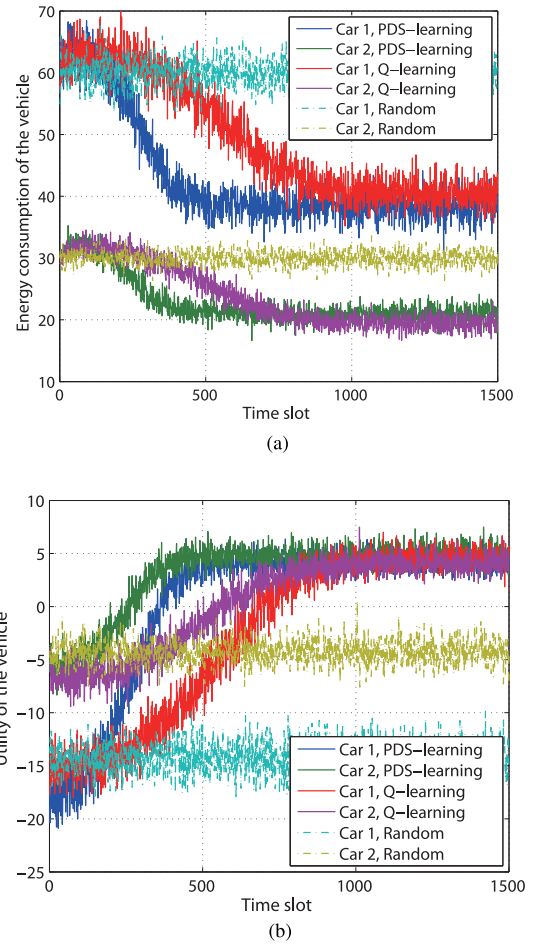

Fig. 9. Performance of vehicles in a dynamic vehicular crowdsensing game for a best-quality sensing task, with $M = 3$ vehicles, $c_1 = 10$, $c_2 = 5$, $c_3 = 1$, and $\beta_i \in [1, 20]$, $\forall 1 \leq i \leq 3$. (a) Energy consumption of the vehicle. (b) Utility of the vehicle.

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 $M = 80$.

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 MCS 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 the 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 Q-learning.

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 that of the Q-learning-based strategy.
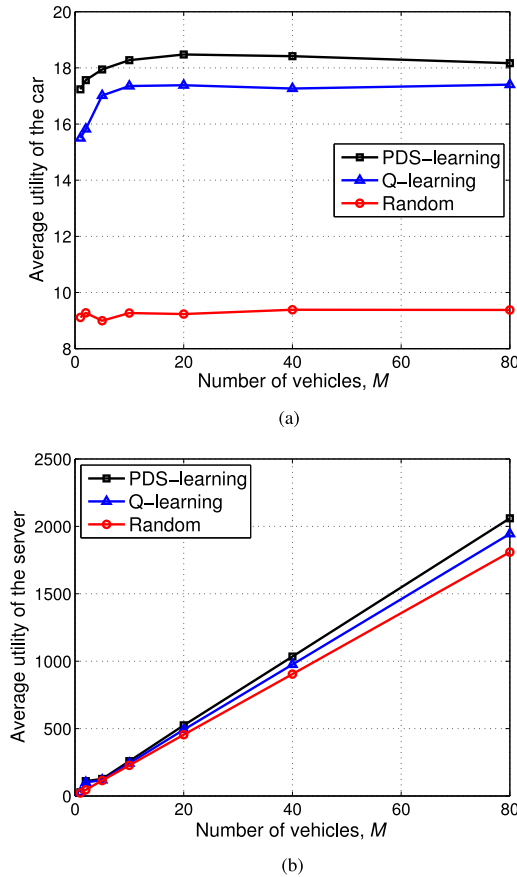
Fig. 10. Performance of the dynamic vehicular crowdsensing game with $c = 10$ and $\beta_i = 10$, $\forall 1 \leq i \leq M$ for the accumulative sensing task. (a) Average utility of the vehicle. (b) Utility of the server.

The performance of the dynamic game with an accumulative sensing task and the random payment with $c = 10$ and $\beta_i = 10$, $\forall 1 \leq i \leq M$, is shown in Fig. 10, showing that the proposed sensing strategies can improve the utility of the vehicle. As shown in Fig. 10(a), the Q-learning-based sensing strategy increases the average utility of the vehicle by 85%, and the PDS-learning-based sensing strategy further increases the utility by 93%, compared with the random benchmark. They also increase the server's utility, which increases with the number of participant vehicles. For example, in Fig. 10(b), the crowdsensing with Q-learning increases the server's utility from 491 to 1944 as M increases from 20 to 80, which is 7% higher than the benchmark and is further improved by 14% by PDS-learning-based sensing with 80 vehicles.

VIII. CONCLUSION

We have formulated a mobile crowdsensing game, in which each vehicle chooses its sensing effort according to its sensing and transmission costs and the expected payment, while the server pays each vehicle based on its sensing accuracy. The NEs of the static vehicular crowdsensing game have been derived for both accumulative sensing tasks and best-quality sensing tasks, and we have provided condition under which each NE exists. Reinforcement learning based MCS schemes have been

proposed for the server in the dynamic MCS game without being aware of the sensing model. Simulation results show that the Q-learning-based MCS scheme can significantly improve the utility of the MCS server. By exploiting the partially known radio channel model, the PDS-learning-based sensing strategy exceeds Q-learning-based sensing, with 5% higher utility for the vehicle in the dynamic game with 20 vehicles.

In order to successfully implement the proposed solution in practice, it is necessary to address several challenges. For example, the deployment of this solution requires repeated interactions between the MCS server and the vehicles during a time period with a constant cost and channel gain. Therefore, it is necessary to further accelerate the learning speed of the MCS system. In addition, as we have assumed accurate evaluation of the sensing reports at the server in this work, it is necessary to consider the evaluation error of the MCS server, possibly due to packet loss. We plan to address these issues in our future work in this area, and also to performance evaluation via experiments on actual vehicular networks.

REFERENCES

[1] U. Lee, B. Zhou, M. Gerla, E. Magistretti, P. Bellavista, and A. Corradi, "MobEyes: Smart mobs for urban monitoring with a vehicular sensor network," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 52–57, Oct. 2006.

[2] J. Eriksson *et al.*, "The pothole patrol: Using a mobile sensor network for road surface monitoring," in *Proc. ACM Int. Conf. Mobile Syst. Appl. Serv.*, Breckenridge, CO, USA, Jun. 2008, pp. 29–39.

[3] P. Mohan, V. Padmanabhan, and R. Ramjee, "Nericell: Rich monitoring of road and traffic conditions using mobile smartphones," in *Proc. ACM Conf. Embedded Netw. Sensor Syst.*, Raleigh, NC, USA, Nov. 2008, pp. 323–336.

[4] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: Current state and future challenges," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 32–39, Nov. 2011.

[5] B. Guo, H. Chen, Z. Yu, X. Xie, S. Huangfu, and D. Zhang, "FlierMeet: A mobile crowdsensing system for cross-space public information reposting, tagging, and sharing," *IEEE Trans. Mobile Comput.*, vol. 14, no. 10, pp. 2020–2033, Oct. 2015.

[6] W. Alasmary, H. Sadeghi, and S. Valaee, "Crowdsensing in vehicular sensor networks with limited channel capacity," in *Proc. IEEE Int. Conf. Commun.*, Budapest, Hungary, Jun. 2013, pp. 1833–1838.

[7] L. Xiao, T. Chen, C. Xie, and J. Liu, "Mobile crowdsensing game in vehicular networks," in *Proc. 2015 IEEE Region 10 Conf.*, Macau, Nov. 2015, pp. 1–6, invited talk.

[8] Z. Feng, Y. Zhu, Q. Zhang, L. Ni, and A. Vasilakos, "TRAC: Truthful auction for location-aware collaborative sensing in mobile crowdsourcing," in *Proc. IEEE Int. Conf. Comput. Commun.*, Toronto, ON, Canada, Apr. 2014, pp. 1231–1239.

[9] Y. Wen *et al.*, "Quality-driven auction based incentive mechanism for mobile crowd sensing," *IEEE Trans. Veh. Technol.*, vol. 64, no. 9, pp. 4203–4214, Sep. 2015.

[10] D. Yang, G. Xue, X. Fang, and J. Tang, "Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing," in *Proc. ACM Int. Conf. Mobile Comput. Netw.*, Istanbul, Turkey, Oct. 2012, pp. 173–184.

[11] L. Xiao, J. Liu, Q. Li, and H. V. Poor, "Secure mobile crowdsensing game," in *Proc. IEEE Int. Conf. Commun.*, London, U.K., Jun. 2015, pp. 7157–7162.

[12] M. Y. S. Uddin, H. Wang, F. Saremi, G. J. Qi, T. Abdelzaher, and T. Huang, "PhotoNet: A similarity-aware picture delivery service for situation awareness," in *Proc. IEEE Real-Time Syst. Symp.*, Vienna, Austria, Nov. 2011, pp. 317–326.

[13] X. He, H. Dai, and P. Ning, "Improving learning and adaptation in security games by exploiting information asymmetry," in *Proc. IEEE Int. Conf. Comput. Commun.*, Hong Kong, Apr. 2015, pp. 1787–1795.

[14] M. Dikaiakos, S. Iqbal, T. Nadeem, and L. Iftode, "VITP: An information transfer protocol for vehicular computing," in *Proc. ACM Int. Workshop Veh. Ad Hoc Netw.*, Cologne, Germany, Sep. 2005, pp. 30–39.

[15] X. Hu *et al.*, "SAfeDJ: A crowd-cloud codesign approach to situation-aware music delivery for drivers," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 12, no. 1, pp. 21:1–24, Oct. 2015.

[16] K. Lan, C. Chou, and H. Wang, "An incentive-based framework for vehicle-based mobile sensing," *Procedia Comput. Sci.*, vol. 10, no. 2, pp. 1152–1157, Aug. 2012.

[17] S. Santini, A. Salvi, A. Valente, A. Pescap, M. Segata, and R. L. Cigno, "A consensus-based approach for platooning with inter-vehicular communications," in *Proc. IEEE Int. Conf. Comput. Commun.*, Hong Kong, Apr. 2015, pp. 1158–1166.

[18] Z. He, J. Cao, and X. Liu, "High quality participant recruitment in vehicle-based crowdsourcing using predictable mobility," in *Proc. IEEE Int. Conf. Comput. Commun.*, Hong Kong, Apr. 2015, pp. 2542–2550.

[19] S. Sur and X. Zhang, "Bridging link power asymmetry in mobile whitespace networks," in *Proc. IEEE Int. Conf. Comput. Commun.*, Hong Kong, Apr. 2015, pp. 1176–1184.

[20] N. Rahbari-Asr, M.-Y. Chow, J. Chen, and R. Deng, "Distributed real-time pricing control for large-scale unidirectional V2G with multiple energy suppliers," *IEEE Trans. Ind. Informat.*, vol. 12, no. 5, pp. 1953–1962, May 2016.

[21] 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.

[22] A. Tarable, A. Nordio, E. Leonardi, and M. Marsan, "The importance of being earnest in crowdsourcing systems," in *Proc. IEEE Int. Conf. Comput. Commun.*, Hong Kong, Apr. 2015, pp. 2821–2829.

[23] X. Zhang, G. Xue, R. Yu, D. Yang, and J. Tang, "Truthful incentive mechanisms for crowdsourcing," in *Proc. IEEE Int. Conf. Comput. Commun.*, Hong Kong, Apr. 2015, pp. 2830–2838.

[24] X. Wang, W. Cheng, P. Mohapatraand, and T. Abdelzaher, "Artsense: Anonymous reputation and trust in participatory sensing," in *Proc. IEEE Int. Conf. Comput. Commun.*, Turin, Italy, Apr. 2013, , pp. 2517–2525.

[25] A. C. Weaver, J. P. Boyle, and L. Besaleva, "Applications and trust issues when crowdsourcing a crisis," in *Proc. 21st Int. Comput. Commun. Netw.*, Munich, Germany, Jul. 2012, pp. 1–5.

[26] Q. Li, G. Cao, and T. F. Porta, "Efficient and privacy-aware data aggregation in mobile sensing," *IEEE Trans. Depend. Sec. Comput.*, vol. 11, no. 2, pp. 115–129, Apr. 2014.

[27] Y. Hua, W. He, X. Liu, and D. Feng, "SmartEye: Real-time and efficient cloud image sharing for disaster environments," in *Proc. IEEE Int. Conf. Comput. Commun.*, Hong Kong, Apr. 2015, pp. 1616–1624.

[28] I. Koutsopoulos, "Optimal incentive-driven design of participatory sensing systems," in *Proc. IEEE Int. Conf. Comput. Commun.*, Turin, Italy, Apr. 2013, pp. 1402–1410.

[29] Y. Wei, Y. Zhu, H. Zhu, Q. Zhang, and G. Xue, "Truthful online double auction for dynamic mobile crowdsourcing," in *Proc. IEEE Int. Conf. Comput. Commun.*, Hong Kong, Apr. 2015, pp. 2074–2082.

[30] D. Zhao, X.-Y. Li, and H. Ma, "How to crowdsource tasks truthfully without sacrificing utility: Online incentive mechanisms with budget constraint," in *Proc. IEEE Int. Conf. Comput. Commun.*, Toronto, ON, Canada, Apr. 2014, pp. 1213–1221.

[31] S. He, D.-H. Shin, J. Zhang, and J. Chen, "Toward optimal allocation of location dependent tasks in crowdsensing," in *Proc. IEEE Int. Conf. Comput. Commun.*, Toronto, ON, Canada, Apr. 2014, pp. 745–753.

[32] D. Moltchanov, "Distance distributions in random networks," *Ad Hoc Netw.*, vol. 10, no. 6, pp. 1146–1166, Aug. 2012.

[33] N. Mastronarde 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.
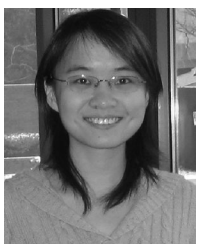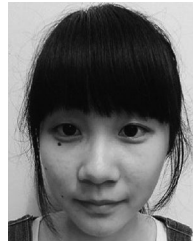
Liang Xiao (M'09–SM'13) received the B.S. degree in communication engineering from the Nanjing University of Posts and Telecommunications, Nanjing, China, the M.S. degree in electrical engineering from Tsinghua University, Beijing, China, and the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, USA, in 2000, 2003, and 2009, respectively.

She is currently a Professor in the Department of Communication Engineering, Xiamen University, Fujian, China. Her current research interests include smart grids, network security, and wireless communications.

Tianhua Chen received the B.S. degree in communication engineering in 2014 from Xiamen University, Xiamen, China, where she is currently working toward the M.S. degree in communication engineering.

Her research interests include network security and wireless communications.

Caixia Xie received the B.S. degree in communication engineering in 2015 from Xiamen University, Xiamen, China, where she is currently working toward the M.S. degree in communication engineering.

Her research interests include network security and wireless communications.

Huaiyu Dai (F'17) received the B.E. and M.S. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1996 and 1998, respectively, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 2002.

He was with Bell Labs, Lucent Technologies, Holmdel, NJ, USA, in summer 2000, and with AT&T Labs-Research, Middletown, NJ, USA, in summer 2001. He is currently a Professor of electrical and computer engineering at North Carolina State University, Raleigh, NC, USA. His research interests include the general areas of communication systems and networks, advanced signal processing for digital communications, and communication theory and information theory. His current research interests include networked information processing and crosslayer design in wireless networks, cognitive radio networks, wireless security, and associated information-theoretic and computation-theoretic analyses.

Prof. Dai was an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS, SIGNAL PROCESSING, AND WIRELESS COMMUNICATIONS. . He coedited two special issues of the EURASIP journals on distributed signal processing techniques for wireless sensor networks, and on multiuser information theory and related applications, respectively. He was the Co-Chair of the Signal Processing for Communications Symposium of IEEE GLOBECOM 2013, the Communications Theory Symposium of IEEE ICC 2014, and the Wireless Communications Symposium of IEEE GLOBECOM 2014. He is currently an Area Editor in charge of wireless communications for the IEEE TRANSACTIONS ON COMMUNICATIONS.

H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, USA, in 1977.

From 1977 to 1990, he was on the faculty of the University of Illinois at Urbana-Champaign, Urbana, IL, USA. Since 1990, he has been on the faculty at Princeton, where he is currently the Michael Henry Strater University Professor of Electrical Engineering. From 2006 to 2016, he was the Dean of Princeton's School of Engineering and Applied Science. His research interests include the areas of information theory, statistical signal processing and stochastic analysis, and their applications in wireless networks and related fields. Among his publications in these areas is the book *Mechanisms and Games for Dynamic Spectrum Allocation* (Cambridge University Press, 2014).

Dr. Poor is currently a member of the National Academy of Engineering and the National Academy of Sciences, and a foreign member of the Royal Society. He is also a Fellow of the American Academy of Arts and Sciences, the National Academy of Inventors, and other national and international academies. He received the Marconi and Armstrong Awards of the IEEE Communications Society in 2007 and 2009, respectively. Recent recognition of his work includes the 2016 John Fritz Medal, the 2017 IEEE Alexander Graham Bell Medal, a Doctor of Science *honoris causa* from Syracuse University (2017), and Honorary Professorships at Peking University and Tsinghua University, both conferred in 2016.