Game Theoretic Analysis of Urban E-Taxi Systems: Equilibria and Efficiency

Yukun Yuan, Yue Zhao, Lin Chen*, and Shan Lin
Department of Electrical and Computer Engineering, Stony Brook University,
*School of Computer Science and Engineering, Sun Yat-sen University

***Abstract*—With increasing deployment of electric vehicles in urban mobility-on-demand systems, electric taxis (e-taxi) drivers need to compete with each other not only for passengers but also for limited charging points due to frequent and time-consuming charging activities. This paper focuses on two crucial research questions in this context: (1) What is the strategy of each e-taxi driver for charging and searching passengers in a non-cooperative environment, and what is the collective system outcome of competing e-taxis? (2) How can the mobility-on-demand service platforms (e.g., Uber and Lyft) push self-interested e-taxi drivers to improve the overall system efficiency. Technically, we study the non-cooperative mobility-on-demand system consisting of e-taxis from a game theoretic perspective. We formulate a mobility-on-demand system with competition among drivers as a stochastic game, analyze the Nash Equilibrium (NE) of the game, and design an approximation algorithm to obtain the NE. Moreover, we show that the NE is not necessarily efficient for the platform and propose a pricing scheme from the platform's perspective which induces the new NE to be efficient. We use a trace-driven simulation to evaluate the design based on datasets consisting of more than 7,000 fuel vehicles and nearly 700 e-taxis, 37 working charging stations, and more than 60,000 passenger trips per day. We show that, compared with the state-of-the-art which optimizes the system efficiency by coordinating e-taxis but is not an equilibrium, the NE achieves a system efficiency of merely 73.5% of that of the cooperative state-of-the-art, and the designed pricing scheme improves the price of anarchy to 95.5%.**

I. Introduction

In recent years, mobility-on-demand services, e.g., Uber, Didi, and Lyft have emerged to offer new transportation services for passengers. In such systems, every self-interested driver owns and operates a vehicle to maximize his/her daily profit. In this work, we study the scenarios when electric taxis (e-taxis) are operated with such a mobility-on-demand platform: self-interested drivers determine the cruising routes and charging locations of their e-taxis to maximize their respective daily profits. Meanwhile, they upload their real-time locations and occupancy status to the mobility-on-demand service platform, and the platform matches the passengers with the nearby unoccupied e-taxis. The platform takes a fixed percentage cut from the fare cost of rides as its income. The total income obtained from all fare costs is used to measure the efficiency of the platform [1].

This work was supported in part by the National Science Foundation under Grant ECCS-2025152, CNS-1952096 and CNS-1553273 (CAREER). Part of the work of Lin Chen was supported by National Science Foundation of China (NSFC) under Grant 62172455, Guangdong Basic and Applied Basic Research Foundation under Grant 2022A1515012309, and the Pearl River Talent Program under Grant 2019QN01X140.

New problems arise with the expansion of mobility-on-demand systems consisting of electric taxis. Firstly, the operational sustainability of an e-taxi requires relatively frequent and time-consuming charging in a day since the range of an EV is less than that of a fuel vehicle and it can cost up to an hour to fully charge an EV [2], [3]. Meanwhile, given the distribution of finite charging points over a city, self-interested e-taxi drivers need to compete with each other not only for passengers but also for charging points. Two crucial research questions arise in this context:

- **Q1**. What is the strategy of each driver for charging and searching passengers in a non-cooperative environment, and what is the collective system outcome of competing e-taxis?
- **Q2**. How can a mobility-on-demand service platform induce self-interested e-taxi drivers to improve the system-wise efficiency?

To answer Q1, we formulate the competition among e-taxi drivers for passengers and charging points as a non-cooperative stochastic game. The game models e-taxi system dynamics as stochastic processes due to randomness, e.g., passenger mobility pattern, traffic condition, and e-taxis' waiting time at charging stations. The Nash Equilibrium (NE) of the game is analyzed and an efficient algorithm approximately computing the NE is proposed. For Q2, recognizing the potential system inefficiency of the NE, a pricing scheme is designed for the mobility-on-demand service platforms to induce e-taxis to a socially optimal NE. The reason for utilizing the pricing scheme to engage drivers is that the service platforms cannot directly control the e-taxis to enhance the system-wise efficiency, and that penalty or reward changes drivers' utility and hence indirectly influences their actions.

Methodologically, modeling the competitive taxi systems as a game is not new [4]–[8]. [5] investigates the problem of how an e-taxi driver relocates for charging and serving passengers to maximize the long-term cumulative reward and designs a reinforcement learning algorithm for each driver. Different from these works, this paper studies a new setting, i.e., competition among e-taxi drivers with mobility-on-demand platforms. Dynamic fare schemes for taxi systems have also been studied before. E.g., [9] models drivers' strategic decisions as a game and proposes a time-dependent fare structure pushing taxi drivers to work during the peak time by increasing fare price. In contrast, our work addresses the platforms' inefficiency due to drivers' self-interested actions by penalizing or rewarding drivers, whereas, the related works either focus on improving drivers' utility [4]–[8] or design

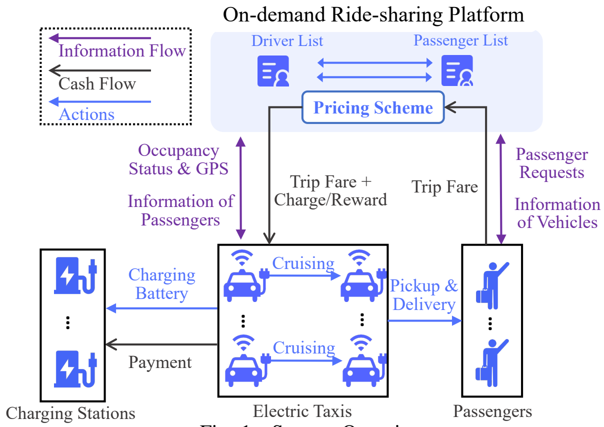

Fig. 1. System Overview

dynamic trip fare to incentivize drivers to work in hours when taxi supply is low [9], [10], which do not resolve the system inefficiency due to drivers' non-cooperative behaviors.

The contributions of this work are summarized as follows.

- To the best of our knowledge, we are the first to study the e-taxi competition in a mobility-on-demand system with a game-theoretic perspective. Specifically, we model the competition among e-taxi drivers as a stochastic game, and prove that it is a stochastic potential game. Given the intractable computation complexity with large state and action spaces, an effective algorithm is designed to approximately compute the NE among drivers.
- We explore the potential system inefficiency of the NE from the perspective of a mobility-on-demand service platform. Specifically, we design a pricing scheme that penalizes and rewards drivers to change their utility and influence their decisions, resulting in new NE of the competitive game among e-taxi drivers. It is proved that the new NE maximizes the efficiency of the platform.
- Our simulation utilizes datasets of 7,000+ fuel vehicles and nearly 700 e-taxis, 37 charging stations, and 60,000+ passenger trips per day. The results demonstrate that, compared with the state-of-the-art which optimizes the system efficiency by coordinating e-taxis but is not an equilibrium, the NE achieves a system efficiency of merely 73.5% of that of the cooperative state-of-the-art, and the proposed pricing scheme improves the price of anarchy to 95.5%.

II. Background and System Overview

An overview of the mobility-on-demand system is shown in Figure 1. In a modern mobility-on-demand system, electric taxis use the equipped wireless sensing devices to communicate with the online platform and periodically upload their information, e.g., the real-time occupancy status, and GPS locations. They also visit charging stations distributed in the city to charge their batteries to support their daily operations. Passengers send the service requests (e.g., origin, destination, and pick-up time) to the online platform when they need the transportation service, and then they can see the estimated ride fares. After receiving the requests, the platform matches each service request with a nearby unoccupied taxi using order dispatch strategies, e.g., combinatorial optimization algorithms [11], and then sends the information of passengers (vehicles) to the vehicles (passengers). The online platform usually takes a fixed percentage cut from the fare cost of a ride as its income.

An electric taxi driver makes following decisions, i.e., cruising routes and charging decisions. Since only nearby unoccupied e-taxis are considered for serving the passengers, the drivers should efficiently plan their cruising routes to increase the probability of matching with passengers. On the other hand, due to the limited driving range of an electric taxi, the e-taxi drivers need to charge the battery several times a day to support their operation [12], [13]. Meanwhile, the number of charging points is limited for e-taxis in a city, especially in the central business areas with high passenger demand [12]. A driver also needs to carefully make the charging decisions to maximize the time duration of working on the road.

The primary goal of the on-demand ride-sharing platform is to maximize its income, which is proportional to the total fare cost of rides that are served. Hence the platforms should understand whether the self-interested actions of drivers are system-wise efficient, and if not, how to enhance the system efficiency. With the consideration of competition among drivers for passengers and charging points, we model a mobility-on-demand system as a competitive game among electric taxi drivers. We analyze and show that the Nash Equilibrium (NE) solutions of the e-taxi game do not achieve system optimum for the platform. We then design a pricing scheme for the platform, which penalizes or rewards the drivers based on their actions by adding to or subtracting from the standard trip fare to induce the new NE to achieve efficiency for the platform.

We note that such a pricing scheme is essential for an online platform to achieve system efficiency. Because the platform cannot force self-interested drivers to take system-wise optimal actions that achieve the best platform-wise quality of service in the real world. The fare cost of rides is transferred from the passengers to the platform first, and then from the platform to the drivers, providing the opportunity to apply pricing schemes to indirectly influence the drivers' actions. In the real world, many platforms have designed dynamic fare cost of rides to influence drivers [9], [10]: e.g., Uber increases the fare of rides in the areas with high passenger demand to attract the e-taxis. However, such schemes may increase the level of competition in these areas, i.e., over-supply, resulting in lack of rides in other regions and system inefficiency. Our work provides a theoretical analysis and data-driven evaluation.

III. System Model and E-taxi Game

In this section, we model a mobility-on-demand system with e-taxis as a stochastic game. The dynamics of the system are partly under the control of e-taxis' actions and partly influenced by random factors such as passenger mobility patterns, urban traffic conditions, and the waiting time of the e-taxis at charging stations, etc. These features of the system are modelled by the following stochastic game.

A city is partitioned into M regions based on, e.g., administrative sub-districts [14] and grid file [15], where $\mathcal{M}$ represents the set of regions. We discretize a day into a number of time slots, indexed by t. More specifically, for any two regions i and i', there are passengers that request taxi service for traveling from region i to i' at time t, and the amount of passengers is denoted by $d^t_{i,i'}$. Let $\mathbf{d}^t = \{d^t_{i,i'}\}_{i,i' \in \mathcal{M}}$ denote the passenger demand of a city during slot t. Let $\mathcal{N}$ be the set of charging stations deployed ($|\mathcal{N}| = N$) in the city, and there are p_j chargers installed in charging station j.

The status of an e-taxi is defined as: (i) vacant: an e-taxi is idly driving on the road and searching next passengers; (ii) waiting for charging: an e-taxi is waiting at a charging station for a free charging point; (iii) charging: an e-taxi is charging its battery at a charging station; (iv) occupied: an e-taxi is delivering passengers to the destination. It is noted that similarly to [13], we allow e-taxis to quit the waiting queue at a charging station to become vacant again for serving passengers if this action introduces more utility to itself.

The definitions of players, states, actions, state transition functions, and utility functions are given as follows.

Players: An e-taxi is a player. As such, there are m players.

System State $\mathbf{s}^t \in \mathcal{S}$: The joint state of an e-taxi system at the beginning of slot t is the concatenation of all e-taxis' state i.e., $\mathbf{s}^t = \{s^t_1, ..., s^t_m\}$, where s^t_k represents the state of e-taxi k at the beginning of slot t. $s^t_k \in \mathcal{S}_k$ includes the working status, location, and remaining energy. Let $status^t_k$ denote the working status of e-taxi k at the beginning of slot t, which is one of four possible working statuses. Let $loc^t_k \in \mathcal{M} \cup \mathcal{N}$ represent the location of e-taxi k at the beginning of slot t. We discretize the remaining energy of an e-taxi into L levels and use $energy^t_k \in [1, L]$ to describe the remaining energy of an e-taxi. Accordingly, the state of e-taxi k is defined as $s^t_k = (status^t_k, loc^t_k, energy^t_k)$.

Action $\mathbf{a}^t \in \mathcal{A}$: a joint action $\mathbf{a}^t = \{a^t_1, ..., a^t_m\}$ demonstrates the actions of all m e-taxis at time slot t. The set of actions for any e-taxi k is the same, denoted as $\mathcal{A}_k = \mathcal{M} \cup \mathcal{N} \cup \{continuing\}$, and $a^t_k \in \mathcal{A}_k$ represents the action of the e-taxi k during slot t. If the e-taxi k is unoccupied, the action $a^t_k \in \mathcal{M} \cup \mathcal{N}$ is going to a region for serving passengers or driving to a charging station for charging the battery. If an e-taxi is occupied, the action $a^t_k \in \{continuing\}$ is to deliver the passenger to the destination. We use $I_i(\mathbf{a}^t)$ ($I_j(\mathbf{a}^t)$) to denote the number of e-taxis going to region i (or charging station j) based on $\mathbf{a}^t$.

Notably, an e-taxi's action, a^t_k, is constrained by its state, s^t_k. Firstly, due to the limited driving speed and duration of a time slot, the driving distance of an e-taxi during a time slot is limited. An e-taxi cannot go to a region or charging station that cannot be reached within a fixed time period (e.g., a time slot) from its current location. In reality, a driver can go to a region that is reached by several time slots. In this case, the decision of the driver at each time slot can be the intermediate regions in the route. Secondly, an e-taxi cannot go to a region for serving passengers or a charging station when it is occupied. Finally, if the remaining energy of an unoccupied e-taxi is low, it must go to a charging station to avoid using up energy on the road. According to the three constraints, we define $w(s^t_k, a^t_k) : \mathcal{S}_k \times \mathcal{A}_k \rightarrow \{0, 1\}$ to denote whether the action a^t_k can be taken if the state of the e-taxi k is s^t_k, where $w(s^t_k, a^t_k) = 1$ if a^t_k can be taken; otherwise, it is 0.

State transition probability $\phi(\mathbf{s}^{t+1}|\mathbf{s}^t, \mathbf{a}^t) : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow [0, 1]$: It describes the probability of transiting to $\mathbf{s}^{t+1}$ given the joint action $\mathbf{a}^t$ in the system state $\mathbf{s}^t$. The state transition probability function of e-taxi k is denoted as $\phi_k(s^{t+1}_k|\mathbf{s}^t, \mathbf{a}^t)$.

Utility function $R_k(\mathbf{s}^t, \mathbf{a}^t) : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$: each e-taxi is associated with a utility function R_k when m e-taxis take the joint action $\mathbf{a}^t$ during slot t for given the system state $\mathbf{s}^t$.

The intuition behind the expression of $R_k(\mathbf{s}^t, \mathbf{a}^t)$ are threefold as described by the conditions **C1**, **C2**, and **C3** as follows.

(i) The vacant e-taxi k goes to a region for serving passengers. The platform matches the e-taxi with a passenger request in the same region to reduce the passenger waiting time. There are different types of matching algorithms, we employ a random matching algorithm for simplicity. According to [4] and [16], if the e-taxi k goes to region i, we define its utility as **C1** : $R_k(\mathbf{s}^t, \mathbf{a}^t) = r^t_i - \tau^t_i \cdot I_i(\mathbf{a}^t)/d^t_i$, where r^t_i is the average monetary reward of a trip starting at region i during slot t. The e-taxi supply over passenger demand in a region, i.e., $I_i(\mathbf{a}^t)/d^t_i$, represents the level of competition for passengers. τ^t_i is a parameter denoting the cost of waiting for a passenger with per unit level of competition in region i during slot t, and it can be learned from the historical e-taxi trajectory data. The main idea is that the utility of an e-taxi is equal to the reward from serving passengers minus the potential loss of reward due to competition.

(ii) The e-taxi k selects to go to a charging station for charging the battery. Similarly to (i), we define the utility of the e-taxi k going to charging station j as **C2** : $R_k(\mathbf{s}^t, \mathbf{a}^t) = -\big(\rho^t_j \cdot (1 - \hat{\tau}^t_j \cdot I_j(\mathbf{a}^t)/p_j)\big)$, where ρ^t_j is the electricity price at the charging station j during slot t, and e is the amount of energy that is charged during a time slot. $I_j(\mathbf{a}^t)/p_j$ is the level of competition for charging points. $\hat{\tau}^t_j$ is a parameter denoting the cost of waiting for a charging point with per unit level of competition at charging station j during slot t, which can be learned from historical e-taxi trajectory data.

(iii) An e-taxi k is occupied and can only take the action "continuing". The utility of an occupied e-taxi is 0, i.e., **C3** : $R_k(\mathbf{s}^t, \mathbf{a}^t) = 0$. In the platform, the estimated fare of each ride is calculated and shown to the drivers when passenger requests are matched. Therefore, the utility for a ride is added to the taxi when its driver picks up the passenger.

We note that the actual utility a taxi gets can be different from the above as post-action realizations of the uncertainties take place. Nonetheless, at the time an e-taxi makes a decision, the above formulation is a priori representation of its knowledge of the posterior utility it will get. This can be understood as the utility from *the players' perspective* at the time of decision making.

We assume that an e-taxi considers the future H time slots as the time horizon for making decisions. Without loss of generality, the first slot is defined as the current time slot.

Hence, each e-taxi k needs to determine the actions, $a_k^{1:H}$ for the future H slots. To simplify the description, we use $\mathbf{a}_k$ to represent the actions of e-taxi k during the future H slots, and $\mathbf{a}_{-k}$ to denote the actions of the e-taxis except k. E-taxi k aims to maximize the expected long-term cumulative utility given the current system state $\mathbf{s}^1$, which is defined as follows.

$$V_k(\mathbf{s}^1, \mathbf{a}_k, \mathbf{a}_{-k}) = \mathbb{E}\big[\sum_{t=1}^{H} \beta^t R_k(\mathbf{s}^t, \mathbf{a}^t)\big] \tag{1}$$

where $\beta \in (0, 1]$ is the discount factor for the future utility.

Strategy: The strategy of an e-taxi during the future H time slots is non-stationary, i.e., it changes over the time, i.e., $\pi_k^{t_1} \neq \pi_k^{t_2}$. In the stochastic game of e-taxis, the strategy of e-taxi k is denoted by $\pi_k = \{\pi_k^1, ..., \pi_k^H\}$. A strategy $\pi_k^t(\mathbf{s}^t, a_k^t) : \mathcal{S} \times \mathcal{A}_k \to \{0, 1\}$ maps from the state of e-taxis at the beginning of slot t and an action of the e-taxi k for this slot t to a binary decision of whether this action should be taken. Based on the constraints of an e-taxi's actions for given its state s_k^t, we also have some constraints on the strategy. First, any e-taxi k should not select the action that is not feasible for given its state s_k^t, i.e., $\pi_k^t(\mathbf{s}^t, a_k^t) = \pi_k^t(s_k^t, \mathbf{s}_{-k}^t, a_k^t) \leq w(s_k^t, a_k^t)$.

Furthermore, for a time slot, an e-taxi can only select one action, i.e., $\sum_{a_k^t \in \mathcal{A}_k} \pi_k^t(\mathbf{s}^t, a_k^t) = 1$.

Given the strategies of all m e-taxis, denoted as $\pi = \{\pi_k^t\}_{1 \leq k \leq m, 1 \leq t \leq H}$ and the initial state $\mathbf{s}^1$, the utility of e-taxi k is rewritten as $J_k^{\pi}(\mathbf{s}^1) = \mathbb{E}\big[\beta^t \sum_{t=1}^{H} R_k(\mathbf{s}^t, \mathbf{a}^t)|\pi\big]$.

IV. Analysis of the Stochastic E-taxi Game

In the non-cooperative game, an important question is to predict the outcome of the game for which the NE is a first-order approximation. It is thus important to compute the NE if it exists. Notably, compared with the related resource competition game [4], [17], [18], in the stochastic e-taxi game, each player makes sequential decisions over time, and the NE is defined among the individual e-taxi drivers as opposed to the aggregate e-taxi distribution.

The organization of this section is as follows. We firstly propose the definition of Nash Equilibrium (NE) and stochastic potential games. Then we describe three lemmas, showing that the stochastic e-taxi game is a potential game when only considering one future time slot. Furthermore, we prove that the stochastic e-taxi game considering future H slots is a stochastic potential game in Theorem 1. Finally, in Theorem 2, we show that the NE of the stochastic e-taxi game can be obtained by solving the optimization problem, which aims to maximize the potential function.

Definition 1 (Nash Equilibrium)**.** *In a stochastic game, a Nash Equilibrium is an optimal strategy $\pi = \{\pi_k\}_{k=1}^m$, such that for all initial state $\mathbf{s}^1$, if any e-taxi k changes the strategy from π_k to $\hat{\pi}_k$, the following holds:*

$$J_k^{\pi_k, \pi_{-k}}(\mathbf{s}^1) \geq J_k^{\hat{\pi}_k, \pi_{-k}}(\mathbf{s}^1) \tag{2}$$

Definition 2 (Stochastic potential game [19], [20])**.** *A stochastic game is called an exact Stochastic potential game if there exists a potential function Φ for $\forall k$, $\forall \mathbf{s}^1 \in \mathcal{S}$ such that*

$$J_k^{\pi_k, \pi_{-k}}(\mathbf{s}^1) - J_k^{\hat{\pi}_k, \pi_{-k}}(\mathbf{s}^1) = \Phi^{\pi_k, \pi_{-k}}(\mathbf{s}^1) - \Phi^{\hat{\pi}_k, \pi_{-k}}(\mathbf{s}^1) \tag{3}$$

Since there are typically a large number of taxis in metropolitan areas, e.g., there are more than 13,000 yellow taxis in NYC, we consider the e-taxi game as a non-atomic game [9], [21], meaning that the impact of a single e-taxi's action on the e-taxi system is negligible [9]. We then have the following two lemmas.

Lemma 1. *For any e-taxi k, the joint state transition probability of the other $m-1$ e-taxis depends on their initial joint states and actions, and the probability does not depend on the action of the e-taxi k, i.e., $\phi(\mathbf{s}_{-k}^{t+1}|s_k^t, \mathbf{s}_{-k}^t, a_k^t, \mathbf{a}_{-k}^t) = \phi(\mathbf{s}_{-k}^{t+1}|\bar{s}_k^t, \mathbf{s}_{-k}^t, \bar{a}_k^t, \mathbf{a}_{-k}^t)$ holds for all $\bar{s}_k^t$ and $\bar{a}_k^t$.*

Lemma 2. *For any e-taxi k, the joint strategy of the other $m-1$ e-taxis outputs the same joint actions if the joint states of these $m-1$ e-taxis do not change, i.e., $\pi_{-k}^t(\mathbf{s}_{-k}^t, s_k^t, \mathbf{a}_{-k}^t) = \pi_{-k}^t(\mathbf{s}_{-k}^t, \bar{s}_k^t, \mathbf{a}_{-k}^t)$ holds for all $\bar{s}_k^t$.*

Based on the two lemmas, we propose the potential function for the stochastic e-taxi game when the future time horizon is only one time slot in the following lemma. Due to page limit, we have assembled the proof of all the lemmas and theorems in the technical report [22].

Lemma 3. *Consider the function $F(s^t, \mathbf{a}^t)$ defined as*

$$\begin{aligned} F(\mathbf{s}^t, \mathbf{a}^t) = \tfrac{1}{2}\Big(&\sum_{i=1}^{M} I_i(\mathbf{a}^t)\big(f_r(d_i^t, I_i(\mathbf{a}^t)) + f_r(d_i^t, 1)\big) \\ &+ \sum_{j=1}^{N} I_j(\mathbf{a}^t)\big(f_c(p_j, I_j(\mathbf{a}^t)) + f_c(p_j, 1)\big)\Big) \end{aligned} \tag{4}$$

$$f_r(d_i^t, c) = r_i^t - \tau_i^t c/d_i^t, \quad f_c(p_j, c) = -(\rho_j^t e - \hat{\tau}_i^t c/p_j).$$

where $I_i(\mathbf{a}^t)$ is the number of e-taxis that go to region i during slot t given their joint action $\mathbf{a}^t$, and $I_j(\mathbf{a}^t)$ is the number of e-taxis that go to charging station j during slot t given their joint action $\mathbf{a}^t$. For any $\mathbf{s}^t$, the following equation holds for any e-taxi k if the e-taxi k changes its strategy from π_k to $\hat{\pi}_k$:

$$\begin{aligned} &R_k(\mathbf{s}^t, \hat{a}_k^t, \mathbf{a}_{-k}^t) - R_k(\mathbf{s}^t, a_k^t, \mathbf{a}_{-k}^t) \\ &= F(\mathbf{s}^t, \hat{a}_k^t, \mathbf{a}_{-k}^t) - F(\mathbf{s}^t, a_k^t, \mathbf{a}_{-k}^t), \end{aligned} \tag{5}$$

where $\pi_k^t(\mathbf{s}^t, a_k^t) = 1$ and $\hat{\pi}_k^t(\mathbf{s}^t, \hat{a}_k^t) = 1$.

Based on the above, we now have the following theorem.

Theorem 1. *The non-atomic non-cooperative game among e-taxis is a stochastic potential game with a potential function defined as:*

$$\Phi^{\pi_k, \pi_{-k}}(\mathbf{s}^1) = \sum_{t=1}^{H} \mathbb{E}[\beta^t F(\mathbf{s}^t, \mathbf{a}^t)|\pi_k, \pi_{-k}] \tag{6}$$

Given the potential function (6), we define the following optimization problem:

$$\begin{aligned} \max\nolimits_{\pi} \quad & \Phi^{\pi}(\mathbf{s}^1) = \sum_{t=1}^{H} \mathbb{E}[\beta^t F(\mathbf{s}^t, \mathbf{a}^t)|\pi, \phi] \\ \textbf{s.t.} \quad & \pi_k^t(\mathbf{s}^t, a_k^t) \leq w(s_k^t, a_k^t) \ \ \sum_{a_k^t \in \mathcal{A}_k} \pi_k^t(\mathbf{s}^t, a_k^t) = 1 \end{aligned} \tag{7}$$

Notably, the optimization is over the *policies* $\pi_k^t(\mathbf{s}^t, a_k^t), \forall t \in [1, H], k \in [1, m], \mathbf{s}^t \in \mathcal{S}$. As such, the number of decision variables is $mH|\mathcal{S}|(M + N + 1)$. In addition, the constraints of the decision variables are all linear. As a result, the above problem is a stochastic linear optimization problem. We now have the following theorem that links the optimal solution of the above optimization problem to the NE of the game.

Theorem 2. *The solution of the optimization problem* (7) *is the NE among the m e-taxis over the future H time slots.*

The above stochastic optimization problem is however not computationally tractable due to the curse of dimensionality in the joint state and action space. Next, we investigate efficient computation methods to accurately approximate the NE.

V. Approximate Computation of Nash Equilibrium

In this section, we propose an efficient method to approximately compute the NE of the stochastic potential game of e-taxis. First, we formulate a Markov Decision Process (MDP) by aggregating e-taxis and prove that the optimal strategy of the MDP can be mapped to an NE of the stochastic potential game. Secondly, we approximate the optimal strategy of the MDP by sampling a new MDP with a smaller size of actions and next-states and solving this new MDP.

A. Markov Decision Process via Aggregating E-taxis

We formulate a new MDP by aggregating e-taxis as follows.

State: We use $\tilde{\mathbf{s}}^t = \{\mathbf{V}^t, \mathbf{O}^t\} \in \tilde{\mathcal{S}}$ to denote the system state at the beginning of slot t by aggregating e-taxis. In detail, let $\mathbf{V}^t \in \mathbb{N}^{(M+N)\times L}$ represent the distribution of unoccupied e-taxis at the beginning of slot t, where $V^t_{i,l}$ ($V^t_{j+M,l}$) is the number of unoccupied e-taxis at region i (in charging station j) with remaining energy l at the beginning of slot t. Let $\mathbf{O}^t \in \mathbb{N}^{M\times L}$ represent the distribution of occupied e-taxis at the beginning of slot t, where $O^t_{i,l}$ is the number of occupied e-taxis at region i with remaining energy l at the beginning of slot t. It is clear that given the system state defined for every e-taxi, i.e., $\mathbf{s}^t$, by aggregating the e-taxis with the same occupancy status, remaining energy, and location, a corresponding system state $\tilde{\mathbf{s}}^t$ is well defined.

Action: The actions of m e-taxis during slot t is denoted as $\mathbf{X}^t = \{X^t_{i,i',l}, X^t_{i,j,l}, X^t_{j,i',l}, X^t_{j,j',l}\}_{i,i'\in\mathcal{M},j,j'\in\mathcal{N},1\le l\le L} \in \mathcal{X}$. In detail, $X^t_{i,i',l}, X^t_{i,j,l}, X^t_{j,i',l}, X^t_{j,j',l} \in \mathbb{N}$ describe the number of e-taxis with remaining energy l that go to region i' or charging station j from region i, or go to region i or charging station j' from charging station j during time slot t.

As discussed previously, given the system state $\tilde{\mathbf{s}}^t$, some actions cannot be taken since they require e-taxis to drive to a region or charging station that cannot be reached within a time slot, or may make low-energy e-taxis run out of energy. We use $\tilde{w}(\tilde{\mathbf{s}}^t, \mathbf{X}^t) \in \{0,1\}$ to describe whether the action $\mathbf{X}^t$ can be taken for the system state $\tilde{\mathbf{s}}^t$, where $\tilde{w}(\tilde{\mathbf{s}}^t, \mathbf{X}^t) = 1$ if the action can be taken, 0 otherwise.

State Transition: The state transition function of the aggregated system can be computed as:
$\tilde{\phi}(\tilde{\mathbf{s}}^{t+1}|\tilde{\mathbf{s}}^t, \mathbf{X}^t) = \sum_{\mathbf{s}^{t+1}\in\mathcal{S}_{\tilde{\mathbf{s}}^{t+1}}, \mathbf{s}^t\in\mathcal{S}_{\tilde{\mathbf{s}}^t}, \mathbf{a}^t\in\mathcal{A}^t_{\mathbf{X}^t}} \phi_k(\mathbf{s}^{t+1}|\mathbf{s}^t, \mathbf{a}^t)$, where $\mathbf{s}^t \in \mathcal{S}_{\tilde{\mathbf{s}}^t}$ represents the city states that are equal to $\tilde{\mathbf{s}}^t$ after aggregation, and $\mathbf{a}^t \in \mathcal{A}^t_{\mathbf{X}^t}$ are the actions that are equal to $\mathbf{X}^t$ after aggregation.

Strategy: The non-stationary and pure strategy is denoted by $\tilde{\pi}^t(\tilde{\mathbf{s}}^t, \mathbf{X}^t) \in \{0,1\}$, dictating whether the action $\mathbf{X}^t$ is taken for given the system state $\tilde{\mathbf{s}}^t$. If the action is taken, $\tilde{\pi}^t(\tilde{\mathbf{s}}^t, \mathbf{X}^t) = 1$; otherwise, it is 0.

Given this new MDP, we formulate another optimization problem similar to (7):

$$\begin{aligned} \max_{\tilde{\pi}} \quad & \textstyle\sum_{t=1}^{H} \mathbb{E}\big[\beta^t \tilde{F}(\tilde{\mathbf{s}}^t, \mathbf{X}^t)|\tilde{\pi}, \tilde{\phi}\big] \\ \textbf{s.t.} \quad & \tilde{\pi}(\tilde{\mathbf{s}}^t, \mathbf{X}^t) \le \tilde{w}(\tilde{\mathbf{s}}^t, \mathbf{X}^t),\ \textstyle\sum_{\mathbf{X}^t} \tilde{\pi}(\tilde{\mathbf{s}}^t, \mathbf{X}^t) = 1 \end{aligned} \tag{8}$$

where $\tilde{F}(\tilde{\mathbf{s}}^t, \mathbf{X}^t)$ is defined as

$$\begin{aligned} \tilde{F}(\tilde{\mathbf{s}}^t, \mathbf{X}^t) = \tfrac{1}{2}\Big(& \textstyle\sum_{i=1}^{M} X^t_i\big(f_r(d^t_i, X^t_i) + f_r(d^t_i, 1)\big) \\ &+ \textstyle\sum_{j=1}^{N} X^t_j\big(f_c(p_j, X^t_j) + f_c(p_j, 1)\big)\Big) \end{aligned} \tag{9}$$

$$X^t_i = \textstyle\sum_{i',j,l} X^t_{i',i,l} + X^t_{j,i,l}, \quad X^t_j = \sum_{i,j',l} X^t_{i,j,l} + X^t_{j',j,l}$$

First, we observe that the optimal strategy of Problem (7) can be obtained from the optimal strategy of Problem (8). Given the optimal strategy of Problem (8), i.e., $\tilde{\pi}^t(\tilde{\mathbf{s}}^t, \mathbf{X}^t)$, the optimal strategy of any e-taxi k during slot t with the system state $\mathbf{s}^t$, i.e., $\pi^t_k(\mathbf{s}^t, a^t_k)$ ($\forall a^t_k \in \mathcal{A}_k$) can obtained as follows. First, we find the corresponding aggregated system state of $\mathbf{s}^t$, denoted as $\hat{\mathbf{s}}^t \in \tilde{\mathcal{S}}$. Secondly, we obtain the action of m e-taxis during slot t based on $\tilde{\pi}^t(\hat{\mathbf{s}}^t, \mathbf{X}^t)$, which is denoted as $\hat{\mathbf{X}}^t$, and $\tilde{\pi}^t(\hat{\mathbf{s}}^t, \hat{\mathbf{X}}^t) = 1$. Finally, for the e-taxis with the same occupancy status (e.g., unoccupied), remaining energy (e.g., l), and location (e.g., region i), their deterministic actions are determined based on $\{\hat{X}^t_{i,i',l}, \hat{X}^t_{i,j,l}\}_{i'\in\mathcal{M},j\in\mathcal{N}}$. Since e-taxis are identical, these e-taxis with the same local state can arbitrarily select an action as long as their actions satisfy the optimal aggregated action $\{\hat{X}^t_{i,i',l}, \hat{X}^t_{i,j,l}\}_{i'\in\mathcal{M},j\in\mathcal{N}}$. It is straightforward to show that the so obtained strategies of all the e-taxis constitute an optimal solution of Problem (7).

B. Approximation of Markov Decision Process by Sampling

Although the aggregation reduces the action and the state space, the resulting MDP remains computationally intractable due to the still large number of actions and states. We now propose an efficient algorithm to approximate the optimal strategy of Problem (8). Similar to [23], the high-level idea is to sample a part of the MDP to construct a smaller MDP.

Given the initial state $\tilde{\mathbf{s}}^t$ of the e-taxi system at the beginning of slot t, we can first obtain the optimal action $\mathbf{X}^t \in \mathcal{X}$ that can maximize $\tilde{F}(\tilde{\mathbf{s}}^t, \mathbf{X}^t)$ for slot t (as opposed to the future H time slots) by solving the following optimization problem:

$$\max_{\mathbf{X}^t} \quad \tilde{F}(\tilde{\mathbf{s}}^t, \mathbf{X}^t), \quad \textbf{s.t.} \quad \tilde{w}^t(\tilde{\mathbf{s}}^t, \mathbf{X}^t) = 1 \tag{10}$$

Given the optimal value of the above optimization problem, denoted as $\tilde{F}_{max}(\tilde{\mathbf{s}}^t)$, we sample a set of actions, denoted as $\mathcal{X}(\tilde{\mathbf{s}}^t)$, where $\forall\ \mathbf{X}^t \in \mathcal{X}(\tilde{\mathbf{s}}^t)$, $\tilde{F}(\tilde{\mathbf{s}}^t, \mathbf{X}^t) \ge \tilde{F}_{max}(\tilde{\mathbf{s}}^t) - \epsilon(\tilde{\mathbf{s}}^t)$. The intuition is that, while taking the optimal action of the current slot may not maximize the objective function in the long-term, taking an action that significantly decreases the current utility (i.e., more than $\epsilon(\tilde{\mathbf{s}}^t)$) is unlikely to be optimal.

Given a system state $\tilde{\mathbf{s}}^t$ and an action $\mathbf{X}^t \in \mathcal{X}(\tilde{\mathbf{s}}^t)$, there is a distribution of the system state of the next slot, i.e., $\tilde{\phi}(\tilde{\mathbf{s}}^{t+1}|\tilde{\mathbf{s}}^t, \mathbf{X}^t)$. We sample the K most probable next-states, denoted as $\mathcal{S}(\tilde{\mathbf{s}}^t, \mathbf{X}^t)$, the probability of transiting to each next-state $\tilde{\mathbf{s}}^{t+1} \in \mathcal{S}(\tilde{\mathbf{s}}^t, \mathbf{X}^t)$ being $\tilde{\phi}_s(\tilde{\mathbf{s}}^{t+1}|\tilde{\mathbf{s}}^t, \mathbf{X}^t) = \frac{\tilde{\phi}(\tilde{\mathbf{s}}^{t+1}|\tilde{\mathbf{s}}^t, \mathbf{X}^t)}{\sum_{\tilde{\mathbf{s}}\in\mathcal{S}(\tilde{\mathbf{s}}^t, \mathbf{X}^t)} \tilde{\phi}(\tilde{\mathbf{s}}^{t+1}|\tilde{\mathbf{s}}^t, \mathbf{X}^t)}$. In summary, given the initial system state at the beginning of slot t, we sample a subset of actions $\mathcal{X}(\tilde{\mathbf{s}}^t)$ and a subset of next-states after taking an action $\mathcal{S}(\tilde{\mathbf{s}}^t, \mathbf{X}^t)$, and re-compute system state transition probability.

We define the state-action value function as: if $t = H-1$, $Q^t(\tilde{\mathbf{s}}, \mathbf{X}^t) = \tilde{F}(\tilde{\mathbf{s}}, \mathbf{X}^t)$; otherwise, $Q^t(\tilde{\mathbf{s}}, \mathbf{X}^t) = \tilde{F}(\tilde{\mathbf{s}}, \mathbf{X}^t) + \sum\limits_{\tilde{\mathbf{s}}\in\mathcal{S}(\tilde{\mathbf{s}}, \mathbf{X}^t)} \beta\tilde{\phi}_s(\tilde{\mathbf{s}}'|\tilde{\mathbf{s}}, \mathbf{X}^t)V^{t+1}(\tilde{\mathbf{s}})$, where $t < H-1$.

The state value function $V^{t+1}(\tilde{\mathbf{s}})$ is defined as:

$$V^{t+1}(\tilde{\mathbf{s}}) = \max_{\mathbf{X}\in\mathcal{X}(\tilde{\mathbf{s}})} Q^{t+1}(\tilde{\mathbf{s}}, \mathbf{X}). \quad (11)$$

After sampling, we obtain a much smaller MDP, and can then solve it by value iteration. The pseudo-code of the detailed MDP approximation algorithm is shown in Algorithm 1. The output of this algorithm is an approximately optimal strategy of Problem (8), which is also an approximate NE of the stochastic potential game of e-taxis (cf. Theorem 2). The computation complexity of this algorithm is $O(HA(AK)^{2H})$, where A is the maximum number of elements in $\mathcal{X}(\tilde{\mathbf{s}}^t)$ for all possible $\tilde{\mathbf{s}}^t$.

Algorithm 1: MDP approximation by sampling

Input: Initial state $\tilde{\mathbf{s}}^t$ and the slot t

Output: $\forall \mathbf{X}^t \in \mathcal{X},\ \tilde{\pi}^t(\tilde{\mathbf{s}}^t, \mathbf{X}^t)$

Solve Problem (10) to obtain the optimal value $\tilde{F}_{max}(\tilde{\mathbf{s}}^t)$.
Sample the set of actions $\mathcal{X}(\tilde{\mathbf{s}}^t)$.
for each $\mathbf{X}^t \in \mathcal{X}(\tilde{\mathbf{s}}^t)$ **do**
  Sample the K most probable next-states $\mathcal{S}(\tilde{\mathbf{s}}^t, \mathbf{X}^t)$.
  Update the state transition function $\tilde{\phi}_s(\tilde{\mathbf{s}}^{t+1}|\tilde{\mathbf{s}}^t, \mathbf{X}^t)$.
  Compute $Q^t(\tilde{\mathbf{s}}^t, \mathbf{X}^t)$.
end for
The action should be taken under $\tilde{\mathbf{s}}^t$ is $\tilde{\mathbf{X}} = \arg\max_{\mathbf{X}^t\in\mathcal{X}(\tilde{\mathbf{s}}^t)} Q^t(\tilde{\mathbf{s}}^t, \mathbf{X}^t)$
for $\mathbf{X}^t \in \mathcal{X}$ **do**
  if $\mathbf{X}^t = \tilde{\mathbf{X}}$ **then**
    $\tilde{\pi}^t(\tilde{\mathbf{s}}^t, \mathbf{X}^t) = 1$
  else
    $\tilde{\pi}^t(\tilde{\mathbf{s}}^t, \mathbf{X}^t) = 0$
  end if
end for
return $\forall \mathbf{X}^t \in \mathcal{X},\ \tilde{\pi}^t(\tilde{\mathbf{s}}^t, \mathbf{X}^t)$

VI. System Efficiency

A. System Efficiency Analysis

For a non-cooperative mobility-on-demand system consisting of e-taxis, an important question to the online platform is how close the joint actions of the drivers are to maximize the income of the online platform. The platform income is formulated as: $\sum_{k=1}^{m} \alpha \hat{J}_k^{\pi}(\mathbf{s}^1)$, where π is the concatenation of m e-taxis' strategies and $\hat{J}_k^{\pi}(\mathbf{s}^1)$ is the total fare of rides that the e-taxi k serves. It is assumed that the platform takes a cut from the fare cost of a ride and the ratio is α. Note that the system efficiency does not consider the e-taxis' payment for charging, which is different from [24] that uses the sum of all players' utility to measure the system efficiency.

Given the definition of system efficiency, we observe that the NE of the stochastic e-taxi game does not always maximize the system efficiency. To demonstrate this, let us consider an example with two regions and two e-taxis. In this example, we only consider the future one time slot, i.e., $H = 1$. The parameters of each region are set as $r_1 = 22, r_2 = 16, \tau_1 = \tau_2 = 4, d_1 = d_2 = 1$. For each e-taxi, there are two actions.

TABLE I

Table of each e-taxi's utility and value of potential function under different actions

		Action of e-taxi #2	
		Region 1	Region 2
Action of e-taxi #1	Region 1	(14,14), 32	(18,12), 30
	Region 2	(12,18), 30	(8,8), 20

The utility of each e-taxi and the value of the potential function under the different actions of two e-taxis are shown in Table I. The first two values in the parentheses represent the utility of each e-taxi and the last value is the potential function value. It is clear that both e-taxis would go to the first region at the NE, maximizing the potential function. However, the NE actions do not achieve the system optimum as 12+18>14+14.

B. Achieving System Efficiency by Pricing Schemes

Given that drivers are self-interested, we now design a pricing scheme for the platform managers to induce drivers to reach the NE that is system optimum. In the real world ride-sharing systems, the platforms have already applied incentives to impact the actions of drivers. As the platforms are responsible for distributing revenue to drivers, this provides opportunities for the designed pricing scheme to be applied.

The high-level idea of the pricing scheme is that if a driver's actions are similar to his/her actions in a system optimal strategy compared with other drivers, this driver will get some reward; otherwise, this driver will be charged a penalty. There might be multiple strategies that maximize the system efficiency simultaneously. When applying a system-wise optimal strategy, a driver may have the minimum utility compared with other drivers. We select a strategy from these strategies, which can maximize the minimum utility of drivers and denote it as $\bar{\pi}$. This max-min operation is widely used to enhance the fairness among agents [25].

The pricing function for each e-taxi k during slot t for given the system state $\mathbf{s}^t$ is designed as

$$P(\mathbf{s}^t, \pi_k^t) = \frac{1}{2}\big(D^{\pi_k^t,\bar{\pi}_k^t}(\mathbf{s}^t) - \frac{1}{m-1}\sum_{k'\neq k} D^{\pi_{k'}^t,\bar{\pi}_{k'}^t}(\mathbf{s}^t)\big) \quad (12)$$

where $D^{\pi_k^t,\bar{\pi}_k^t}(\mathbf{s}^t)$ is the difference when taking two different actions under the two strategies, i.e.,

$$D^{\pi_k^t,\bar{\pi}_k^t}(\mathbf{s}^t) = \sum_j \pi_k^t(\mathbf{s}^t, j)\big(f_c(p_j, 1) + f_c(p_j, \sum_j \pi_k^t(\mathbf{s}^t, j))\big)$$
$$+ \sum_i \pi_k^t(\mathbf{s}^t, i) f_r(d_i^t, 1) - \Big(\sum_i \bar{\pi}_k^t(\mathbf{s}^t, i) f_r(d_i^t, 1)$$
$$+ \sum_j \bar{\pi}_k^t(\mathbf{s}^t, j)\big(f_c(p_j, 1) + f_c(p_j, \sum_j \bar{\pi}_k^t(\mathbf{s}^t, j))\big)\Big)$$

It is noted that given system state $\mathbf{s}^t$, the total payment to the e-taxi manager is $\sum_{k=1}^{m} P(\mathbf{s}^t, \pi_k^t) = 0$, i.e., the budget of the pricing scheme is balanced.

After applying the pricing scheme, the utility of an e-taxi k taking action a_k^t during slot t for system state $\mathbf{s}^t$ changes as:

$$\bar{R}_k(\mathbf{s}^t, a_k^t, \mathbf{a}_{-k}^t) = R_k(\mathbf{s}^t, a_k^t, \mathbf{a}_{-k}^t) - P(\mathbf{s}^t, \pi_k^t)$$

We use $RI_k(\mathbf{s}^t, \mathbf{a}^t)$ to denote the utility of the e-taxi k during slot t for serving passengers under the joint action $\mathbf{a}^t$. It is noted that if the e-taxi k selects to charge the battery or continues to deliver passengers, $RI_k(\mathbf{s}^t, \mathbf{a}^t)$ is equal to 0. Then we have $\sum_{k=1}^{m} RI_k(\mathbf{s}^t, \mathbf{a}^t) = \sum_{i=1}^{M} I_i(\mathbf{a}^t) f_r(d_i^t, I_i(\mathbf{a}^t))$.

Theorem 3. *The stochastic e-taxi game with pricing is a stochastic potential game with a potential function defined as*

$$\bar{\Phi}^{\pi}(\mathbf{s}^1) = \frac{1}{2}\sum_{t=1}^{H}\mathbb{E}\big[\beta^t \sum_{k=1}^{m} RI_k(\mathbf{s}^t, \mathbf{a}^t)|\pi\big] \quad (13)$$

Corollary 1. *With the pricing scheme, the NE of the e-taxi game obtained from maximizing the potential function, i.e., Eq.* (13)*, achieves system efficiency.*

After applying the pricing scheme, the utility functions of drivers change, introducing a new non-cooperative game. The drivers have different NE strategies in the new game. Theorem 3 implies that the new e-taxi game is also a potential game, and the potential function is equal to half of the total fare cost of rides that all drivers serve. From Theorem 2, the strategies of drivers that maximize the potential function (i.e., optimizing the system efficiency) is the NE. Therefore, the NE obtained by maximizing Eq. (13) achieves system efficiency. It is noted that this pricing scheme can be implemented in an online way, where the on-demand system managers need to notify drivers the pricing function ahead.

VII. Trace-based Simulations

A. Simulation Setting

We use three datasets to conduct the trace-based simulations. The first dataset includes the trajectory data of nearly 700 e-taxis. Each e-taxi has a wireless communication module and a GPS device, which are used to upload the real-time location and the occupancy status every 30 seconds. The second dataset is for charging stations. There are 37 charging stations deployed in the city, which have different number of charging points. They are built only for e-taxis to encourage the deployment of the e-taxi system. The dataset includes the GPS location and the number of charging points in each charging station. The last dataset contains the information of passenger trips, including when and where a passenger is picked up and dropped off, and the fare cost of the trip. The three datasets are collected from the same city, where the e-taxi system and the charging stations only for e-taxis are deployed.

There are five strategies compared in the evaluation. (i) Efficient pricing schemes (EPS): after applying the pricing scheme, the utility function of drivers changes, introducing the different NE of the game. We assume that the drivers follow the new NE. (ii) Nash Equilibrium (NE): it is assumed that each driver follows the Nash Equilibrium strategy without the designed pricing scheme. (iii) Reactive to passenger demand (R2D): an e-taxi first filters the regions that can be reached within a time slot, and then selects the target region from the filtered regions based on a probability distribution, which is defined based on the passenger demand in these filtered regions. (iv) Reactive to trip fare (R2F): an e-taxi selects the region that has the maximum average trip fare in the historical data. (v) Oracle [13]: it is assumed that there is a centralized controller that optimizes and controls all e-taxis' actions to maximize the system efficiency over a future time horizon, demonstrating the upper bound of system efficiency. This solution assumes that e-taxis are cooperative and approximately characterizes the system optimum, which is however not the case in the competitive mobility-on-demand systems. In R2D and R2F, the strategy for charging is the same as that in [26] by which an e-taxi only charges its battery when the remaining energy is below 15% and selects the charging station with the minimum waiting time. It is noted that R2D and R2F consider optimizing the utility of the current time slot rather than the long-term cumulative utility. We do not compare our pricing scheme with the existing ones [9], [10] because they have the different objectives, e.g., incentivizing drivers to work in hours when taxi supply is low or increasing the fare cost to maximize market revenue rather than addressing the inefficient competition among e-taxis.

We use the following metrics to compare the performance of the five strategies. (i) Daily utility per e-taxi: it is equal to the total income of an e-taxi minus the payment for charging during a day. In the evaluation, we use the time-varying electricity price from [27] as charging prices. (ii) System efficiency: it is defined as the fare cost of all rides during a day times the rate that is taken by the online platforms, i.e., 35% [28]. Since we use the data from Shenzhen, a city in China, for our evaluation, the unit of system efficiency is the Chinese Yuan (CNY). (iii) Number of served passengers: it is the number of passengers that are served by e-taxis during a day. The length of a time slot is 20 minutes. We consider the future four time slots, i.e., $H = 4$, and we set $K = 5$.

B. Results

1) Comparison of different strategies: We measure the system efficiency of the different strategies and present the results in Figure 2. There are two approximated NE strategies evaluated in this figure: the first one is the approximated NE of the stochastic game (i.e., NE), and the second one is the improved approximated NE of the stochastic e-taxi game with the proposed pricing scheme (i.e., EPS). Since oracle focuses on maximizing the fare cost of all rides by coordinating their actions, it represents the maximum system efficiency. We observe that the system efficiency of the original NE is 73.5% of that of the oracle, meaning that the NE of the stochastic e-taxi game is somewhat far from achieving the system optimum due to the non-cooperative environment. Nonetheless, the new NE with the proposed pricing scheme achieves a similar (95.5%) system efficiency compared with oracle, demonstrating that our design of the pricing scheme is effective in improving the NE of the e-taxi system. The second observation is that the number of served passengers of the original NE is 86.2% of that of the oracle, meaning the NE of the competitive game misses some passengers due to competition for passengers with high fare cost of rides. In spite of that, the new NE with the pricing scheme increases the number of served passengers, i.e., 91.4% of that of the oracle, showing that our pricing scheme indirectly enhances the service quality for passengers.

We plot the average daily utility per e-taxi (with standard deviation marked) and the distribution of e-taxis' daily utility in Figures 3 and 4. We have several observations. First, the new NE with the pricing scheme achieves a similar (96.1%)

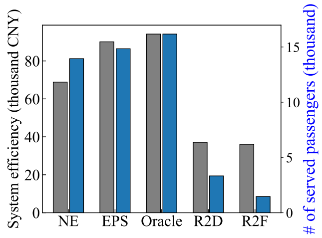

Fig. 2. System efficiency & number of served passengers

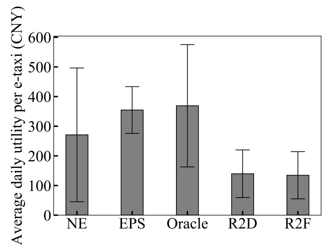

Fig. 3. Average utility of each e-taxi

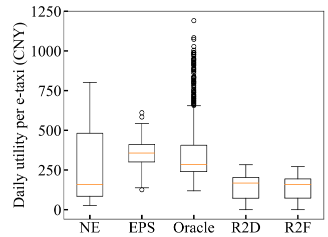

Fig. 4. E-taxi utility distribution

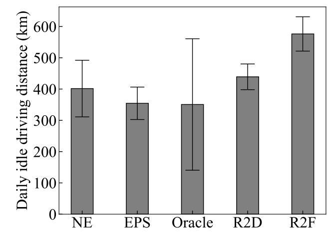

Fig. 5. Idle driving distance

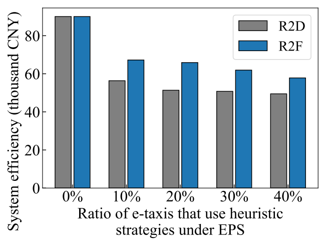

Fig. 6. Impact of e-taxis using heuristic strategies with pricing

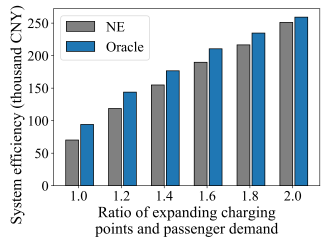

Fig. 7. Performance with expansion of charging points and passenger demand

average daily utility per e-taxi compared with oracle, showing that our pricing scheme reduces the inefficient competition for limited passengers and charging points, and enhances drivers' utility. Second, our pricing scheme introduces a lower spread of daily utility among drivers compared with oracle and the original NE, demonstrating that the pricing scheme can provide fairer utility for drivers. The reason is that compared with the original NE, the pricing scheme penalizes and rewards drivers based on the differences between their actions and the system optimal actions that optimize fairness. Whereas, oracle only focuses on optimizing the system efficiency. The last observation is that the NE with the pricing scheme (EPS) improves the average daily utility per e-taxi by 30.6%, 154.7%, and 164.2% compared with the original NE, R2D, and R2F since the two heuristic algorithms only concentrate on the regions with many passengers or with high fare cost without considering the potential competition.

The idle driving distance per e-taxi is plotted in Figure 5. It is observed that the new NE with the pricing scheme introduces a similar daily idle driving distance compared with oracle and reduces that by 11.7%, 19.3%, and 38.5% compared with the original NE, R2D, and R2F, because both EPS and oracle pick up more passengers, reducing the idle driving distance for searching passengers. Since the e-taxis always seek the regions with the maximum average trip fare by R2F, the drivers may need to drive idly for a long distance.

2) Performance of the NE: We now evaluate the impact of e-taxis using heuristic strategies (i.e., R2D and R2F) rather than the new NE when applying the pricing scheme. Figure 6 plots the system efficiency with the different ratio of e-taxis using the heuristic strategies. We observe that the system efficiency drops by 37.4% (25.2%) when 10% e-taxi drivers use R2D (R2F). This implies that a small number of drivers that deviate from the new NE with the pricing scheme (EPS) can impact the system efficiency a lot. The second observation is that when the ratio of e-taxis using heuristic strategies increases from 10% to 40%, the system efficiency still decreases, indicating the heuristic strategies without considering the competition reduce the system efficiency.

Figure 7 shows the system efficiency at the original NE and that from oracle as we expand the passenger demand and the charging points. The first observation is that the system efficiency improves both at the original NE and with oracle as the passenger demand and charging points increase. The reason is that, with such expansions, each e-taxi has more opportunities to pick up the passengers and can charge the battery with a shorter waiting time, improving the system efficiency. We also calculate the ratio between the system efficiency of the NE and that of the oracle, i.e., the (approximate) price of anarchy. We observe that the system efficiency of the NE gets closer to the optimal system efficiency as the passenger demand and charge points increase. The intuition is that, with such expansions, the e-taxis have more flexibility and the competition among them becomes less opportunistic.

VIII. Related Work

Vehicle competition: In the field of transportation, several works have studied the competition among vehicles for passengers [4]–[7] or road network resources [8]. The non-cooperative game among taxi drivers for passengers is usually considered as a stochastic game. [4] studies how to compute the Wardrop Equilibrium distribution of taxis with the complete information. [6] designs a fictitious play algorithm for each taxi driver to converge to the NE with the assumption that each driver has the same non-stationary mixed strategy. [5] and [7] focus on solving the decision problem for a taxi driver to maximize the expected long-term cumulative revenue for the individual taxi driver in a non-cooperative setting. [8] introduces a reinforcement learning-based scheme to obtain the optimal route choice strategy of vehicles, which is Nash Equilibrium. In summary, this work differs from these related works in two aspects. (i) This work studies a new setting, i.e., competition among e-taxi drivers with mobility-on-demand platforms. (ii) This work addresses the platforms' inefficiency due to drivers' self-interested actions by penalizing or rewarding drivers, whereas, the related works focus on improving drivers' utility [4]–[8].

Dynamic fare cost of rides: Some works explore the design of dynamic fare cost of rides to impact the distribution of taxi supply in spatial-temporal dimensions, which further maximizes the market revenue [9], [10]. [9] models drivers' strategy making process as a game and proposes a dynamic time-dependent fare structure that incentivizes taxi drivers to work during the peak time by increasing fare price. [10] designs the time-of-day pricing framework for taxi systems aiming to maximize total market revenue by utilizing the temporal non-stationary nature of the taxi market. However, these works design dynamic trip fare costs to incentivize drivers to work in hours when taxi supply is low that do not resolve the system inefficiency issue due to drivers' competition.

Vehicle coordination: A handful of works optimize the efficiency of urban taxi systems by coordinating or guiding taxis' behaviors [13], [29]–[36]. Some works [13], [30], [33], [35] assume that a centralized agent controls and coordinates taxis' activities to improve service quality. [31] and [34] use the multi-agent reinforcement learning to recommend the charging station or the service request for the individual taxi to enable coordination and cooperation among taxis. However, we consider non-cooperative setting that each taxi driver is self-interested to maximize its own utility instead of optimizing the performance of the entire taxi system cooperatively.

IX. Conclusion

We first investigate the behaviors of electric taxis for serving passengers and charging batteries in a non-cooperative environment, and propose an efficient algorithm to find the Nash Equilibrium (NE) of the e-taxi system. Next, from the perspective of the service platform, to induce e-taxi drivers to optimize the system efficiency at equilibria, a pricing scheme is designed that provably induces the NE to achieve the system optimum. Finally, trace-driven simulations show that, compared with the state-of-the-art which optimizes the system efficiency by coordinating e-taxis but is not an equilibrium, the NE achieves a system efficiency of merely 73.5% of that of the cooperative state-of-the-art, and the proposed pricing scheme improves the price of anarchy to 95.5%.

References

[1] K. Lin, R. Zhao, Z. Xu, and J. Zhou, "Efficient large-scale fleet management via multi-agent deep reinforcement learning," in *ACM KDD*, 2018.

[2] How long does it take to charge an electric car? [Online]. Available: https://pod-point.com/guides/driver/how-long-to-charge-an-electric-car

[3] Compare electric cars: Ev range, specs, pricing more. [Online]. Available: https://insideevs.com/reviews/344001/compare-evs/

[4] D. Calderone and S. Sastry, "Markov decision process routing games," in *ACM/IEEE ICCPS*, 2017.

[5] E. Wang, R. Ding, Z. Yang, H. Jin, C. Miao, L. Su, F. Zhang, C. Qiao, and X. Wang, "Joint charging and relocation recommendation for e-taxi drivers via multi-agent mean field hierarchical reinforcement learning," *IEEE Transactions on Mobile Computing*, 2020.

[6] P. Varakantham, S. Cheng, G. Gordon, and A. Ahmed, "Decision support for agent populations in uncertain and congested environments," in *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.

[7] A. Ahmed, P. Varakantham, and S. Cheng, "Uncertain congestion games with assorted human agent populations," in *UAI*, 2012.

[8] B. Zhou, Q. Song, Z. Zhao, and T. Liu, "A reinforcement learning scheme for the equilibrium of the in-vehicle route choice problem based on congestion game," *Applied Mathematics and Computation*, 2020.

[9] J. Gan and B. An, "Game-theoretic considerations for optimizing taxi system efficiency," *IEEE Intelligent Systems*, 2017.

[10] X. Qian and S. V. Ukkusuri, "Time-of-day pricing in taxi markets," *IEEE Transactions on Intelligent Transportation Systems*, 2017.

[11] Z. Xu, Z. Li, Q. Guan, D. Zhang, Q. Li, J. Nan, C. Liu, W. Bian, and J. Ye, "Large-scale order dispatch in on-demand ride-hailing platforms: A learning and planning approach," in *ACM KDD*, 2018.

[12] G. Wang, X. Chen, F. Zhang, Y. Wang, and D. Zhang, "Experience: Understanding long-term evolving patterns of shared electric vehicle networks," in *ACM Mobicom*, 2019.

[13] Y. Yuan, D. Zhang, F. Miao, J. Chen, T. He, and S. Lin, "p^ 2charging: Proactive partial charging for electric taxi systems," in *ICDCS*, 2019.

[14] D. Zhang, J. Huang, Y. Li, F. Zhang, C. Xu, and T. He, "Exploring human mobility with multi-source data at extremely large metropolitan scales," in *ACM Mobicom*, 2014.

[15] J. Nievergelt, H. Hinterberger, and K. Sevcik, "The grid file: An adaptable, symmetric multikey file structure," *ACM TODS*, 1984.

[16] S. Li, Y. Yu, D. Calderone, L. Ratliff, and B. Açrkmeşe, "Tolling for constraint satisfaction in markov decision process congestion games," in *2019 American Control Conference (ACC)*, 2019.

[17] F. Wu, N. Singh, N. Vaidya, and G. Chen, "On adaptive-width channel allocation in non-cooperative, multi-radio wireless networks," in *2011 Proceedings IEEE INFOCOM*.

[18] J. Zheng, Y. Cai, N. Lu, Y. Xu, and X. Shen, "Stochastic game-theoretic spectrum access in distributed and dynamic environment," *IEEE transactions on vehicular technology*, vol. 64, no. 10, 2014.

[19] S. Valcarcel Macua, J. Zazo, and S. Zazo, "Learning parametric closed-loop policies for markov potential games," in *ICLR*, 2018.

[20] J. Holler, "Learning dynamics and reinforcement in stochastic games," Ph.D. dissertation, 2020.

[21] Y. Xiong, J. Gan, B. An, C. Miao, and Y. Soh, "Optimal pricing for efficient electric vehicle charging station management," in *AAMAS*, 2016.

[22] Technical report. [Online]. Available: https://drive.google.com/file/d/1VcGXTLuTJHEq6zbedOccuvq3sqp1aQ49/view?usp=sharing

[23] A. Sidford, M. Wang, X. Wu, L. Yang, and Y. Ye, "Near-optimal time and sample complexities for solving markov decision processes with a generative model," *Advances in Neural Information Processing Systems*, 2018.

[24] F. Wu, S. Zhong, and C. Qiao, "Globally optimal channel assignment for non-cooperative wireless networks," in *IEEE INFOCOM*, 2008.

[25] D. Nace and M. Pioro, "Max-min fairness and its applications to routing and load-balancing in communication networks: a tutorial," *IEEE Communications Surveys Tutorials*, vol. 10, no. 4, pp. 5–17, 2008.

[26] Z. Dong, C. Liu, Y. Li, J. Bao, Y. Gu, and T. He, "Rec: Predictable charging scheduling for electric taxi fleets," in *IEEE RTSS*, 2017.

[27] G. Wang, X. Xie, F. Zhang, Y. Liu, and D. Zhang, "bcharge: Data-driven real-time charging scheduling for large-scale electric bus fleets," in *2018 IEEE Real-Time Systems Symposium (RTSS)*, 2018.

[28] As rideshare prices skyrocket, uber and lyft take a bigger piece of riders' payments. [Online]. Available: https://missionlocal.org/2021/07/as-rideshare-prices-skyrocket-uber-and-lyft-take-a-bigger-bite-of-the-pie/

[29] Y. Yuan, Y. Zhao, and S. Lin, "Sac: Solar-aware e-taxi fleet charging coordination under dynamic passenger mobility," in *IEEE CDC*, 2021.

[30] F. Miao, S. Han, S. Lin, J. Stankovic, D. Zhang, S. Munir, H. Huang, T. He, and G. Pappas, "Taxi dispatch with real-time sensing data in metropolitan areas: A receding horizon control approach," *IEEE Transactions on Automation Science and Engineering*, 2016.

[31] W. Zhang, H. Liu, F. Wang, T. Xu, H. Xin, D. Dou, and H. Xiong, "Intelligent electric vehicle charging recommendation based on multi-agent reinforcement learning," in *ACM WWW*, 2021.

[32] Y. Yuan, Y. Zhao, and S. Lin, "Poet: Towards power-system-aware e-taxi coordination under dynamic passenger mobility," in *ACM e-Energy*, 2022.

[33] S. He, L. Pepin, G. Wang, D. Zhang, and F. Miao, "Data-driven distributionally robust electric vehicle balancing for mobility-on-demand systems under demand and supply uncertainties," in *IROS*, 2020.

[34] M. Li, Z. Qin, Y. Jiao, Y. Yang, J. Wang, C. Wang, G. Wu, and J. Ye, "Efficient ridesharing order dispatching with mean field multi-agent reinforcement learning," in *The World Wide Web Conference*, 2019.

[35] S. Jha, S. Cheng, M. Lowalekar, N. Wong, R. Rajendram, T. Tran, P. Varakantham, N. Trong, and F. Abd Rahman, "Upping the game of taxi driving in the age of uber," in *AAAI*, 2018.

[36] Y. Yuan, Y. Zhao, and S. Lin, "Source: Towards solar-uncertainty-aware e-taxi coordination under dynamic passenger mobility," in *American Control Conference (ACC)*, 2022.