Transportation Research Part B 121 (2019) 114–134

Contents lists available at ScienceDirect

Transportation Research Part B

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

A Markov decision process approach to vacant taxi routing with e-hailing

Xinlian Yu[a,1,*], Song Gao[b,2], Xianbiao Hu[c,3], Hyoshin Park[d,3]

[a] University of Massachusetts, Amherst, United States
[b] University of Massachusetts, Amherst, United States
[c] Missouri University of Science and Technology, United States
[d] North Carolina Agricultural and Technical State University, United States

ARTICLE INFO

Article history:
Received 27 April 2018
Revised 16 November 2018
Accepted 28 December 2018
Available online 15 January 2019

ABSTRACT

The optimal routing of a vacant taxi is formulated as a Markov Decision Process (MDP) problem to account for long-term profit over the full working period. The state is defined by the node at which a vacant taxi is located, and action is the link to take out of the node. State transition probabilities depend on passenger matching probabilities and passenger destination probabilities. The probability that a vacant taxi is matched with a passenger during the traversal of a link is calculated based on temporal Poisson arrivals of passengers and spatial Poisson distributions of competing vacant taxis. Passenger destination probabilities are calculated directly using observed fractions of passengers going to destinations from a given origin. The MDP problem is solved by value iteration resulting in an optimal routing policy, and the computational efficiency is improved by utilizing parallelized matrix operations.

The proposed model and an efficient implementation of the value iteration algorithm are tested in a case study with parameters derived from GPS trajectories of over 12,000 taxis in Shanghai, China for a study period of 5:30 - 11:30 am on a typical weekday. The optimal routing policy is compared with three heuristics based on simulated trajectories. Results show that the optimal routing policy improves average unit profit by 23.0% and 8.4% over the random walk and local hotspot heuristic respectively; and improves occupancy rate by 23.8% and 8.3% respectively. The improvement is larger during higher demand periods.

1. Introduction

1.1. Background and motivation

Taxis play an important role in providing on-demand mobility. Compared to other forms of public transportation, the advantages of taxis include speediness, privacy, comfort, door-to-door service and longer operation hours with wide spatial

* Corresponding author.
E-mail addresses: xinlianyu@umass.edu (X. Yu), sgao@umass.edu (S. Gao), xbhu@mst.edu (X. Hu), hpark1@ncat.edu (H. Park).

[1] Research Assistant
[2] Associate Professor
[3] Assistant Professor

https://doi.org/10.1016/j.trb.2018.12.013
0191-2615/

coverage. Traditionally, vacant taxis cruise on roads searching for customers. In recent years, thanks to the proliferation of GPS-enabled smartphones, e-hailing applications (e.g., Uber, Lyft, and Didi Chuxing) are widely adopted by ride-sourcing drivers and in some cases, traditional taxi drivers (Didi Chuxing in China) to receive requests from nearby customers. The driver of a hired taxi usually aligns his/her routing objective with the passenger's, given the paramount importance of customer service (taking a detour to get higher fare is unprofessional and rare), and such a problem has been well studied. The interesting question is how to route vacant taxis. Taxis cruising on roads not only result in wasted gas and time for taxi drivers but also generate additional traffic in a city. Therefore, how to improve the utilization of taxis is of importance to both taxi drivers and the society.

In an earlier study by the co-authors (Hu et al., 2012), a dynamic programming model of routing vacant taxis was proposed to depict the decisions at intersections according to the passenger arrival rate. However, the expected search time is only minimized for the next customer, which might be inefficient in the long run. For example, driving to the airport might not minimize the search time for the next customer, but it brings in a higher chance of a long trip for the next customer and thus the profit might be higher overall. For this reason, experienced taxi drivers would not simply make their customer-search decisions depending on the current searching time/profit, but would also consider the subsequent possible states that could be encountered. The majority of vacant taxi routing studies in the literature, in fact, only considers an optimization problem until meeting the next customer (e.g., Zhang et al., 2015; Qu et al., 2014; Dong et al., 2014; Hwang et al., 2015; Huang et al., 2015), and in some cases, the revenue from the next customer (e.g., Yuan et al., 2013).

This study formulates the vacant taxi routing problem as a Markov Decision Process (MDP) so that long-term objectives can be taken into account instead of the immediate one of meeting the next customer. Some studies apply reinforcement learning (RL) (Sutton and Barto, 1998) and adopt MDP formulations (e.g., Han et al., 2016; Verma et al., 2017), yet their treatments are usually not fully developed, in that important modeling issues such as competition from other taxis and e-hailing are ignored and space is highly aggregated. A typical RL algorithm also is purely data driven without taking advantage of the understanding of the underlying physical process, that is, no state transition probabilities are derived.

1.2. Objective and contributions

The objective of the study is to develop a methodology for the vacant taxi routing optimization problem to achieve better optimality, practicality and computationally efficiency. Towards this end, contributions in modeling, problem formulation and solution algorithm design are made, detailed as follows.

Modeling A queueing theory-based model for matching taxis and passengers is proposed to account for competition from other taxis and use of e-hailing apps. The routing decisions are based on the physical road network in contrast to cell/zone based, which enables more practical implementations including the generation of turn-by-turn guidance.

Problem formulation The MDP formulation optimizes long-term expected profit over the complete working period, accounting fully for the impact of current decisions on future return over multiple pickups and drop-offs, and thus is able to integrate the array of factors (e.g., searching distance, searching time, pick-up probability, competition from other taxis, revenue from the next passenger) considered by other studies (e.g., Yuan et al., 2013; Hwang et al., 2015) in a single, theoretically appealing formulation. In the Shanghai case study, the MDP formulation improves unit profit up to 27% and 8% over the random walk and local hotspot heuristic respectively; and improve occupancy rate up to 27% and 15% respectively.

Efficient implementation of the solution algorithm An enhanced value iteration for solving the MDP problem is proposed making use of efficient matrix operations, and its efficiency is tested in a mega city transportation network with reasonable running time. Existing studies handle computational efficiency by either adopting a cell/zone-based approach, or limiting the search to a local range if the physical road network is used. This study achieves computational efficiency for the complete network of a mega city at the highest level of spatial resolution through a recursive problem formulation and efficient implementation of the solution algorithm.

The proposed implementation does not improve the worst-case complexity of the value iteration algorithm. In fact, there is no trivial way to improve the worst-case complexity of the value iteration algorithm as each of the action-dependent transitions have to be visited at least once. Indeed, we focus on an efficient implementation from the point of view of data parallelism, which offers the possibility to solve large MDPs efficiently for the complete network of a mega city at the highest level of spatial resolution.

1.3. Paper organization

The remainder of the paper is organized as follows. Section 2 provides a review of the literature. Section 3 formulates an MDP problem for the vacant taxi routing problem, defining states, actions and transition probabilities, followed by the presentation of an efficient solution algorithm in Section 4. Section 5 presents the computational experiments using GPS data from Shanghai, China to evaluate the merits of the proposed methodology. Finally, Section 6 concludes the study and discusses potential directions for future work.

Table 1
An overview of single vacant taxi routing studies with network optimization.

	Network Representation	Decisions	Planning Horizon	Matching Probability	Objective
Hu et al. (2012)	Physical	Routing policy	Half cycle	Queuing theory; no competition	Minimal searching time
Zhang et al. (2015)	Physical	Route	Half cycle	Empirical frequency; with competition	Shortest path with at least one expected passenger
Yuan et al. (2013)	Physical	“Parking” place and the route to it	One cycle	Empirical frequency; competition by queue length	Combinations of 1) revenue miles per unit searching time, 2) searching time, 3) pick-up probability, 4) queue length at “parking” place
Qu et al. (2014)	Physical	Route	N/A	Empirical frequency	Maximal “profit” with a fixed number of segments
Dong et al. (2014)	Physical	Route	N/A	N/A	Maximal score (related to revenue) with distance constraint
Huang et al. (2015)	Locations and their connections	A sequence of fixed number of pick-up points	Half cycle	Empirical frequency	Minimal searching distance
Verma et al. (2017)	Cell-based	Routing policy	Multi-cycle	Empirical frequency	Maximal revenue
This Study	Physical	Routing policy	Multi-cycle	Queuing theory; with competition	Maximal profit

2. Literature review

2.1. Single vacant taxi routing problem

The single vacant taxi routing problem is also known as the taxi recommender problem in the literature. The emergence of GPS tracking has facilitated the study of taxi routing problems, and massive taxi GPS datasets have attracted the attention of researchers from various fields with expected cross-fertilization of methods from transportation engineering, operations research, and computer science. The general problem statement is that given the location of a vacant taxi in an urban area, find the optimal decisions regarding its spatial movements.

Some studies recommend one or multiple pick-up locations without solving a network-based optimization problem. The attractiveness of a location with respect to the current location of the taxi is calculated using a number of factors, such as the distance between the current location and the pick-up location, expected revenue of trips from the pick-up location, waiting time for the next passenger at the pick-up location, and the probability of getting matched with a passenger on the way from the current location to the pick-up location. Some or all of the factors are manipulated to generate a single metric of attractiveness, and locations are ranked accordingly (e.g., Powell et al., 2011; Hwang et al., 2015; Zhang et al., 2016). The approach abstracts away the taxi cruising process on roads, and thus does not have a physically meaningful objective in the problem formulation, such as maximizing profit or minimizing search time, although the evaluation of the methods is usually based on such measures.

Table 1 presents a list of studies (including the current one) where some form of network-based optimization is carried out. Note that the network can be either the physical road network, cell/zone-based or a number of locations connected by abstract links (see the “Network Representation” field).

There are three broad categories of taxi routing decisions (see the “Decisions” field):

- A route is defined as a sequence of connected links (physical road segments or abstract connections between locations) without metrics attached to the end node (labeled “Route” in the table). Different from a regular commuter routing problem, a vacant taxi does not have a definitive destination, and thus constraints are added to ensure that the optimal route does not become unrealistic, such as within a certain distance or time (Yuan et al., 2013; Dong et al., 2014), with a fixed number of segments (Qu et al., 2014; Huang et al., 2015), and with at least one expected passenger (Zhang et al., 2015).
- A route with metrics attached to the end node, the so-called “parking” place where taxies queue for passengers (hotels, transportation hubs) as in Yuan et al. (2013). Metrics attached at the end node (such as revenue miles from the next passenger) are included in the optimization together with those from the route.
- A routing policy which maps any state (node) to an action (link) (Hu et al., 2012; Han et al., 2016; Verma et al., 2017). An MDP formulation generates policies and in this paper, they are called routing policies. Note that the MDP framework allows natural extension to include other important features of the transportation network that might affect the optimal routing, such as time-of-day, traffic condition, special events, by expanding the state space. All the MDP problems in the literature, as well as this study, are with infinite horizon, and thus either a terminal state is defined, such as finding a passenger (Hu et al., 2012), or a discount factor of less than 1 is applied to returns from the future so that a convergent policy can be found (Han et al., 2016).

The "Planning Horizon" field in the table summarizes how far into the future a study considers the routing problem. Define a cycle as the vacant taxi trip from a drop-off location to the pick-up location for the next passenger, succeeded by the hired trip to the passengers' destination. Some studies consider only the first half cycle ("Half cycle"), that is, until a passenger is picked up, while a few consider the revenue from the next passenger ("One cycle"). Some studies do not have a clear probabilistic description of the cruising process ("N/A"). This study adopts a multi-cycle approach to account for the long-term profit, similar to studies based on RL (Han et al., 2016; Verma et al., 2017).

The matching of a vacant taxi and a passenger (see the "Matching Probability" field) is usually set simply as the observed fraction of matched taxis over all taxis present at a link/cell/zone, with the exception of Hu et al. (2012). Clustering sometimes is carried out to resolve the issue of data sparsity. Competition from other vacant taxis is modeled in Zhang et al. (2015) by accounting for relative locations of multiple taxis on the same link in calculating the matching probability, while Yuan et al. (2013) use the queue length at the "parking" place as in indirect measure of competition, in that it is not factored in matching probability but used in optimization as an either objective or constraint. This study uses a space Poisson distribution for vacant taxis and factors it directly in calculating matching probabilities.

The objective of the optimization problems also vary in the literature. Closely related to the planning horizon, when a half-cycle is considered, the objective function is usually the search travel time or distance, while when a full cycle is considered, revenue from the next passenger can be included. Studies without a clear model of the matching process (Qu et al., 2014; Dong et al., 2014) have objectives that are not well defined, despite that sometimes the name suggests otherwise. A planning horizon of multiple cycles allows for long-term objective to be included, and resolves the issue of the inconsistency between the optimization objective and evaluation criteria. Often times, a half-cycle or one cycle is adopted so that optimization can be done quickly, yet in evaluation, taxis are simulated for multiple cycles. The MDP formulation ensures that the correct objective is optimized.

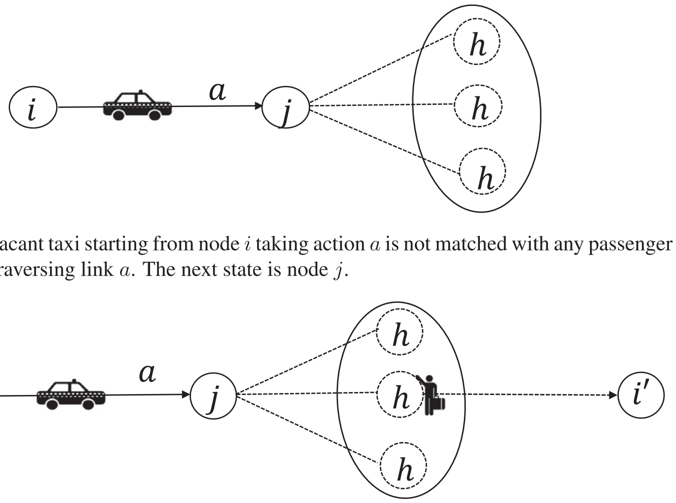
2.2. Other related work

In this section, related studies that are not directly addressing vacant taxi routing problem is reviewed. They either might provide insights into the formulation and solution of the taxi routing problem (e.g., the vehicle routing problem), or can provide support to a future extension of the current approach (taxi demand and destination prediction).

Vehicle Routing Problems Models and algorithms developed for non-myopic vehicle routing problem (VRP) under uncertainty with look-ahead policies and rolling horizons (e.g., Mitrović-Minić et al., 2004; Thomas and White, 2004; Ferrucci et al., 2013) might provide insights for taxi routing problems in terms of accounting for future unknown demand and efficient solution algorithms. Thomas and White (2004) formulated a Markov Decision Process (MDP) in which known customers may ask for service with a known probability. Mitrović-Minić et al. (2004) included double-horizon heuristic that minimizes route distance for customers served in the near-term. Ferrucci et al. (2013) presented a tabu search approach for the delivery of newspapers and applied temporal and spatial clustering of future requests, assumed to be known as a time-space Poisson distribution, which guides vehicles into request-likely areas. It is however recognized that the taxi problem is different. In a typical VRP, the service of a customer does not bring the vehicle to another location, while a taxi does and the destination is not known until the request is taken. This significantly increases the geographic spread of taxi movements. In addition, a taxi (without carpooling service) can serve only one quest at one time and a new request does not come up until the old request is finished (unless a dispatcher is sending request during the previous ride).

Taxi Demand and Destination Prediction Accounting for future states in taxi searching behavior requires sound models of geographic and temporal distributions of taxi demand and destination prediction. Several approaches have been proposed to predict taxi demand distribution which could be combined with the optimal taxi routing model including the time-series forecasting techniques such as the time-varying Poisson model and the autoregressive integrated moving average (ARIMA) (Li et al., 2012; Moreira-Matias et al., 2013), the multi-level clustering technique where demand over neighboring cells are aggregated, and the neural network based algorithms (e.g., Ke et al., 2017; Xu et al., 2017). For instance, Zhao et al. (2016) implemented and compared three models, i.e., the Markov algorithm, Lempel-Ziv-Welch algorithm, and neural network. The results showed that neural network performed better with the lower theoretical maximum predictability while the Markov predictor had better performance with the higher theoretical maximum predictability. Some socio-demographical and built-environment variables have also been in use for predicting taxi passenger demand (Qian and Ukkusuri, 2015).

Taxi Driver Search Behavior. Another category of related work aims to model accurately the actually observed customer-search behavior of vacant taxi drivers. These studies provide comprehensive and quantitative insight into factors affecting taxi drivers' incomes and assist in developing effective optimization algorithms for taxi operations. Yang and Wong (1998) developed a model to determine the taxi movements on a given road network. Their study was further improved to capture congestion effects (Wong et al., 2001; Yang et al., 2005a), multiple user classes (Wong et al., 2008), stochastic searching processes (Wong et al., 2005; Yang et al., 2010), day-to-day learning processes (Kim et al., 2005) and search frictions between vacant taxis and taxi customers (Yang and Yang, 2011; Yang et al., 2014). Wong et al. (2014a) formulated and validated multinomial logit (MNL) models to predict vacant taxi drivers' zone choices for customer searching in both peak and off-peak hours. Szeto et al. (2013) further extended the consideration to every hour in a day. Wong et al. (2014b) formulated a cell-based network and modeled the local customer-search movement of vacant taxi drivers based on the probability of successfully meeting the next taxi customers. Wong et al. (2015b) proposed a sequential logit-based



(a) A vacant taxi starting from node i taking action a is not matched with any passenger while traversing link a. The next state is node j.

(b) A vacant taxi starting from node i taking action a is matched with a passenger at node h with a destination i' while traversing link a. The next state is node i'.

Fig. 1. Illustration of the passenger matching process on a link.

vacant taxi behavior model predicting searching paths as a sequence of choices of adjacent zones while heading to their designated zones as compared with the model of Wong et al. (2014b). Wong et al. (2015a) further combined the two proposed models to a two-stage modeling approach, cell-based model for local (within zone) search decisions and ESL for zonal decisions to predict vacant taxi movements in searching for customers. Qin et al. (2017) categorize taxi drivers into three levels according to their revenue and develop a generalized multilevel ordered logit (GMOL) model to find the significant factors that influence revenue.

Others. A large number of studies have been conducted to better understand and improve the taxi market, focusing an array of topics, such as taxi equilibrium assignment analysis (e.g., Yang et al., 2005b, 2010; Yang and Yang, 2011; Long et al., 2017), taxi fleet dispatching systems (e.g., Seow et al., 2010; Hou et al., 2013; Lowalekar et al., 2016), ride-sharing/carpooling problems (e.g., Hosni et al., 2014; Lee and Savelsbergh, 2015; Qian et al., 2017; Masoud and Jayakrishnan, 2017), pricing in taxi/ride-sourcing market (e.g., Qian and Ukkusuri, 2017; Zha et al., 2017), validating user equilibrium with taxi trajectory data (Xie et al., 2017) and route planning through taxi trajectory mining (Yang et al., 2017).

3. Formulation of the non-myopic optimal taxi routing problem

The defining characteristics and assumptions of the optimal taxi routing problem are:

- The vehicle routing problem is applicable for a single taxi;
- When a taxi is hired, the routing problem is reduced to a fastest path problem from the passenger's origin to destination and is not studied explicitly in this paper;
- Passenger arrivals and vacant taxi distribution are assumed independent;
- Passenger arrivals at different nodes are assumed independent;
- A passenger is matched with the nearest vacant taxi.

Accordingly, an MDP formulation for the taxi routing problem is presented.

3.1. States and actions

A taxi driver's routing decisions over a time horizon on a given day is modeled as an MDP. A taxi travels in a traffic network $G = (N, A)$. N is the set of nodes and A the set of links. There is at most one directional link, a, from the source node i to sink node j. $A(i)$ is the set of downstream links of i. The taxi is actively searching for, or carrying passengers during a planning horizon.

When a taxi is hired, the routing problem is reduced to a fastest path problem from the passenger's origin to destination, and is not studied explicitly in this paper. The state of a vacant taxi is defined as the current node $i \in N$. The action set for state i is the set of outgoing links $A(i)$. For a given state i and action $a \in A(i)$, two types of transition to a new state i' could happen (see Fig. 1). a) The taxi is not matched with any passenger when traversing link $a = (i, j)$, and the next state is j, the sink node of a. b) The taxi is matched with a passenger at h (not necessarily the same as j due to e-hailing) when traversing link a, and the next state is the destination node of the passenger i'. To calculate state transition probabilities, the passenger matching probability on a link (Section 3.2) and passenger destination probabilities (Section 3.3) are needed.

3.2. Passenger arrival and matching probability on a link

Passengers arrive at link a following a one-dimensional space-time Poisson process with rate λ_a per hour per mile. For modeling convenience, these are simplified as homogeneous time Poisson processes at each node with a constant average passenger demand rate, and the arrival rate at node j (per hour), $\lambda_j = \sum_{a \in B(j)} \lambda_a l_a$, where l_a is the length of link a and $B(j)$ the set of incoming links. The combined process over all nodes is also a Poisson process with arrival rate $\lambda = \sum_{j \in N} \lambda_j$. In practice, demand rate λ_j is often approximated by observed met demand rate. Statistical analysis can be carried out to build a predictive model for the demand rate as a function of built environment variables (e.g., residential density, and employment by business type such as hotel and nightclub), time of day, and weather condition (e.g., Phithakkitnukoon et al., 2010; Moreira-Matias et al., 2012).

When e-hailing is used, it is assumed that the nearest vacant taxi to a passenger gets matched to the passenger. Vacant taxis around node n at any given point of time follow a two-dimensional spatial Poisson distribution with density γ_n. For a given node n, the probability of a vacant taxi r miles away (based on right-angle travel) being the nearest vacant taxi is the probability of no vacant taxi in a square (denoted by S) rotated at 45 ° centered at node n with area equal to $2r^2$ (Larson and Odoni, 1981), namely,

$$P_n(r)\{X(S) = 0\} = \exp(-2\gamma_n r^2). \tag{1}$$

where $X(S)$ denotes the number of empty taxis contained in the square S.

The set of potential pick-up nodes is limited to those that are within a certain distance to the vacant taxi. Let $N(j)$ denotes the sets of nodes within a certain matching distance R to node j. That is, R is the farthest distance between a passenger and a taxi where a request can go through. The combined process over all nodes in $N(j)$ is also a Poisson process with arrival rate $\lambda_{N(j)} = \sum_{j \in N} \lambda_j$. The matching distance can be different for different areas, and it probably changes as a function of time as well. For example, during slow hours, drivers are willing to pick up a passenger who is far away. During busy hours, drivers are less likely to accept trips with long pick-up time. In this study, the rates are assumed to be static within a study period (say, 2 hours), and future research will address time-varying rates.

Consider a vacant taxi with e-hailing traversing link a. It gets matched with a passenger at node h when the following conditions are all satisfied.

- A passenger arrives at node h during the traversal time τ_a,
- The arrival at node h is earlier than arrivals on all other nodes in $N(j)$,
- The taxi in question is the nearest vacant taxi to node h.

The probability of having at least one arrival from any node in $N(j)$ during τ_a is $1 - \exp(-\lambda_{N(j)}\tau_a)$. The probability that an arrival from node h is earlier than all other nodes is $\frac{\lambda_h}{\lambda_{N(j)}}$ (Larson and Odoni, 1981). The product of the two probabilities is the probability that the earliest arrival during τ_a happens at node h. The matching probability, $p_{a,h}$, is the product of the probability that the earliest arrival during τ_a happens at node h and the probability that the taxi in question is the nearest vacant taxi to node h, namely,

$$p_{a,h} = \begin{cases} \dfrac{\lambda_h}{\lambda_{N(j)}}\left(1 - \exp(-\lambda_{N(j)}\tau_a)\right)\exp(-2\gamma_h \mathcal{L}^2_{a\to h}), & \text{if } h \in N(j) \\ 0, & \text{otherwise} \end{cases}, \tag{2}$$

where $\mathcal{L}_{a\to h}$ is the right-angle distance from link a to node h, which can be approximated as the distance from the middle point of link a.

For those taxis that pick up passengers along the roads without e-hailing, it usually requires that the taxi and passenger to be no more than 1 block away from each other, thus the pick-up node set without e-hailing, $N(j)$, is a subset of the pick-up nodes with e-hailing.

3.3. Passenger destination probabilities

The probability of a passenger picked up at node h having node k as the destination, $p_{h\to k}$, can be approximated by the observed fraction of passengers picked up at node h going to k. When no passenger pick-up is observed at node h, the probability is undefined. To resolve this issue, the study area is divided into zones such that any zone has strictly positive number of pick-ups. Let node h be in zone $\mathcal{H}$ and node k in zone $\mathcal{K}$. Assume each node in zone $\mathcal{K}$ has equal probability of being the destination node, and the destination probability is

$$p_{h\to k} = \begin{cases} \dfrac{p_{\mathcal{H}\to\mathcal{K}}}{m_{\mathcal{K}}}, & \forall \mathcal{H} \neq \mathcal{K} \\ \dfrac{p_{\mathcal{H}\to\mathcal{K}}}{m_{\mathcal{K}} - 1}, & \forall \mathcal{H} = \mathcal{K}, \forall h \neq k \\ 0, & \text{if } h = k \end{cases}, \tag{3}$$

where $p_{\mathcal{H}\rightarrow\mathcal{K}}$ is the probability of a passenger picked up in zone $\mathcal{H}$ having zone $\mathcal{K}$ as the destination zone, and $m_{\mathcal{K}}$ is the number of nodes in zone $\mathcal{K}$. The equal probability assumption can be easily relaxed.

The proposed modeling methodology could be applied to any study area using different sizes of zones. The sizes of the zones should be designed carefully based on the required level of modeling accuracy and the information available to the modeler. An unnecessarily large zone would mask traffic pattern differences that might be important for taxis finding customers. If the sizes were too small, the relevant data collected would be statistically unreliable, and the number of samples in each zone would be insufficient to provide representative means on the model parameters. In practice, it is advised to use the traffic analysis zones (TAZs) in a regional planning model as the basis for calculating passenger destination probabilities.

3.4. State transition probabilities

For a given state i and action $a \in A(i)$ with a sink node j, the transition probability, $p_{ii'|a}$ is defined as follows:

$$p_{ii'|a} = \begin{cases} 1 - \sum\limits_{h\in N(j)} p_{a,h} + \sum\limits_{h\in N(j)} p_{a,h}\; p_{h\rightarrow j}, & \text{if } i' = j \\ \sum\limits_{h\in N(j)} p_{a,h}\; p_{h\rightarrow i'}, & \text{if } i' \neq j \end{cases} \tag{4}$$

In the first case, the next state of the taxi is the sink node j. The probability of arriving at node j is the sum of the probability of arriving at j without getting matched and the probability of picking up a passenger from node h with destination j. In the second case, the next state is not the sink node j. In this case, a passenger from node h with destination $i'(i' \neq j)$ is matched, and the taxi arrives at node i' after picking up the passenger from node h and carrying the passenger from h to i', both following shortest paths. The probabilities are summed over all possible h. The taxi continues the routing process after dropping off the passenger.

3.5. Immediate profit

It follows that the immediate profit of going from state i to i' given action a can be written as follows:

$$g_{ii'|a} = \begin{cases} \dfrac{-\alpha\tau_a(1 - \sum\limits_{h\in N(j)} p_{a,h}) + \sum\limits_{h\in N(j)} \left[F(d_{h\rightarrow j}) - \alpha(\tau_a + \mathcal{T}_{j\rightarrow h} + \mathcal{T}_{h\rightarrow j})\right] p_{a,h} p_{h\rightarrow j}}{1 - \sum\limits_{h\in N(j)} p_{a,h} + \sum\limits_{h\in N(j)} p_{a,h}}, & \text{if } i' = j \\ \dfrac{\sum\limits_{h\in N(j)} \left[F(d_{h\rightarrow i'}) - \alpha(\tau_a + \mathcal{T}_{j\rightarrow h} + \mathcal{T}_{h\rightarrow i'})\right] p_{a,h} p_{h\rightarrow i'}}{\sum\limits_{h\in N(j)} p_{a,h}\; p_{h\rightarrow i'}}, & \text{if } i' \neq j \end{cases} \tag{5}$$

where α is the taxi operating cost per unit time, and $\mathcal{T}_{j\rightarrow h}$ ($\mathcal{T}_{h\rightarrow i'}$) is the fastest path travel time from node j to h (h to i'). $F(d_{h\rightarrow i'})$ is the taxi fare of an occupied trip from pick-up node h to destination node i', where $d_{h\rightarrow i'}$ is the occupied travel distance from node h to i'. $F(d_{h\rightarrow i'})$ can be calculated by Eq. (6), a piecewise linear fare structure used in the study area of Shanghai in this study and can be adapted to other forms that depends on distance and/or travel time.

$$F(d_{h\rightarrow i'}) = \begin{cases} f_0, & \text{if } d_{h\rightarrow i'} \leq d_0 \\ f_0 + \beta(d_{h\rightarrow i'} - d_0), & \text{if } d_0 \leq d_{h\rightarrow i'} \leq d_1 \\ f_0 + \beta(d_1 - d_0) + \gamma(d_{h\rightarrow i'} - d_0), & \text{if } d_{h\rightarrow i'} \geq d_1 \end{cases}, \tag{6}$$

Similar to the state transition equation, the expected payoff is calculated for two different cases. In the case where the next state is j, the payoff is either the negative of the operating cost of traversing link a, which is $-\alpha\tau_a$, with the probability of not matched with a passenger, $\Sigma_{h\in N(j)} p_{a,h}$, or the taxi fare of going from h to j minus the operating cost of traversing a, going from j to h and from h to j, which is $F(d_{h\rightarrow j}) - \alpha(\tau_a + \mathcal{T}_{j\rightarrow h} + \mathcal{T}_{h\rightarrow j})$, with the probability of getting matched with a passenger whose destination is j. Note that the probabilities are normalized. In the second case where the taxi is matched with a passenger whose destination is $i' \neq j$, the same calculation of fare minus operating cost is carried out, with normalized probabilities.

3.6. The Bellman equation

Let $V^*(i)$ denote the optimal expected payoff starting from state i. The taxi driver chooses the action at each state i to maximize the expected payoff that is the sum of the expected immediate payoff and the expected downstream payoff, which is the expectation of the payoff over all possible next state i'. The optimal expected payoff is obtained by solving the Bellman equation (Bellman, 1957) as follows:

$$V^*(i) = \max_{a\in A(i)} \sum_{i'\in N} \left[g_{ii'|a} + \rho V^*(i')\right] p_{ii'|a}, \forall i \in N. \tag{7}$$

where ρ is discount factor, and $0 \le \rho \le 1$. Since this is an infinite horizon problem, ρ is set to be a number slightly smaller than 1 to ensure the existence of finite optimal expected payoff.

The optimal routing policy is then written as follows:

$$\mu^*(i) = \arg\max_{a\in A(i)} \sum_{i'\in N} \left[g_{ii'|a} + \rho V^*(i')\right] p_{ii'|a}, \forall i \in N. \tag{8}$$

4. Solving the Bellman equation

The Bellman equation can be solved by value iteration (Bellman, 1957). $V^*(i)$ is termed the value function, and at each iteration, the value function at each state is updated by Eq. (7) where the value function estimates from the previous iteration are substituted into the right-hand side of the equation to obtain new estimates at the left-hand side. The time complexity is $O(|A| \cdot |N|^2)$ per iteration, with $|A|$ actions and $|N|$ states.

Vectorizing batch operations avoid expensive for-loops and significantly improves computational performance (van der Walt et al., 2011). The Bellman equation is thus re-formulated as a series of matrix operations.

The transition probability matrix, $P(|NA| \times |N|)$, and immediate payoff matrix, $G(|NA| \times |N|)$, are defined as follows:

$$P_{ia,i'} = \begin{cases} p_{ii'|a}, & \text{if } a \in A(i) \\ 0, & \text{otherwise} \end{cases}, \quad \forall i, i' \in N, a \in A \tag{9}$$

and

$$G_{ia,i'} = \begin{cases} g_{ii'|a}, & \text{if } a \in A(i) \\ 0, & \text{otherwise} \end{cases}, \quad \forall i, i' \in N, a \in A \tag{10}$$

Define an expected payoff vector $V(|N| \times 1)$ with entry V_i as the expected payoff starting from state i, $i \in N$. According to Eq. (7), define

$$Y_{ia,1} = \begin{cases} \sum_{i'} g_{ii'|a} p_{ii'|a} + \rho \sum_{i'} V(i') p_{ii'|a}, & \text{if } a \in A(i) \\ 0, & \text{otherwise} \end{cases}, \quad \forall i, i' \in N, a \in A \tag{11}$$

To re-write Eq. (11) as matrix operations, define vector $b(|N| \times 1)$ where $b_{i,1} = 1$, $\forall i \in N$. Note that $((G \circ P) \cdot b)_{ia,1} = \sum_{i'} g_{ii'|a} p_{ii'|a}$, and $(P \cdot V)_{ia,1} = \sum_{i'} V(i') p_{ii'|a}$, where $\circ$ is the element-by-element multiplication operator and $\cdot$ the matrix multiplication operator. Eq. (11) then can be written as

$$Y = (G \circ P) \cdot b + \rho P \cdot V. \tag{12}$$

The max operator in Eq. (7) needs to operate on a matrix where a row represents a state and the columns represent expected payoffs from all feasible actions from that state (all outgoing links from the node). Thus Y is reshaped into a matrix $U(|N| \times |A|)$, such that

$$U_{i,a} = Y_{ia,1} \tag{13}$$

Note that $U_{i,a} = 0$, if $a \notin A(i)$, that is, a is not an outgoing link of node i. The expected payoff of an outgoing link could be negative (taxi driver losing money), and if a max operator is directly applied to U_i, which takes the maximum over all columns, an infeasible action with an expected payoff of 0 could be chosen as the optimal. Therefore an exponential transformation is applied to $U_{i,\,a}$, if $a \in A(i)$. Define the operator $\exp(U)$ such that

$$\exp(U)_{i,a} = \begin{cases} \exp(U_{i,a}), & \text{if } a \in A(i), \\ 0, & \text{otherwise.} \end{cases} \quad \forall i \in N, a \in A, \tag{14}$$

and let $W = \exp(U)$. Eq. (7) can then be written as

$$V_i = \ln(\max_a(W_{i,a}), \forall i \in N, a \in A \tag{15}$$

Algorithm 1 shows the procedure for solving the Bellman equation via matrix operations. At each iteration, the expected payoff vector V is updated (Lines 6–9). The iteration is stopped when $||V_{prev}/V - 1||_\infty < \epsilon$ for a given threshold $\epsilon > 0$, where $||\cdot||_\infty$ is the maximum norm.

Algorithm 1: Solving Bellman's equation.

Input: transition probability matrix P, immediate payoff matrix G, all-ones vector b and convergence criteria ϵ
begin
  1. Initializing the value functions and the discount factor
  $V = 0;\ 0 \leq \rho \leq 1;$
  2.Value iteration
  do
    $V_{prev} = V;$
    $Y = (G \circ P) \cdot b + \rho P \cdot V;$
    Define matrix $U(|N| \times |A|)$, where $U_{i,a} = Y_{ia},\ i \in N, a \in A;$
    $W = \exp(U);$
    $V = \ln(\max_a W),\ \mu = \arg\max_a U;$
  while $||V_{prev}/V - 1||_\infty < \epsilon;$
  return V, μ
end

5. Computational tests

The objectives of the computational tests are three-fold:

- To demonstrate the efficiency of the solution algorithm in solving the proposed MDP problem in a real-life, large-scale network,
- To understand the differences between the proposed MDP and baseline heuristics in terms of unit profit and occupancy rate, and
- To understand the solution patterns of the proposed MDP problem, in relation to baseline heuristics.

5.1. The network, GPS data and experiment setup

The Study Area and Network. Shanghai is the most populated metropolitan area in China, with a land area of more than 6300 km^2, with a population of over 24 million. Shanghai urban area spreads broadly, including the city center landmark area (the Bund and People's square) and several other central areas. Shanghai roadway network is comprised of 13,531 nodes and 30,167 directed links, excluding connectors. The travel time on each link of the network is computed based on the speed limit by road type. Shortest paths and travel times between all nodes are pre-calculated and stored in a look-up table.

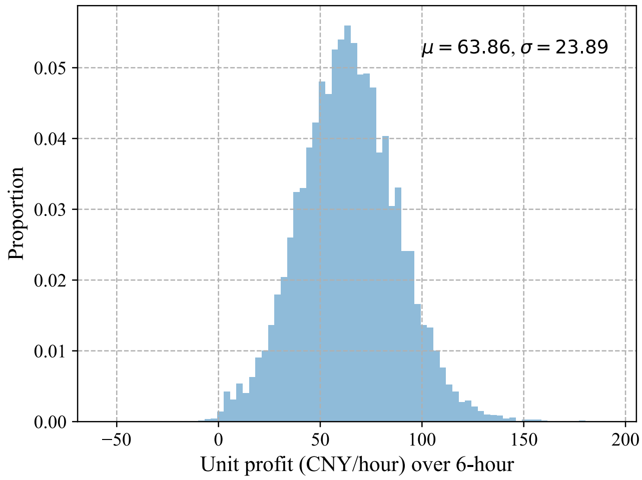
The study area is divided into 4518 zones with smaller zones in populated urban areas and larger zones in suburban areas. The zones are small enough so that it is realistic to assume that all nodes within a destination zone have an equal probability of being the destination node, an assumption required in Eq. (3). Note that the states, actions and state transitions are network-based instead of zone-based. The purpose of the zones is to ensure a large enough sample size in calculating passenger destination probabilities.

GPS Data. GPS trajectories with 10-second gaps and indicators of hired vs. vacant status are available from one of the major taxi companies (market share of approximately 25%) in Shanghai. The market share is deemed sufficiently large to deduce the movements of taxis in general and provide adequate evaluation platform for the proposed methodology.

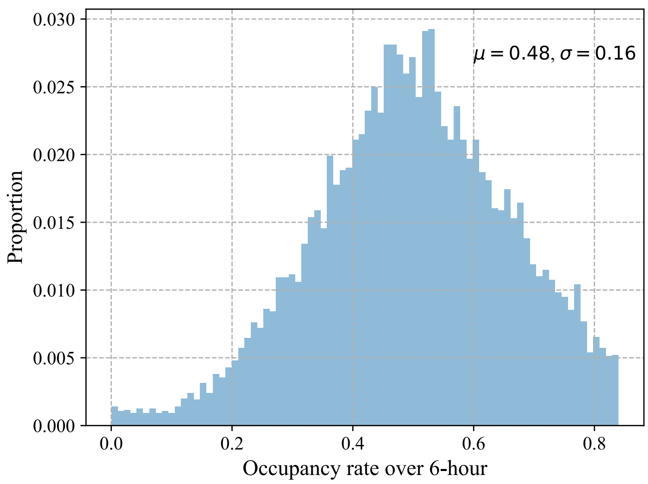
Data cleaning is carried out to remove obvious mistakes, for example, hired trips with exceptionally short travel distances over a long period, and very short travel times. The mistakes could be due to the GPS device malfunctions, poor connectivity to satellites in the urban areas surrounded by highrise buildings, or human errors by taxi drivers who operated the devices. Occupied trips are eliminated if its distance is shorter than 500 m, and/or if its travel time is shorter than 1 min.

In Shanghai, taxi drivers work for one day and rest for one day. They usually shift between 5–6 am. Average daily working time is 14.8 hours (Lv et al., 2017). An inquiry (Qin et al., 2017) found that taxi drivers in Shanghai have a flexible meal schedule with varying length during the daytime (scattering across 11:00-14:00 and 16:00-19:30), it is thus very complicated to distinguish between the status waiting for passengers and the status taking a meal. We eventually select the period 5:30-11:30 am as our study time to ensure that it's continuous and long enough for making non-myopic routing decisions. It should be noted that the proposed methodology can be directly applied to longer study periods.

After data cleaning, trajectories from 12,017 taxis in 5:30-11:30 am on a representative weekday in April, 2015 are used for the case study. No special events or holidays which may introduce great trip variability were reported during the study period.



(a) Distribution of unit profit



(b) Distribution of occupancy rate

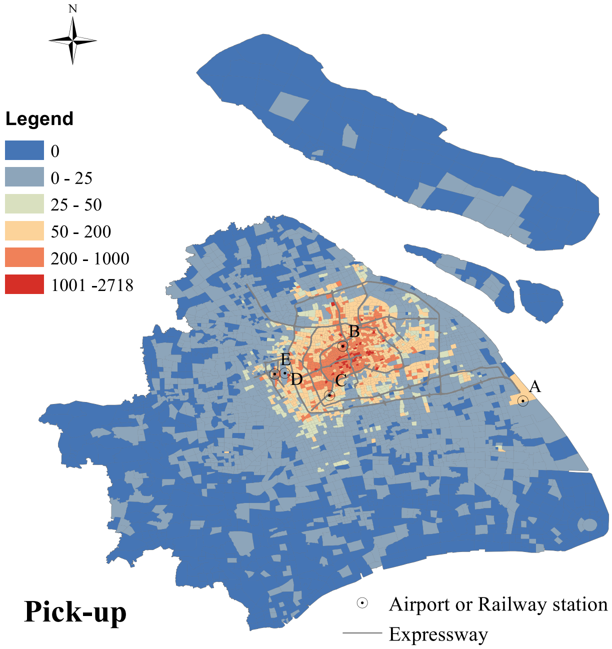
Fig. 2. Distribution of observed unit profit and occupancy rate 5:30-11:30 am.

Observed Taxi Profitability. Taxi fares are charged based on distance traveled. The parameters for calculating taxi fare $F(d_{h\to i'})$ in Eq. (6) are set as $f_0 = 14, d_0 = 3, d_1 = 15, \beta = 2.5$ and $\gamma = 3.6$.[4] The unit operating cost, α, is assumed to be 0.5 CNY/min.

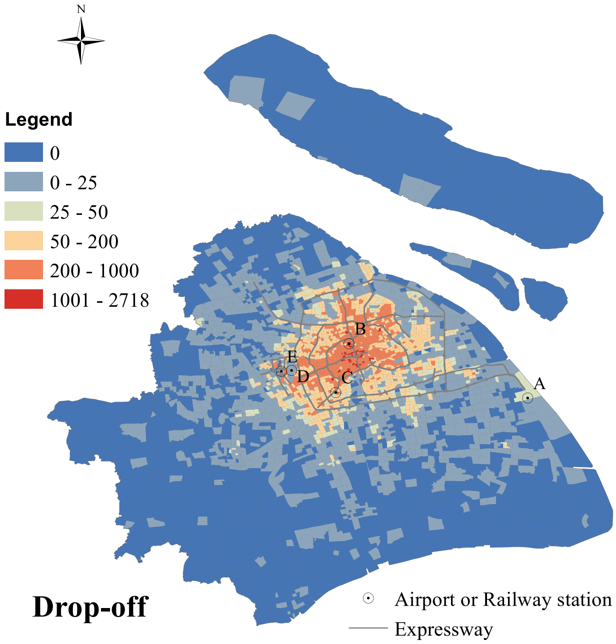
Fig. 2a shows the distribution of unit profit (CNY/hour) during the 6-hour study period with a mean of 63.86 CNY/hour and a standard deviation of 23.89 CNY/hour. Profits are below 47.36 CNY for the lowest 25% of drivers, while above 80.03 CNY for highest 25%, indicating a significant profit difference across drivers. Fig. 2b shows the distribution of occupancy rate (the quotient between occupied time and the total working time) with a mean of 0.48 and a standard deviation of 0.16. Occupancy rates are below 0.39 for the lowest 25% of drivers, while above 0.61 for the highest 25%, showing a similar dispersion among drivers.

Observed Pick-up and Drop-off Densities. The coordinates of pick-up and drop-offs are extracted from the dataset and associated to the closet node. Fig. 3 shows a heat map of the density of pick-up and drop-off during the study period. Major transportation hubs are labeled with A: Pudong international airport; B: Shanghai railway station; C: Shanghai south railway station; D: Hongqiao international airport; E: Hongqiao railway station. The map also shows major expressways in and around the city center and connecting Pudong international airport. As expected, the highest pick-up and drop-off densities are located around the city center and Pudong airport. The downtown area (approx. area: 660 km^2, roughly within the Outer

[4] source: https://www.travelchinaguide.com/cityguides/shanghai/transportation/taxi.html.



(a) Pick-up



(b) Drop-off

Fig. 3. Pick-up and drop-off density 5:30-11:30 am on a weekday in April, 2015 (count per km^2).

Ring Expressway) contributes about 80% of pick-ups and drop-offs and dominates the active operation regions of taxis. It is also noted that most short trips concentrate in downtown area, where most business, social, and cultural activities converged day and night, while long trips spread broadly and originate or end in regions containing airports or railway stations. There are very few trips on the islands and they are thus omitted in the remainder of the paper.

Experiment Setup. The whole study period is divided into three 2-hour time intervals: 5:30-7:30 am, 7:30-9:30 am and 9:30-11:30 am.

For each interval, the passenger arrival rate $\lambda(j)$ and vacant taxi density rate $\gamma(j)$ at any node j are assumed time-invariant and calculated from the historical data. Passenger arrival rate at each node is set as the average count of pick-ups per hour during each time interval. The calculation of vacant taxi density rate is more involved as it is a time-wise average of spatial rate. A buffer is created for each node, which is a circle centered at the node with a radius of 500 m, approximately the 75th percentile of all link length. At any given time instance (say, 8:00am), the vacant taxi density for a node is the number of vacant taxis within the buffer at that time instance (a snapshot) divided by the area of the buffer. Multiple snapshots are taken every 15 minutes over each 2-hour time interval and the average over all snapshots is used as the vacant taxi density for the 2-hour period.

The matching node set $N(j)$ is comprised of nodes within 1 *km* radius of node j, that is, the vacant taxi is eligible to be matched with passengers within 1 *km* crow-fly distance. The matching radius is set such that both picking up along the roads and e-hailing at high demand density areas are accommodated. It is conceivably higher in a low demand area or period where drivers need to drive a relatively long distance to pick up passengers with e-hailing. Given that the data come from a traditional taxi company and the percentage of e-hailing is conceivably small, the radius is set to be relatively short.

Passenger destination probability is calculated directly based on Eq. (3) from the data for each 2-hour time interval.

5.2. Computational performance

The value iteration algorithm is coded in Python 3.5 with NumPy. All computations are carried out on a workstation with an eight-core 3.0GHz Xeon E5-1660 processor and 64GB RAM. The discount factor is set as 0.95. For each 2-hour time interval, the running time per iteration is about 16.15 min (969 sec) when for-loops are used, and is reduced to about 1.98 min (118 sec) with matrix operations (Section 4), an 8x speed-up. It is expected that the speed-up will be higher with more cores to process the matrix operations in parallel, and real-time efficiency can thus be achieved.

Fig. 4a shows the maximum relative difference of value function over successive iterations as a function of the number of iterations. The relative value function difference reaches below 0.01 at about the 25th iteration. Fig. 4b shows the relative change in routing policy as a function of the number of iterations. The policy change is calculated as the fraction of states whose optimal action changes from the last iteration. The relative policy difference reaches below 0.01 at about 22nd iteration. The convergence patterns are similar across time intervals.

5.3. Evaluation

In this section, the optimal routing policy is compared with a number of baseline heuristics using two metrics: unit profit and occupancy rate. The metrics are generated by executing the optimal policy and heuristics in the network for the study period of 5:30 through 11:30 am, and the developed taxi/passenger matching model and passenger destination model are postulated as the true model based on which taxi and passenger matching is simulated. Comparisons between the optimal policy and heuristics provide insights into the value of the proposed method and are precursor to real world evaluation.

These comparisons aim to confirm the advantage of the proposed method over some commonly used heuristics, under the condition that perfect models of matching and passenger destination choice are available, which gives an indication of its potential performance in the real world, when such perfect models usually do not exist.

The comparison of the optimal routing policy with the observed taxi driver routing choices cannot be made until the optimal routing policy is implemented in the real world. Another way is to build a high-fidelity traffic simulation testbed where taxi driver behavioral models are calibrated using observed trajectory data, and then evaluate the optimal routing policy against the calibrated, simulated drivers in the simulation testbed. The comparison with real world drivers or simulated drivers calibrated against real world data is important in assessing the method's value. However, implementing the proposed method in the field or developing and calibrating a high-fidelity traffic simulation model is beyond the scope of the study, and left for future research.

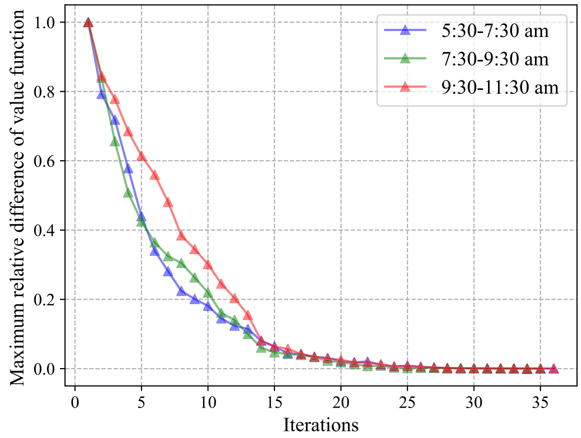
5.3.1. Heuristics

Three heuristics are defined as follows in increasing order of sophistication.

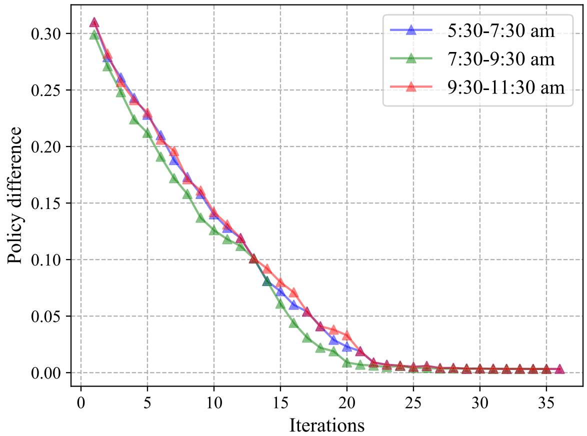
Random walk. A vacant taxi chooses an outgoing link randomly and once matched with a passenger, takes the fastest path to deliver the passenger. This is the simplest strategy.

Global hotspot. A vacant taxi heads toward the zone with the highest demand density (number of pick-ups per km^2), following the fastest path to the central node of the zone (the node closet to the centroid of the zone), and do random walks in this zone until matched with a passenger. Once matched, it takes the fastest path to deliver the passenger. Note that the taxi could get matched on the way to the highest demand density zone.

Local hotspot. The global hotspot strategy can be inefficient when the taxi is far from the highest demand density zone. A more sensible strategy is to move to higher demand zones sequentially (see, e.g., Wong et al., 2015a, 2015b) taxi driver behavioral studies by. This strategy can be viewed as a spatially aggregated, partially myopic approximation of the proposed optimal routing policy. Note that a majority of previous taxi routing studies (see, e.g., Yuan et al., 2013; Dong et al., 2014; Qu



(a) Value Function Difference vs. Iterations



(b) Policy Difference vs. Iterations

Fig. 4. Convergence of the Value Iteration Algorithm.

et al., 2014; Huang et al., 2015) recommend cruising routes with a maximum cruising distance or time, a constraint that is needed to avoid unrealistically long routes, necessitated by the myopic nature of their methods. The local hotspot heuristic is in the general family of partially myopic strategies.

The study area is divided into equal-sized square cells to represent roughly the range of a local area. The length of each cell is set at the 75th percentile of straight line distance from drop-off location to the next pick-up location (not the actual search distance), which equals 5 km approximately. A zone (which is smaller than a cell) is considered a member of a cell if its central node falls in the cell. The queuing-based model developed in Section 3 is applied to match a vacant taxi with passengers while the taxi moves following the Local hotspot strategy, and the Local hotspot strategy is implemented as follows:

Step 0: Calculate the demand density in each zone during each of the 2-hour period.

Step 1: For a vacant taxi start from node i, calculate the shortest path P_z (in terms of a sequence of nodes) from i to the centroid of the zone with the highest demand density in the current cell.

Step 2: Check if the taxi is matched with a passenger at each node while moving along path P_z

Step 2-1: If the taxi is not matched with any passenger, the next state is the subsequent node on path P_z, and time t is updated.

Step 2-2: If the taxi is matched with a passenger during the movement, the next state is the destination of the passenger and time t is updated; go to Step 1.

Step 3: Let the taxi move randomly within the current zone for 15 min, and check if it is matched with a passenger at each node.

Step 3-1: If the taxi not matched with any passenger, the next state is the sink node of the link it takes, and time t is updated.

Step 3-2: If the taxi is matched with a passenger during the movement, the next state is the destination of the passenger and time t is updated; go to Step 1.

Step 4: Pick a neighboring cell C that contains the highest demand density zone Z among all zones in all neighboring cells. Calculate the shortest path P_c (in terms of a sequence of nodes) from current node to the centroid of zone Z with the highest demand density in the chosen cell C.

Step 5: Check if the taxi is matched with a passenger at each node while moving along path P_c.

Step 5-1: If the taxi is not matched with any passenger, the next state is the subsequent node on path P_c, and time t is updated.

Step 5-2: If the taxi is matched with a passenger, the next state is the destination of the passenger and time t is updated; go to Step 1.

Step 6: Steps 1–5 are repeated until t reaches the end of the whole study period.

5.3.2. Trajectory simulation

A simulation of a single taxi's trajectory starting from various locations for the 6-hour study period is conducted according to each of the routing strategies. The parameters of the taxi/passenger matching model and the passenger destination probability model are the same as those used in generating the optimal routing policy (Section 5.1). The simulation of the optimal routing policy is needed because 1) the optimal value function obtained from solving the Bellman equation has a discount factor of 0.95 for convergence reason, yet the profit in real life should not be discounted given the relatively short time intervals, that is, one dollar earned now and one hour later should be treated as equal valued; 2) the occupancy rate is not available from solving the Bellman equation.

A strategy $\mu(i, H)$ specifies a probability vector associated with outgoing links, i.e., $\mu(i,H) \to (p_1, \ldots, p_a, \ldots, p_{|A(i)|}), a \in A(i),$ for each node i and routing history H. For the optimal routing policy, random walk and global hotspot heuristics, H is empty as the action does not depend on the routing history. For the local hotspot heuristic, the routing history H represents whether the cruising time in the current zone has reached 15 minutes. The optimal routing policy and the global hotspot heuristic are deterministic strategies, in that exactly one of the outgoing links at any node is assigned probability 1. The random walk and local hot spot heuristics are random strategies, in that outgoing links of certain nodes are assigned probabilities other than 0 or 1. A clock is advanced along the simulated trajectory to determine whether the simulation has reached the end of the study period.

The execution of a strategy from any given node i could result in multiple realizations of trajectories due to the random processes of passenger arrival, competition with other vacant taxis, and passenger destination choice. Following are the steps to simulate a single taxi trajectory for a specific strategy:

1. For a taxi at node i, an action $a = (i, j)$ is chosen according to the specific strategy.
2. The location of a matched passenger, h, is sampled according to matching probability, as in Eq. (2).
3. If no matching happens (with probability $1 - \sum_{h\in N(j)} p_{a,h}$), the taxi moves to node j and the clock is advanced by t_a. The sampling of the location of a matched passenger, i.e. Step 2, is then repeated.
4. If matching happens, the taxi moves to the passenger location, node h, and the clock is advanced by $\tau_a + \mathcal{T}_{j\to h}$.
5. The passenger destination i' is sampled according to Eq. (3). The taxi moves to i', and the clock is advanced by $\mathcal{T}_{h\to i'}$.
6. Steps 1 through 5 are repeated until the clock reaches the end of the study period.

For any given strategy, 1200 trajectories are simulated from each initial state (node) where there is a positive number of observed vacant taxies at that node at 5:30 am. For each trajectory, the unit profit (CNY/hour), occupancy rate (percentage of time with a passenger onboard) and other relevant measures (such as time spent in each pre-defined cells) are calculated. The average unit profit and occupancy rate for each node is an estimate of the expected unit profit and occupancy rate. The sample size has been increased until the sample average at each node stabilizes.

5.3.3. Results

Unit Profit and Occupancy Rate Over the Network. First the average unit profit and occupancy rate over the network for each strategy is presented. Taxis do not start randomly over the network. Fig. 5 shows a heat map of the starting locations of vacant taxis at 5:30 am. The high density starting locations are in the city center and its surrounding areas. The most flourishing ones are indicated on Fig. 5: (1) Jiali Sleepless City in Zhabei District, the Central Ring Commercial Circle and Zhenru Commercial Center to the North; (2) Around Shanghai rail way station; (3) Around Hongqiao airport and Hongqiao railway station to the West; and (4) Meilongzhen Plaza, the Commercial Center in Minhang District to the South, with some university campuses. To account for the starting position distribution, the empirical distribution of average unit profit (or occupancy rate) over the network is weighted by the starting position distribution. Practically, each sample average for a given node is duplicated by the number of observed vacant taxis from that node at 5:30 am, and the resulting data points are used to plot the histogram in Fig. 6 and the average is taken over all data points to generate Table 2.

Table 2 shows that the optimal policy performs the best in each interval in terms of both average unit profit and occupancy rate. Specifically, over the 6 hours, the average unit profit for the optimal policy is 23.0% higher than the random

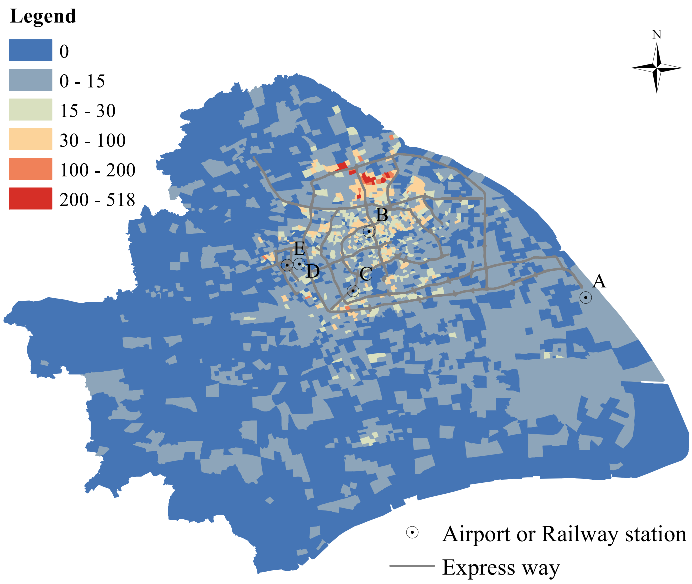


Fig. 5. Vacant taxi density at 5:30am (count per km^2).

Table 2
Average unit profit and occupancy rate in the morning time intervals.

Strategy	Unit profit (CNY/hour)				Occupancy rate			
	5:30-7:30	7:30-9:30	9:30-11:30	All intervals	5:30-7:30	7:30-9:30	9:30-11:30	All intervals
Random walk	61	78	72	71	0.40	0.45	0.42	0.42
Global hotspot	68	77	77	74	0.44	0.46	0.47	0.45
Local hotspot	69	86	85	80	0.46	0.47	0.48	0.48
Optimal policy	72	93	92	87	0.47	0.54	0.54	0.52

walk, 17.0% higher than the global hotspot and 8.4% higher than the local hotspot strategy. This suggests that it is beneficial to take into account subsequent pick-ups and drop-offs beyond the immediate next customer. The increases are higher during higher-demand time intervals, probably due to more room for improvement.

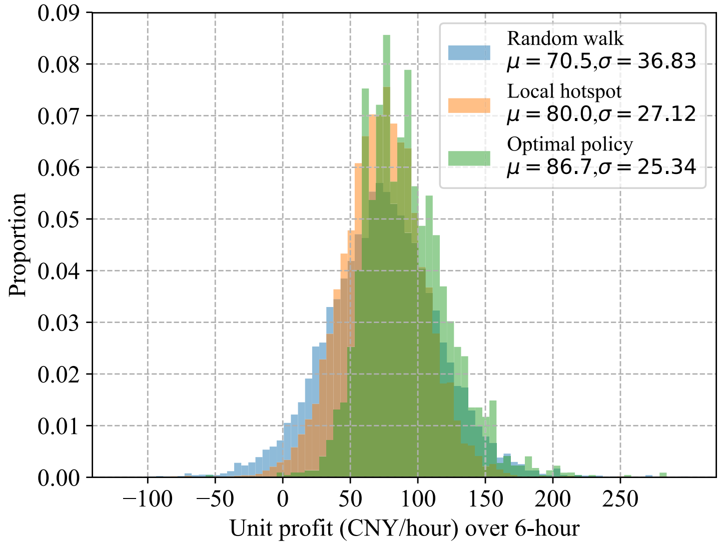
It is noted that the global hotspot strategy generates smaller unit profit than the local hotspot strategy, suggesting that travel time/distance to the next potential customer is a very important factor. Given its inferior performance, the global hotspot strategy will be omitted in the more detailed analyses to follow. The random walk strategy is kept for later analysis as it provides bottom line performance.

While maximizing taxi utilization is not the optimization criterion in the proposed optimization problem, it is still observed that the optimal policy is able to increase the average occupancy rate by 23.8% over the random walk, 15.6% over the global hotspot and 8.3% over the local hotspot strategy. It is intuitive that these two metrics are highly positively correlated, as less time spent on searching for passengers suggests more time spent on making money.

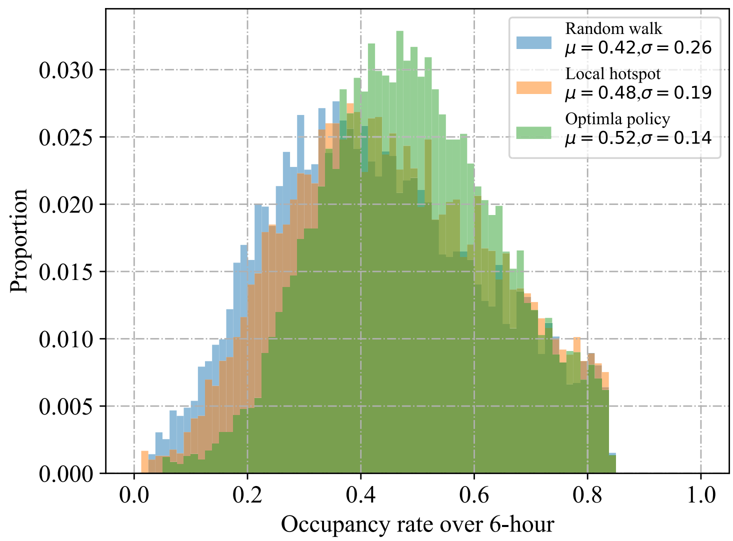
For a given strategy, the average unit profit is the highest during the morning peak interval (7:30 - 9:30 am) due to higher demand, and it becomes only slightly lower for late morning interval (9:30 - 11:30 am), suggesting that taxi trips might have a less pronounced morning peak than regular commuter trips. The flat pattern is also present for the average occupancy rate.

Fig. 6 shows the distribution of unit profit and occupancy rate for the 6-hour study period for each of the three remaining strategies: optimal policy (green), local hotspot (orange) and random walk (blue). It can be seen that the distributions for the optimal policy shift to the right, suggesting an across-the-board increase instead of isolated extremely large increases from certain locations.

To further understand the spatial pattern differences among the different strategies, analyses are done by starting locations. Fig. 7(a) shows the difference in average unit profit by starting location between the optimal policy and random walk strategy. Not surprisingly, a taxi following the optimal policy can make more profit than random walk no matter where it starts from (yellow through red). Fig. 7(b) shows the difference in average unit profit by location between the optimal policy and local hotspot strategy. From most starting locations, the average unit profit of taxis taking the optimal policy is higher. The relationship is reversed for some locations (blue), due to the fact that the local hotspot strategy has a state space larger than the optimal policy, expanded by including routing history in its state. The optimal policy is optimal among all policies defined based on the same state space, that is, the nodes, but is not necessarily so compared to a more flexible strategy.



(a) Unit profit (CNY/hour)



(b) Occupancy rate

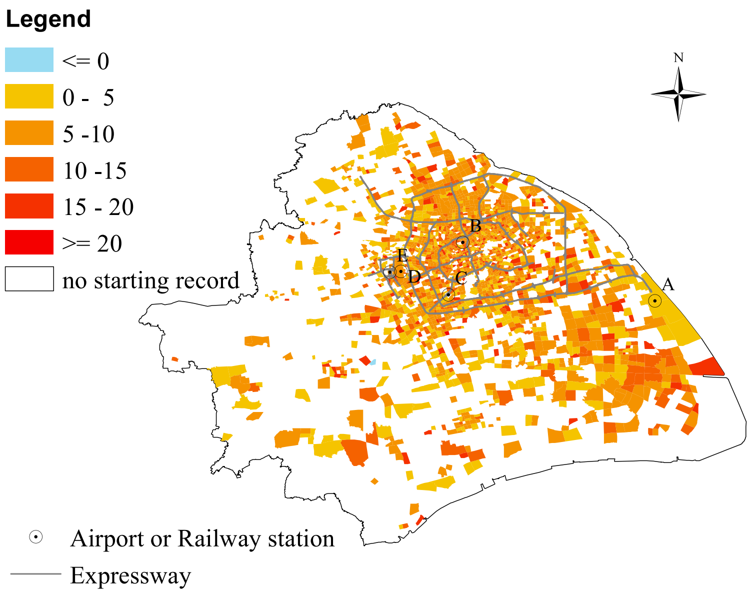
Fig. 6. Distribution of unit profit and occupancy rate for three strategies 5:30-11:30 am.

Spatial Patterns of Solutions. The position of a taxi at any time starting from a given location following a given strategy can be described by the coordinates (X, Y) in a two-dimensional plane, where X and Y are the projected longitude and latitude (in meters) respectively (Maling, 2013). X and Y are continuous random variables given the underlying passenger matching and destination choice process. To obtain probability distributions and summary statistics of the two random variables, the study area is discretized into a grid where each square cell is 3 km long, and the fraction of time the taxi spent in each cell over the 6-hour study period is the empirical probability of (X, Y) in that cell, $p_{x,\,y}$.

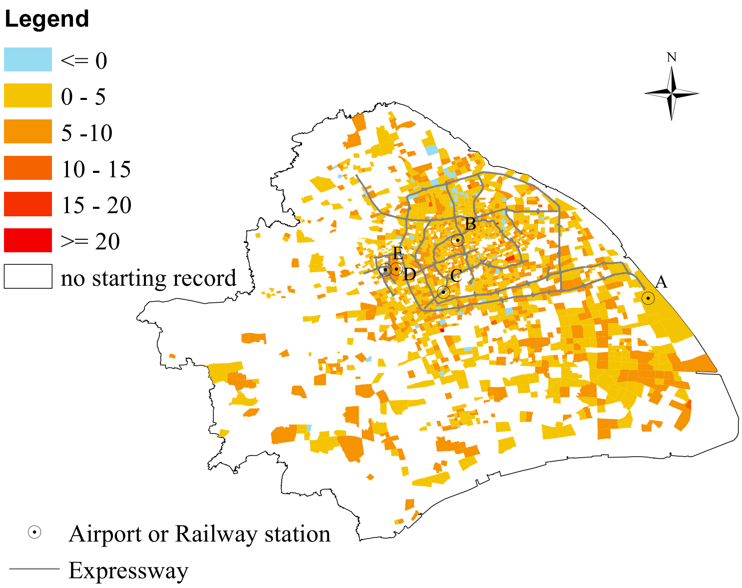
The Hellinger distance (Le Cam and Yang, 2012) is used to measures the difference between two probability distributions P and P^*, that is,

$$H(P, P^*) = \frac{1}{\sqrt{2}}\sqrt{\sum_{(x,y)\in G_{(x,y)}} (\sqrt{p_{x,y}} - \sqrt{p^*_{x,y}})^2}. \tag{16}$$

The Hellinger distance is between 0 and 1, and a larger value indicates a larger difference. Table 3 presents the Hellinger distance between the distribution of visited locations of the optimal policy and random walk, and between that of the



(a) Optimal Policy vs Random Walk



(b) Optimal Policy vs Local Hotspot

Fig. 7. Differences in unit profit of optimal vs. heuristic routing by starting zones.

optimal policy and local hotspot for four different starting locations, of which two are major transportation hubs and two are major commercial areas. As expected, the difference between the optimal policy and random walk is larger than the difference between the optimal policy and local hotspot.

Fig. 8 shows the empirical distributions of visited locations from four starting locations. While the movement pattern differs across starting locations, a similar trend can be observed, that is, the trajectories are more widely distributed for the two heuristic strategies. For a taxi starting from a transportation hub (Pudong Airport, around Hongqiao Airport or Hongqiao Railway Station), the trajectory following the optimal policy is mostly between the airports and railway station. The trajectory starting from either of two commercial areas is distributed more widely, probably due to more diversity in the origins and destinations of passengers in commercial areas compared to transportation hubs.

Table 4 shows summary statistics of visited locations for the three strategies starting from four different locations. A larger X means more east and a larger Y means more north. A positive covariance suggests the trajectory is more north-

Fig. 8. Empirical probability density function of the visited location from a starting location (red star). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Difference between distributions of visited locations.

Starting location	Optimal vs Random walk	Optimal vs Local hotspot
Pudong Airport	0.4117	0.3023
Around Hongqiao Airport and Hongqiao Railway Station	0.3928	0.3653
Central Ring Commercial Circle and Zhenru Commercial Center	0.5424	0.4969
Meilongzhen Plaza, the South Commercial Center	0.5128	0.4879

Table 4
Summary statistics on spatial distributions of visited locations (*X, Y* are in meter).

Starting location	Strategy	$E[X]$	$E[Y]$	$Var[X]$	$Var[Y]$	$Cov(X, Y)$
Pudong Airport	Optimal policy	1,353,473	3,634,788	2,144,078	3,990,332	-1309082
	Random walk	1353955	3,633,768	4360890	5,317,421	-2126156
	Local hotspot	1353821	3,633,815	4016870	4,931,743	-1126138
Around Hongqiao Airport and Hongqiao Railway Station	Optimal policy	1,351,548	3,638,640	1219876	1,780,287	1005313
	Random walk	1,352,130	3638216	2,454,627	2,245,158	-1141663
	Local hotspot	1351999	3,638,527	2377582	1750020	-1541638
Central Ring Commercial Circle and Zhenru Commercial Center	Optimal policy	1,352,473	3,634,619	4,141,008	3190332	-2309082
	Random walk	1,352,004	3,645,415	5252598	4,246,513	-1396572
	Local hotspot	1,352,015	3645085	4065228	4,005,653	-1854668
Meilongzhen Plaza, the South Commercial Center	Optimal policy	1,351,575	3632361	3745147	4343224	1402910
	Random walk	1,351,861	3,631,450	5,557,328	4,820,842	1499701
	Local hotspot	1,351,703	3,631,959	4,010,645	5,268,561	2473367

east/southwest than northwest/southeast, and a negative covariance suggests that the trajectory is more northwest/southeast than northeast/southwest.

For a taxi starting from one of the transportation hubs and the Meilongzhen Plaza, the expected location following the optimal policy is to the northwest of that following the random walk or local hotspot strategy (smaller $E[X]$ and larger $E[Y]$), while for a taxi starting from Central Ring Commercial Circle and Zhenru Commerical Center, the expected location following the optimal policy is to the northeast of that following the two heuristic strategies. Nor surprisingly, random walk has the largest variance among the three strategies for each of the four locations.

6. Conclusions and future directions

In this paper, the single vacant taxi routing problem is investigated, which aims at maximizing long-term expected profit over the complete working period. Theoretical contributions in modeling and problem formulation as well as practical contributions in computational efficiency are provided, which builds the foundation for real-world implementations of taxi routing. A queueing theory-based model for matching taxis and passengers is proposed to account for competition from other taxis and use of e-hailing apps. The problem is formulated as a Markov decision process, taking into account the impact of current decisions on future return over multiple pickups and drop-offs. To improve computation efficiency, an enhanced value iteration algorithm for solving the MDP is proposed via matrix operations.

Numerical experiments in a mega city suggest that matrix operation helps to achieve 8x speed-up in computation. Simulation experiments are conducted to compare the performance of the proposed strategy with a number of baseline heuristics. The MDP formulation improves unit profit by 23.0% and 8.4% over the random walk and local hotspot heuristic respectively; and improve occupancy rate by 23.8% and 8.3% respectively. Empirical spatial distributions of taxi location from a few starting locations and following the various strategies are obtained from the simulation, and the heuristics are shown to have in general more spread-out spatial patterns than the optimal policy. Specifically the optimal policy concentrates between major transportation hubs if starting from one of them. The trajectory starting from either of two commercial areas is distributed more widely than that from a transportation hub.

The research can be extended in a number of directions. First, time-dependent customer demand and empty taxi distribution will be considered. Formulating the problem as an MDP in a time-space network is computationally challenging given the explosion of the state space. Approximate Dynamic Programming (ADP) is a potential tool (see, e.g., Larsen et al., 2004; George and Powell, 2006), which approximates the value function and avoids the evaluation of all possible future states, the so-called complete sweep. Secondly, the empty taxi density is exogenous in the current approach, which should be relaxed to consider the feedback loop between optimal taxi routing and the matching probability change due to multiple taxies routed to the same location, much like the congestion effect in traditional traffic assignment. Future research can consider dynamically rebalancing the empty taxis by moving cruising taxis from over-supply regions to over-demand regions. Thirdly, to ensure real-time performance and thus its applicability to a deployable system, innovative computing architecture that involves a combination of cloud computing and embedded systems on vehicles can be explored to speed up the computation. Fourthly, this study matches a vacant taxi with passengers based on locality following a greedy method, i.e., finding the closest taxi to serve a passengers request. A queuing strategy is applied to serve the passengers with the principle of first-come-first-served. In real world, such kind of greedy matching could lead to a spatiotemporal mismatch between taxis and passengers in the long run. Future study may look into a more flexible modeling of spatiotemporal matching between passengers and vacant taxis in a coordinated way. Finally, as only fulfilled requests are considered as the demand in the study, the future work may also consider quantifying the actual demand by including the unfulfilled possible requests, such as using the request data from e-hailing platforms (see, e.g., Ke et al., 2017; Tong et al., 2017) and reavling the unconstrained demand from the observed data which has been applied in the demand management in airline industry (Ratliff et al., 2008), as well as for taxi market (Qian and Ukkusuri, 2017).

References

Bellman, R., 1957. A Markovian Decision Process. Technical Report. DTIC Document.

Dong, H., Zhang, X., Dong, Y., Chen, C., Rao, F., 2014. Recommend a profitable cruising route for taxi drivers. In: Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on. IEEE, pp. 2003–2008.

Ferrucci, F., Bock, S., Gendreau, M., 2013. A pro-active real-time control approach for dynamic vehicle routing problems dealing with the delivery of urgent goods. Eur. J. Oper. Res. 225 (1), 130–141.

George, A.P., Powell, W.B., 2006. Adaptive stepsizes for recursive estimation with applications in approximate dynamic programming. Mach. Learn. 65 (1), 167–198.

Han, M., Senellart, P., Bressan, S., Wu, H., 2016. Routing an autonomous taxi with reinforcement learning. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. ACM, pp. 2421–2424.

Hosni, H., Naoum-Sawaya, J., Artail, H., 2014. The shared-taxi problem: formulation and solution methods. Transp. Res. Part B 70, 303–318.

Hou, Y., Li, X., Zhao, Y., Jia, X., Sadek, A.W., Hulme, K., Qiao, C., 2013. Towards efficient vacant taxis cruising guidance. In: Global Communications Conference (GLOBECOM), 2013 IEEE. IEEE, pp. 54–59.

Hu, X., Gao, S., Chiu, Y.-C., Lin, D.-Y., 2012. Modeling routing behavior for vacant taxicabs in urban traffic networks. Transp. Res. Record 2284, 81–88.

Huang, J., Huangfu, X., Sun, H., Li, H., Zhao, P., Cheng, H., Song, Q., 2015. Backward path growth for efficient mobile sequential recommendation. IEEE Trans. Knowl. Data Eng. 27 (1), 46–60.

Hwang, R.-H., Hsueh, Y.-L., Chen, Y.-T., 2015. An effective taxi recommender system based on a spatio-temporal factor analysis model. Inf. Sci. (Ny) 314, 28–40.

Ke, J., Zheng, H., Yang, H., Chen, X.M., 2017. Short-term forecasting of passenger demand under on-demand ride services: a spatio-temporal deep learning approach. Transp. Res. Part C 85, 591–608.

Kim, H., Oh, J.-S., Jayakrishnan, R., 2005. Effect of taxi information system on efficiency and quality of taxi services. Transp. Res. Record (1903) 96–104.

Larsen, A., Madsen, O.B., Solomon, M.M., 2004. The a priori dynamic traveling salesman problem with time windows. Transp. Sci. 38 (4), 459–472.

Larson, R.C., Odoni, A.R., 1981. Urban operations research. Prentice-Hall.

Le Cam, L., Yang, G.L., 2012. Asymptotics in Statistics: some Basic Concepts. Springer Science & Business Media.

Lee, A., Savelsbergh, M., 2015. Dynamic ridesharing: is there a role for dedicated drivers? Transp. Res. Part B 81, 483–497.

Li, X., Pan, G., Wu, Z., Qi, G., Li, S., Zhang, D., Zhang, W., Wang, Z., 2012. Prediction of urban human mobility using large-scale taxi traces and its applications. Front. Comput. Sci. 6 (1), 111–121.

Long, J., Szeto, W., Du, J., Wong, R., 2017. A dynamic taxi traffic assignment model: a two-level continuum transportation system approach. Transp. Res. Part B 100, 222–254.

Lowalekar, M., Varakantham, P., Jaillet, P., 2016. Online spatio-temporal matching in stochastic and dynamic domains. In: AAAI, pp. 3271–3277.

Lv, Z., Wu, J., Yao, S., Zhu, L., 2017. Fcd-based analysis of taxi operation characteristics: a case of shanghai. J. East China Normal Univ. 3, 133–144.

Maling, D.H., 2013. Coordinate systems and map projections. Elsevier.

Masoud, N., Jayakrishnan, R., 2017. A real-time algorithm to solve the peer-to-peer ride-matching problem in a flexible ridesharing system. Transp. Res. Part B 106, 218–236.

Mitrović-Minić, S., Krishnamurti, R., Laporte, G., 2004. Double-horizon based heuristics for the dynamic pickup and delivery problem with time windows. Transp. Res. Part B 38 (8), 669–685.

Moreira-Matias, L., Gama, J., Ferreira, M., Damas, L., 2012. A predictive model for the passenger demand on a taxi network. In: 2012 15th International IEEE Conference on Intelligent Transportation Systems. IEEE, pp. 1014–1019.

Moreira-Matias, L., Gama, J., Ferreira, M., Mendes-Moreira, J., Damas, L., 2013. Predicting taxi–passenger demand using streaming data. IEEE Trans. Intell. Transp. Syst. 14 (3), 1393–1402.

Phithakkitnukoon, S., Veloso, M., Bento, C., Biderman, A., Ratti, C., 2010. Taxi-aware map: Identifying and predicting vacant taxis in the city. In: International Joint Conference on Ambient Intelligence. Springer, pp. 86–95.

Powell, J.W., Huang, Y., Bastani, F., Ji, M., 2011. Towards reducing taxicab cruising time using spatio-temporal profitability maps. In: SSTD. Springer, pp. 242–260.

Qian, X., Ukkusuri, S.V., 2015. Spatial variation of the urban taxi ridership using gps data. Appl. Geogr. 59, 31–42.

Qian, X., Ukkusuri, S.V., 2017. Time-of-day pricing in taxi markets. IEEE Trans. Intell. Transp. Syst. 18 (6), 1610–1622.

Qian, X., Zhang, W., Ukkusuri, S.V., Yang, C., 2017. Optimal assignment and incentive design in the taxi group ride problem. Transp. Res. Part B 103, 208–226.

Qin, G., Li, T., Yu, B., Wang, Y., Huang, Z., Sun, J., 2017. Mining factors affecting taxi drivers incomes using gps trajectories. Transp. Res. Part C 79, 103–118.

Qu, M., Zhu, H., Liu, J., Liu, G., Xiong, H., 2014. A cost-effective recommender system for taxi drivers. In: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, pp. 45–54.

Ratliff, R.M., Rao, B.V., Narayan, C.P., Yellepeddi, K., 2008. A multi-flight recapture heuristic for estimating unconstrained demand from airline bookings. J. Revenue Pricing Manage. 7 (2), 153–171.

Seow, K.T., Dang, N.H., Lee, D.-H., 2010. A collaborative multiagent taxi-dispatch system. IEEE Trans. Autom. Sci. Eng. 7 (3), 607–616.

Sutton, R.S., Barto, A.G., 1998. Reinforcement learning: An introduction. The MIT Press, Cambridge, MA.

Szeto, W.Y., Wong, R.C.P., Wong, S.C., Yang, H., 2013. A time-dependent logit-based taxi customer-search model. Int. J. of Urban Sci. 17 (2), 184–198.

Thomas, B.W., White, C.C., 2004. Anticipatory route selection. Transp. Sci. 38 (4), 473–487.

Tong, Y., Chen, Y., Zhou, Z., Chen, L., Wang, J., Yang, Q., Ye, J., Lv, W., 2017. The simpler the better: a unified approach to predicting original taxi demands based on large-scale online platforms. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, pp. 1653–1662.

Verma, T., Varakantham, P., Kraus, S., Lau, H.C., 2017. Augmenting decisions of taxi drivers through reinforcement learning for improving revenues. In: Proceedings of the International Conference on Automated Planning and Scheduling, Pittsburgh, PA, USA, pp. 18–23.

van der Walt, S.v.d., Colbert, S.C., Varoquaux, G., 2011. The numpy array: a structure for efficient numerical computation. Comput. Sci. Eng. 13 (2), 22–30.

Wong, K., Wong, S.C., Bell, M.G., Yang, H., 2005. Modeling the bilateral micro-searching behavior for urban taxi services using the absorbing markov chain approach. J. Adv. Transport. 39 (1), 81–104.

Wong, K., Wong, S.C., Yang, H., 2001. Modeling urban taxi services in congested road networks with elastic demand. Transp. Res. Part B 35 (9), 819–842.

Wong, K., Wong, S.C., Yang, H., Wu, J., 2008. Modeling urban taxi services with multiple user classes and vehicle modes. Transp. Res. Part B 42 (10), 985–1007.

Wong, R., Szeto, W., Wong, S., 2014b. A cell-based logit-opportunity taxi customer-search model. Transp. Res. Part C 48, 84–96.

Wong, R., Szeto, W., Wong, S., 2015. A two-stage approach to modeling vacant taxi movements. Transp. Res. Part C 59, 147–163.
Wong, R.C., Szeto, W., Wong, S., 2015. Sequential logit approach to modeling the customer-search decisions of taxi drivers. Asian Transp. Studies 3 (4), 398–415.
Wong, R.C.P., Szeto, W.Y., Wong, S., Yang, H., 2014a. Modelling multi-period customer-searching behaviour of taxi drivers. Transportmetrica B 2 (1), 40–59.
Xie, J., Nie, Y.M., Liu, X., 2017. Testing the proportionality condition with taxi trajectory data. Transp. Res. Part B 104, 583–601.
Xu, J., Rahmatizadeh, R., Bölöni, L., Turgut, D., 2017. Real-time prediction of taxi demand using recurrent neural networks. IEEE Trans. Intell. Transp. Syst.
Yang, H., Leung, C.W., Wong, S.C., Bell, M.G., 2010. Equilibria of bilateral taxi–customer searching and meeting on networks. Transp. Res. Part B 44 (8–9), 1067–1083.
Yang, H., Wong, S.C., 1998. A network model of urban taxi services. Transp. Res. Part B 32 (4), 235–246.
Yang, H., Yang, T., 2011. Equilibrium properties of taxi markets with search frictions. Transp. Res. Part B 45 (4), 696–713.
Yang, H., Ye, M., Tang, W.H., Wong, S.C., 2005. Regulating taxi services in the presence of congestion externality. Transp. Res. Part A: 39 (1), 17–40.
Yang, H., Ye, M., Tang, W.H.-C., Wong, S.C., 2005. A multiperiod dynamic model of taxi services with endogenous service intensity. Oper. Res. 53 (3), 501–515.
Yang, L., Kwan, M.-P., Pan, X., Wan, B., Zhou, S., 2017. Scalable space-time trajectory cube for path-finding: a study using big taxi trajectory data. Transp. Res. Part B 101, 1–27.
Yang, T., Yang, H., Wong, S.C., 2014. Taxi services with search frictions and congestion externalities. J. Adv. Transp. 48 (6), 575–587.
Yuan, N.J., Zheng, Y., Zhang, L., Xie, X., 2013. T-Finder: a recommender system for finding passengers and vacant taxis. IEEE Trans. Knowl Data Eng. 25 (10), 2390–2403.
Zha, L., Yin, Y., Du, Y., 2017. Surge pricing and labor supply in the ride-sourcing market. Transp. Res. Procedia 23, 2–21.
Zhang, D., He, T., Lin, S., Munir, S., Stankovic, J.A., 2015. Online cruising mile reduction in large-scale taxicab networks. IEEE Trans. Parallel Distribut. Syst. 26 (11), 3122–3135.
Zhang, K., Feng, Z., Chen, S., Huang, K., Wang, G., 2016. A framework for passengers demand prediction and recommendation. In: IEEE International Conference on Services Computing (SCC), pp. 340–347.
Zhao, K., Khryashchev, D., Freire, J., Silva, C., Vo, H., 2016. Predicting taxi demand at high spatial resolution: Approaching the limit of predictability. In: Big Data (Big Data), 2016 IEEE International Conference on. IEEE, pp. 833–842.