Game Theoretic Considerations for Optimizing Efficiency of Taxi Systems

Jiarui Gan
The Key Lab of Intelligent Information Processing, ICT, CAS
University of Chinese Academy of Sciences
Beijing 100190, China
ganjr@ics.ict.ac.cn

Bo An
School of Computer Engineering
Nanyang Technological University
Singapore 639798
boan@ntu.edu.sg

Abstract

Taxi service is an indispensable part of public transport in modern cities. The taxi system is operated by a large number of self-controlled drivers lacking of centralized scheduling and control, which makes it inefficient, difficult to analyze and optimize. It is thus important to take into account taxi drivers' strategic behavior in order to optimize taxi systems' efficiency. This paper reviews existing taxi system researches for modeling taxi system dynamics, introduces the taxi system efficiency optimization problem, and presents a game theoretic approach for optimizing the efficiency of taxi systems. Challenges and open issues in the taxi system efficiency optimization problem are also discussed.

Introduction

Taxi service has long been an indispensable part of public transport in modern cities. Unlike other types of public transportation systems (e.g., bus and metro systems), the taxi system is a highly decentralized system operated by a large number of self-controlled taxi drivers who can freely choose their operation schedules and movements. Though the decentralized nature makes taxi service more flexible and accessible than other modes of public transport, it makes the taxi system inefficient. Many practical issues thereby arise in taxi markets, and how to optimize the efficiency of taxi systems becomes an important but challenging problem.

The difficulties are two-fold. Firstly, the taxi system can be affected by many factors, such as road condition, customer demand, fare price, etc., which not only depend on each other in very complex ways, but also vary with time. Secondly, the efficiency of the taxi system relies deeply on the taxi drivers' behaviors and their operation strategies. Whereas advanced transportation researches have provided thorough studies on the interdependencies among factors in the taxi system and have built taxi system models (Yang, Wong, and Wong 2002; Yang et al. 2005b), little has been done to answer how taxi drivers, as a group of intelligent agents, behave and choose their strategies, and how their behaviors and strategies affect the system in turn. To address the challenges, a game theoretic approach is proposed to model the taxi drivers' strategic behaviors. In the rest of this paper, beginning with a review of related works in both transportation and Artificial Intelligence (AI), we present the Taxi system Efficiency optimization Problem (TEMP), and then the game theoretic solution approach. Challenges and open issues of TEMP are also discussed in this paper.

Related Work

Researches concerning the models and economics of taxi-i market dated back to 1969, when Orr studied the inadequacy of traditional cost-demand theory in application to the taxi market research, with an analysis to the equilibrium of a competitive taxi market. Later, Douglas (1972) introduced an aggregate demand and supply model which assumes that the amount of customer demand for taxi services depends on the expected monetary and time cost, and the expected customer waiting time depends on the total vacant taxi-hours. The model provided a basis for extensive subsequent studies on the economics of taxi market. For instance, based on the model, Yang et al. incorporated congestion externalities and time variance in service intensity in their study (Yang et al. 2005a; 2005b). Besides, some other studies particularly investigated the taxi fare pricing. For instance, Schaller (1998) studied the effects of fare price increase on taxi market in New York. Kim and Hwang (2008) studied an incremental discount policy on the taxi fare with the objective of maximizing average profit of taxis. Yang et al. (2010) examined a nonlinear fare structure and showed its advantages over the existing linear fare structure in Hong Kong. Unfortunately, a significant limitation of these researches is that taxi drivers' strategic behaviors are unconsidered as they compete with each other to maximize their profits. These behaviors are the key to the decentralized nature of the taxi market, and the main cause of the low efficiency of taxi markets.

AI techniques have provided a powerful tool to analyze human behaviors in a computational approach, and have been applied in many transportation studies, such as traffic control and intersection management (Dresner and Stone 2007; 2008; Bazzan 2009; Au, Shahidi, and Stone 2011; Pulter, Schepperle, and Böhm 2011) and, more relevantly, decision support for taxi drivers (Varakantham et al. 2012) and empirical multi-agent-based taxi system simulations (Alshamsi, Abdallah, and Rahwan 2009; Seow, Dang, and Lee 2010; Cheng and Nguyen 2011).
Taxi System Efficiency Optimization Problem
The goal of TEMP is to maximize the efficiency of a taxi system through adjusting the taxi fare. To this end, the first step is to know how system efficiency is affected by fare price in a taxi market. In this section, we present a multi-period taxi market model based on advanced transportation research (Yang, Wong, and Wong 2002; Yang et al. 2005b), and then formulate TEMP as a bilevel optimization program.

A Multi-period Taxi Market Model
A taxi market is a dynamic time-varying system. To deal with such variances, we discretize the optimization horizon $T$ (e.g., a whole day) into a set of equal-length periods, such that when the duration of each period is sufficiently short, the market can be treated as uniform in each period. Within a single period, say period $i$, the number of customers served by the whole taxi system is determined by average monetary and time cost of a trip (Yang, Wong, and Wong 2002), i.e.,

$$D^i(F^i, L^i, W^i) = \bar{D}^i \exp\{-\beta(F^i/\gamma + \phi_1 L^i + \phi_2 W^i)\}, \quad (1)$$

where $F^i$ is the average fare price, $L^i$ is the average travel time, and $W^i$ is the average customer waiting time; $\beta > 0$ is a sensitivity parameter; $\phi_1$ and $\phi_2$ are parameters used for converting time costs into monetary costs; $\gamma$ is the average number of passengers per ride; $D^i$ is the number of potential customers, which is obtained when the total cost is zero. The waiting time $W^i$ in turn depends on $D^i$ as

$$W^i(D^i, L^i, p^i) = \frac{\omega}{p^i \cdot N_T - D^i L^i / (\gamma \cdot \tau)}, \quad (2)$$

where $\omega > 0$ is a parameter depending on the density of taxi stands; $p^i$ is the Percentage of Working taxis (PoW); $N_T$ is the total number of taxis; and $N_T = D^i L^i / (\gamma \cdot \tau)$ represents vacant taxis in period $i$. It can be proven that when $F^i$, $L^i$ and $p^i$ are fixed, $D^i$ and $W^i$ are uniquely determined by Eqs. (1) and (2) (Yang et al. 2005b). Therefore, $D^i$ and $W^i$ are in fact implicit functions of $F^i, L^i$ and $p^i$. We denote them as $D^i = D^i(F^i, L^i, p^i)$ and $W^i = W^i(F^i, L^i, p^i)$.

Given the average trip distance $d^i$, the travel time can be represented by travel speed $V^i$ as $L^i = d^i / V^i$. Travel speed in a road network can be approximated by a linear function of number of on-road vehicles (Smith and Cruz 2012), which is furthermore linear to PoW $p^i$ as we assume that the number $N^i$ of non-taxi vehicles on the road is a period-specific constant. Thus, $V^i$ is a linear function of $p^i$, i.e., $V^i(p^i) = \mu(p^i \cdot N_T + N^i) + \lambda$, where $\mu$ and $\lambda$ are parameters depending on the road condition. We furthermore denote $L^i, D^i$, and $W^i$ as $L^i = L^i(p^i), D^i = D^i(F^i, p^i)$, and $W^i = W^i(F^i, p^i)$, respectively.

We adopt a distance-based fare structure (the model can also be extended to other fare structures): $F^i = f_0 + f^i \cdot (d^i - d_0)$, where $f_0$ is the initial charge and $d_0$ is the distance covered by $f_0$; $f^i$ is the charge rate for period $i$, i.e., the per unit distance charge. We optimize the fare structure through adjusting the charge rate $f^i$, and thus treat $F^i$ as a function $F^i(f^i)$. Accordingly, all the market factors, particularly the number $D^i$ of served customers, now depend on $f^i$ and $p^i$, i.e., $D^i = D^i(f^i, p^i)$. For ease of description, we denote a market factor over $T$ as a column vector with each component corresponding to a period. For example, we denote charge rate over all periods as $f = (f^i)$.

Model the Taxi drivers’ strategy
From a game theoretic perspective, the taxi drivers have a set $S$ of pure strategies, and they adopt a mixed strategy (strategy for simplicity), which is a probability distribution $x$ over their pure strategies in $S$. Each pure strategy is a schedule that specifies working and resting periods. Formally, a pure strategy is denoted as a 0/1 vector $x \in \{0, 1\}^n$, where $s^i=1$ (resp. $s^i=0$) represents working (resp. resting) in period $i$. Each pure strategy has to satisfy some constraints to ensure its feasibility in practice. For example:

- **Constraint 1 (C1):** A taxi driver should not work for more than $\tilde{n}_w$ periods in any schedule in $S$.
- **Constraint 2 (C2):** A taxi driver should not work continuously for more than $\tilde{n}_w$ periods in any schedule in $S$.

We assume that the taxi drivers adopt the same strategy. Thus when they choose strategy $x$, PoW is determined as:

$$p(x) = \sum_{s \in S} x_s \cdot s. \quad (3)$$

Furthermore, since taxi drivers are profit-driven, they always choose the best strategy which maximizes their utility, i.e.,

$$x^* \in \arg \max_{x \mid |x| > 0} U(f, p(x)). \quad (4)$$

The utility function $U(f, p)$ is defined as the sum of utilities in all periods, i.e., $U = \sum_i U^i$, and $U^i$ is defined as

$$U^i(f^i, p^i) = D^i \cdot F^i / (\gamma \cdot N_T) - p^i \cdot c_g \cdot \tau,$$

where $D^i \cdot F^i / (\gamma \cdot N_T)$ represents the income as $D^i \cdot (\gamma \cdot N_T)$ is the average number of trips each taxi serves, and $c_g$ is the cost of gasoline per unit time. Thus the fare price determines the taxi drivers’ strategy via the optimization in Eq. (4), and taxi drivers’ strategy in turn determines PoW via Eq. (3). $U(f, p)$ is strictly concave with respect to $p$ (Gan et al. 2013), so that given a convex feasible set of $p$ there is only one $p$ that maximizes $U$ (Boyd and Vandenberghe 2004). This means even if there are more than one solutions to Eq. (4), the solutions must all yield the same PoW, and a one-to-one correspondence from $f$ to $p$ is guaranteed.

A Bilevel Optimization Problem
We measure the system efficiency with the total number of served customers, and formulate a TEMP as the following bilevel optimization program.

$$\max_{f, x^*} D(f, p(x^*)) \quad (5)$$

subject to

$$x^* = \arg \max_{x \mid x \geq 0, 1^T x = 1} U(f, p(x)). \quad (6)$$

The model can also handle other measures of the system efficiency with the same form of optimization program as long as the optimization objective is a function of $f$ and $p$. To solve this bilevel optimization problem, a practical way is
to discretize the continuous fare price space into a small set of candidate prices, e.g. \( \{¥1.00, ¥1.20, \ldots, ¥5.00\}^n \), and solve the lower level program (Eq. (6)) under each of the candidate prices to find the optimal fare price. Thus the key is to compute the lower level program. Unfortunately, the lower level program suffers from a scalability issue caused by the exponential explosion of the taxi drivers’ pure strategy space. Next, we introduce the atom schedule method (ASM) to address the scalability issue.

**Solve TEMP—the Atom Schedule Method**

ASM works for TEMP problems with constraints C1 and C2 considered. The idea is to represent each feasible schedule (i.e., schedules satisfying C1 and C2) with a set of *atom schedules* (atom, for short), such that the set of atoms needed for representing all feasible pure strategies has a much smaller size than the pure strategy set. The lower level program is then reformulated as a much compactier one that optimizes over the set of atoms.

\[ s = \{o(j, k) | 1 \leq j \leq k \leq n, \quad s_j = s_{j+1} = \cdots = s_k = 1 \}. \]

Note that for notational simplicity, a pure strategy \( s \) is denoted as either a 0/1 vector or a set of atoms. Let the set of all possible pure strategies be \( S \), and the set of all needed atoms be \( O \). A weight \( w_o \) is assigned to each \( o \in O \) to represent the percentage of drivers who use it. It follows that PoW can be computed as:

\[
p^i = \sum_{o \in O} w_o \cdot \delta(o, i), \quad \forall i = 1, \ldots, n,
\]

where \( \delta(o(j, k), i) = 1, \text{ if } j \leq i \leq k \) (i.e., \( o \) covers period \( i \)), otherwise 0.

Thus \( p \) is now defined as a function of \( w = (w_o) \), denoted as \( p = p(w) \); and the optimization program can be reformulated as a compact one that takes \( w \) instead of \( x \) as the variable. Specially, when C2 are enforced on \( S \), we only need atoms no longer than \( n_c \) periods, so that

\[ O \subseteq \{o(j, k) | 1 \leq j \leq k \leq n_c, \quad 0 \leq k - j < n_c \}. \]

and, similar to PoW, \( q^i(w) \) is the percentage of taxis switching from working to resting in period \( i \). Namely,

\[
q^i(w) = \sum_{o \in O} w_o \cdot \delta'(o, i), \quad \forall i = 1, \ldots, n_c.
\]

Experimental Evaluations

We evaluate the effect of scheduling constraints and the scalability of ASM. This is in addition to experimental evaluations of the existing work (Gan et al. 2013).

Effects of the Constraints

We compare experimental results with and without constraints C1 and C2 to examine their effects. As shown in Figure 2(a), system efficiency peaks at ¥2.60 when constraints C1 and C2 are considered, whereas the curve continues to increase when constraints are ignored, leading to an imprecise optimal fare of ¥5.00 (or even higher). The extra system efficiency improvement is in fact unreachable due to impractical overworking of the taxi drivers. This can be seen from Figure 2(b), where the variances of PoW show that taxi drivers are reluctant to work during the peak time because of the scheduling constraints.

Performance of ASM

The scalability of ASM is examined with problems with up to 100 periods. Figures 3(a) and 3(b) depict the scalability of runtime and memory use of ASM, in comparison with scalability of the naive formulation (Eqs. (5)–(6)). Whereas the naive formulation runs out of memory at 15 periods, ASM can handle problems of up to 100 periods very easily.
Challenges and Open Issues

While the current model and algorithm are capable of handling TEMPs with specific settings, they are still inadequate for more extensive and complex real-world scenarios.

More Scalable Algorithms

In practice, customer demand and road condition might be different in different days of a week. The cycle of the taxi market is more likely to be a week rather than a day. In order to cover a whole week at the same level of granularity, more periods are required. Moreover, when the model needs to be more fine-grained to achieve higher accuracy, shorter periods, such as half an hour or even 10 minutes, are required and the number of periods increases accordingly.

Arbitrary Scheduling Constraints

To enhance the practical effectiveness of the optimized fare, it is also necessary to consider other types of scheduling constraints in some other real-world scenarios. For example, when the model is more fine-grained, each period has a shorter duration (e.g., half an hour), and it is unrealistic that a taxi driver rests or works for only one period each time since it takes time and effort to find parking when they switch. Therefore, the following constraints may need to be considered.

- **C3:** A taxi driver should not work continuously for less than \( n_c \) periods in any schedule in \( S \).
- **C4:** A taxi driver should not rest for less than \( n_r \) periods in any schedule in \( S \).

Besides, in a well regulated market, it is essential to take into account the impact of market regulations. For example, in some cities, taxi drivers are not allowed to switch from working to resting during peak times. The regulation was introduced because it was frequently reported that large numbers of taxi drivers refuse to take passengers in the name of a daily working switch (in these cities, a taxi is usually operated by two drivers working on alternative days, so that when one’s shift ends, the driver needs to hand the taxi to the other driver). When the regulations are effectively implemented (if not in megacities such as Beijing and Shanghai) with the help of intensive supervision and heavy fines, understanding taxi drivers’ reaction to the regulations is essential for the optimization. We consider:

- **C5:** A taxi driver should not switch from working to resting during peak time in any schedule in \( S \).

Unfortunately, ASM is designed for TEMPs with only constraints C1 and C2, and cannot handle the above constraints.

Heterogenous Taxis and Taxi Drivers

The modeling of taxi drivers’ strategic behaviors is currently established based on the assumption that all taxis and their drivers in the system are homogenous. Although this is generally true in taxi systems of many cities, there are exceptions which require special consideration. For example, as mentioned before, in some cities a taxi can be operated by more than one drivers—usually two, such that one works for the daytime shift, and the other works for the nighttime shift.

In this case, taxis might run for a longer time, and constraints C1 and C2 might actually be violated. Besides, in some cities, taxis might not all belong to the same class. Taxis from different companies, or of different car models might adopt a slightly different fare scheme. Another remarkable case is the dispatching mode, instead of roaming mode, adopted by some taxis. All these cases result in heterogenous strategy spaces or utility functions of the taxi drivers, and call for relaxation of the homogeneity assumption.

Uncertainties in Human Behaviors

Uncertainties have always been issues in modeling human being’s intelligent behaviors. In a taxi system, taxi drivers face uncertainties when implementing their strategies. Since a taxi driver cannot decide the time needed to serve the next customer, when it is approaching the end of her schedule, she will need to decide whether to take another customer or not: taking another customer, the taxi driver may have to work overtime when it takes too long to serve the next customer; otherwise, the taxi driver ends her work ahead of the scheduled time, resulting in less revenue generation. Usually, a taxi driver chooses her action according to the time and the market conditions. The way taxi drivers cope with the uncertainties might also differ among individuals. This might also be attributed to the heterogeneity of taxi drivers.

Impact of App-based Services

The fast development of smart phones in recent years has made available vast new apps and services at hand. Ride sharing apps and customer-to-driver taxi-booking apps, such as Uber and Didi Dache, which connect taxi drivers with customers looking for a ride, are reshaping the traditional taxi market. Notably, these services are more than a simple dispatching systems as it also makes available negotiation between customers and taxi drivers, and provides wider choices to not only the customers but also the taxi drivers. Growing uses of these new services suggest the necessity of considering them in taxi system researches.

Spatial Variances

The existing model only considered the time variances of the taxi system, while spatial variances is also common in taxi systems of modern cities, especially megacities. Density of customer demand, and levels of congestion might all vary over different locations. These variances pose significant impact on the taxi system’s performance, and need to be considered in the future work.

Conclusions

This paper presents a game theoretic approach for the taxi system efficiency optimization problem. The approach includes the following key components: 1) a multi-period taxi system model considering taxi drivers’ strategic behaviors; 2) a bilevel optimization program as the formulation for TEMP; 3) ASM—a compact representation to address the scalability issue of the bilevel program. By modeling taxi drivers’ strategic behavior from a game theoretic perspective, the approach opens a new door to taxi system researches.
References


Smith, J. M., and Cruz, F. 2012. State dependent travel time models, properties, & kernels three-phase traffic theory.


