Taxis Strike Back: A Field Trial of the Driver Guidance System

Industrial Applications Track

Shih-Fen Cheng
School of Information Systems
Singapore Management University
Singapore
sfcheng@smu.edu.sg

Shashi Shekhar Jha
Fujitsu-SMU Urban Computing and Engineering (UNiCEN) Corp Lab.
Singapore Management University
Singapore
shashij@smu.edu.sg

Rishikeshan Rajendram
Fujitsu-SMU Urban Computing and Engineering (UNiCEN) Corp Lab.
Singapore Management University
Singapore
rishikeshanr@smu.edu.sg

ABSTRACT

Traditional taxi fleet operators world-over have been facing intense competitions from various ride-hailing services such as Uber and Grab (specific to the Southeast Asia region). Based on our studies on the taxi industry in Singapore, we see that the emergence of Uber and Grab in the ride-hailing market has greatly impacted the taxi industry: the average daily taxi ridership for the past two years has been falling continuously, by close to 20% in total. In this work, we discuss how efficient real-time data analytics and large-scale multi-agent optimization technology could potentially help taxi drivers compete against more technologically advanced service platforms.

Our technology is based on an earlier theoretical work proven to work in a series of simulation studies. Our major contribution in this paper is the demonstration that the proposed design, when coupled with a real-time data feed of close to 20,000 taxis around Singapore, can indeed help drivers to improve their performances. To provide concrete real-world evidence that such technology can indeed benefit taxi drivers, we have tested the driver guidance system (DGS) operationally since September 2017. With 361 recruited drivers and 5 months of operational data, we have demonstrated that when drivers actively follow our guidance during their roaming (more than 60% of roaming time before acquiring a trip), their expected roaming times can be reduced by 22% when compared to the cases where guidances are not followed. By further breaking down the analysis by time periods, workdays, and areas, we point out the spatial-temporal combinations in which the DGS is most useful.

KEYWORDS

taxi driver guidance; multiagent optimization; transportation; mobility-on-demand

ACM Reference Format:

Shih-Fen Cheng, Shashi Shekhar Jha, and Rishikeshan Rajendram. 2018. Taxis Strike Back: A Field Trial of the Driver Guidance System. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018,* IFAAMAS, 8 pages.

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden.

1 INTRODUCTION

In recent years, we have witnessed a surge in the popularity of mobility-on-demand services, most notably Uber[1]. Despite various controversies surrounding Uber and Uber-like services, most critics agree that Uber-like services have indeed significantly improved the delivery of mobility-on-demand services, which were previously dominated by traditional taxis. Compared to traditional taxi services, Uber-like services bring in innovations in three major areas: 1) the use of smartphones in engaging commuters, 2) the use of data science in enabling better supply-demand matching, and 3) the use of price mechanism in nudging both demand and supply levels. These innovations have made mobility-on-demand services more accessible and affordable in many cities, and taxi industry everywhere is facing enormous pressure in catching up.

Not all aforementioned innovations are new, in fact, the taxi industry has long been using GPS-based dispatch systems for matching drivers and commuters, and along the way generated large amount of data (the most notable such data is the New York taxicab dataset, which is openly available). Based on these datasets, researchers from various fields have come up with a wide spectrum of studies, for example, on behavioral studies [1, 2], big-data analytics [7], and guidance systems [13]. However, very few (if not none) of such research outputs have been adopted by the taxi industry. As the operation environment changes with the emergence of Uber-like services, it is the time for taxi industry to evolve.

By surveying the existing research, we can see that there are many promising ways in which the competitiveness of the taxi industry could be improved. One such idea is to perform big-data analytics and generate personalized guidance for individual drivers. In this paper, we have implemented one such system designed by Jha et al. [6] and tested it in a small-scale field trial. The system contains two major components: 1) the demand/supply prediction engine that relies on a real-time data feed which provides taxi status and locations, and 2) a multi-agent optimizer that generates personalized driving recommendations based on real-time demand/supply predictions. We launched our field trial in September 2017 and have since recruited 361 drivers to our field trial in phases. With five months of operational data, we demonstrate that when drivers are following provided guidances actively (according to our definition, this refers to trips where guidances are followed more than 60% of the whole roaming time), the resulting roaming times are 22% less than the cases where the guidances are not followed this closely

[1]Uber is certainly not the only player in the field; other important players in this area include DiDi in China, Ola in India, Grab in the Southeast Asia, Taxify in EU and Africa, and EasyTaxi in Brazil; just to name a few.

(including cases where the system is not used at all). We further break down the analysis by looking at the temporal and spatial dimensions, and we pinpoint the time periods of a day, workday (or not), and regions in Singapore that the guidance would work the best.

Our initial operational experience shows that there is great potential in pushing driver guidance to the real-world operation.

The rest of the paper is organized as follows: in the next section, we present a brief review of the related literature. The succeeding sections present an overview of the deployed guidance system, and the specifics of our field trial. The analysis of the results gathered from the five months of the field trials is discussed next before concluding the paper.

2 RELATED LITERATURE

The development of taxi driver guidance system has been widely studied in the literature [12–16]. Besides providing guidances to drivers, another stream of research is to provide demand-related information to drivers. For example, in the work by Moreira-Matias et al. [9], taxi demands at various taxi stands are predicted and sent to drivers directly. In Zhang et al. [16], the authors propose a model to learn the driving behaviors and decision-making of taxi drivers using Bayesian learning. By analyzing the historical GPS traces of the movement of various taxi drivers, the authors extract several cues such as origin and destination of trips along with a drive-by information. These sources of information are then modeled to supplement the knowledge of taxi drivers (based on their individual experiences) about the expected demand in various parts of the city.

Yuan et al. [14] propose a recommendation system catering to the taxi drivers as well as passengers. Using a probabilistic model, they derive the temporal distribution of taxis' pick-ups and drop-offs. They use the temporal distributions to predict a preferable parking spot to the taxi drivers (based on their current location) with a high likelihood of passengers arrivals. The authors provide these locations to the taxi drivers as a hotspot based on a score returned by their model. This may lead to imbalances as the global demand and supply is not considered for all the taxis in different states. In Yuan et al. [13], the same set of authors derive the fastest travel routes for a source and destination pair at different times of the day using the movement of taxis. They use the movement patterns of different taxis from the GPS traces to build a time-dependent landmark graph and calculate the practically viable fastest route for a user.

Chiang et al. [4] propose a model to predict booking demands. They divide the space in a set of grids and for each grid they use Gaussian mixture models for estimating the distributions for booking or call based demand at different times of the day. Their model only serves the booking based passengers and does not consider the demand from street-hails and taxi queues. Zhang et al. [15] derive a hotness parameter for the passenger demand at different locations and times using historical traces. This parameter is then used to provide a set of top-k recommendations to the taxi drivers based on their current location and the distance of desired locations with demand. The model does the assignment of taxi drivers at different demand location in a greedy manner which could lead to a global imbalance of demand and supply.

In Qu et al. [11], instead of providing a set of potential locations to look of passengers, the authors construct a complete driving route for each taxi drivers with high net profits. The authors use historical GPS traces to construct the taxi movement trajectories which is then represented as a graph. Each trajectory is then evaluated using a net profit objective function. Finally the taxi drivers are provided with complete driving route in order to maximize their earnings. In a similar approach, Ge et al. [5] provide a set of pick-up locations to the taxi drivers. These locations are derived considering the location of the taxi and the amount of distance to be traveled along with the expected revenue.

As discussed above, various solutions proposed in the literature emphasizes mostly on the prediction of demand hotspots which are then provided to the taxi drivers as recommendations. There are two basic flaws with such systems : 1) Usually the demand prediction models are derived from historical GPS traces. Such models are less reactive to the real-world situations as the latest information is usually not considered. 2) The demand is only considered locally based on the current location of the taxi drivers which leads to greedy matching of demands and supplies. This often leads to imbalances of demands and supplies as the taxis often gets crowded in a few selected areas (such as a city center) while other areas are constantly running low on supplies.

When choosing the framework to power our driver guidance system, we carefully evaluate the candidates so that the implemented system would achieve the following two most critical goals: 1) The chosen system must be reactive to the real-world data feed. This is to avoid the first flaw mentioned above, 2) The chosen system must explicitly consider the interaction among tens of thousands of taxi drivers. This is to ensure that the system can easily scale up regardless of number of drivers who we intend to serve, and 3) Finally, the chosen system must also be scalable computationally. This implies that the guidance engine should be constructed so that the execution time of the engine does not depend on the number of users. In the next section, we will briefly describe a driver guidance system that satisfy all these three important properties.

3 THE DRIVER GUIDANCE SYSTEM (DGS)

Through a real-time data link, we are continuously receiving the locations and states of almost all operating taxis in Singapore from the Land Transport Authority (LTA) of Singapore (at the end of September 2017, there are around 24,000 registered taxis in Singapore, and over 99% of them are included in the dataset). The location is reported as a GPS coordinate while the state indicates whether the taxi is Available, Busy, Hired, or On-call (we will elaborate more on how we use these taxi states in the later section).

Our Driver Guidance System (DGS) is implemented following the design described in Jha et al. [6]. The high-level architecture of the DGS is illustrated in Figure 1, and there are four critical components: 1) data stream handler, 2) demand and supply prediction engine, 3) multi-agent recommendation engine, and 4) mobile App that interacts with the drivers. The high-level design principles and necessary details are briefly described in the following subsections.

Figure 1: The design of the Driver Guidance System (DGS).

For complete description on the technical details, we refer interested readers to Jha et al. [6].

3.1 Data Stream Handler

The real-time data stream contains real-time GPS coordinates and states of all currently active taxis in Singapore[2]. As GPS coordinates and status updates both can experience errors due to either hardware or communication issues, we will need to identify and correct (if possible) these errors before feeding the received data to other components. The basic data cleaning step involves removing the obvious GPS errors such as location discontinuity, out-of-bound errors, or null signals. After basic cleaning, the next important task to be performed in this component is the *map matching*, where the GPS coordinate is mapped to an actual road link. This step allows us to sense the activity of taxi movement along the actual road network. All taxis' locations are continuously mapped to physical roads using a Hidden-Markov-Model-based map-matching algorithm [10]. This algorithm also provides necessary corrections to the location sensing where necessary. Our map-matching implementation works in a rolling horizon manner and generates trajectories for all taxis independently.

3.2 Demand and Supply Prediction Engine

The precise knowledge of demands and supplies forms the foundation for the DGS platform. While the real-time supply is readily available from the data stream, the prediction of demand and supply in near future is challenging.

The major innovation in the design of the demand prediction engine is to treat each free-cruising taxi as a *demand probe*. Empirically, we observe that the chance of us seeing demand on a particular link is inversely correlated with the amount of time passed since last visit by a free taxi. In other words, it invalidates the memoryless property of the exponential distribution and thus the demand will not follow the Poisson arrival process, which is commonly assumed in the literature. Based on this insight, we develop independent prediction models for all tuples of (*road link*, *day of week*, *time of day*). These models essentially return the likelihood of a cruising taxi getting a trip (could be from either street hailing or booking).

Formally speaking, our demand prediction engine employs a multilevel logistic regression model to estimate the likelihood of finding a passenger on a street at different times. Let T be the state of a taxi which is 1 if the taxi is hired (occupied), and 0 otherwise. The likelihood of finding a passenger on a street s is given by:

$$P(T = 1|\delta_s) = \text{logit}^{-1}(\alpha_{s,t,d} + \beta_{s,t,d}\ \delta_s), \tag{1}$$

where s is the street, t is the time-interval (30 minutes) of a day, d is the day of the week, and $\alpha_{s,t,d}$ and $\beta_{s,t,d}$ are regression coefficients.

[2] All except 100 taxis that are managed by HDT, the newest fleet that is exempted from the data requirement.

The independent variable is δ_s, which represents the time elapsed since the last arrival of an empty taxi on street s.

By monitoring taxi movements in real-time, we can utilize the above regression models and predict the likelihoods of taxi demands along all links. However, these models do not provide us with predictions on *demand counts* within a given time interval, which would be needed when we engage the optimization model. To predict the demand counts, we simulate the arrivals of empty taxis by utilizing historical data, and calculate the expected demands that would be generated.

3.3 Multi-agent Recommendation Engine

The recommendation engine generates personalized recommendations for all taxi drivers based on their locations, with the objective of balancing overall demands and supplies across the city. This problem can be viewed as a specialized spatio-temporal matching problem, where taxis (agents) are instructed to move around and match with (stationary) passengers. We solve this multiagent matching problem using a centralized multi-period stochastic optimization model proposed by Lowalekar et al. [8]. The decision space for the recommendation is discretized into 1km-by-1km grid cells. These grid cells are also used for aggregating demand and supply predictions. The objective function is defined as:

$$\begin{aligned} \max\Big(&- \sum_{i\in\mathcal{G}}\sum_{j\in\mathcal{G}} Cost^1_{ij}\cdot u^1_{ij}\Big) \\ &+ \frac{1}{|D|}\sum_{k\le|D|}\sum_{t=1}^{T}\sum_{i\in\mathcal{G}}\Big(\sum_{d\in D^k_t}\mathcal{R}^t_{id}\cdot x^{t,k}_{id}\Big) \\ &- \frac{1}{|D|}\sum_{k\le|D|}\sum_{t=2}^{T}\sum_{i\in\mathcal{G}}\Big(\sum_{j\in\mathcal{G}} Cost^t_{ij}\cdot u^{t,k}_{ij}\Big), \end{aligned} \tag{2}$$

where $\mathcal{G}$ is the set of grid cells, D is the number of demand samples considered for the optimization, T is the time horizon, $Cost_{ij}$ is the cost of movement from grid i to grid j, R_{id} is the net revenue for serving demand d in grid i, $x^{t,k}_{id}$ is a decision variable that assigns agents in grid i to demand d at time t within sample k. After the assignment, the number of unassigned taxis moving from grid i to grid j is denoted by u_{ij}.

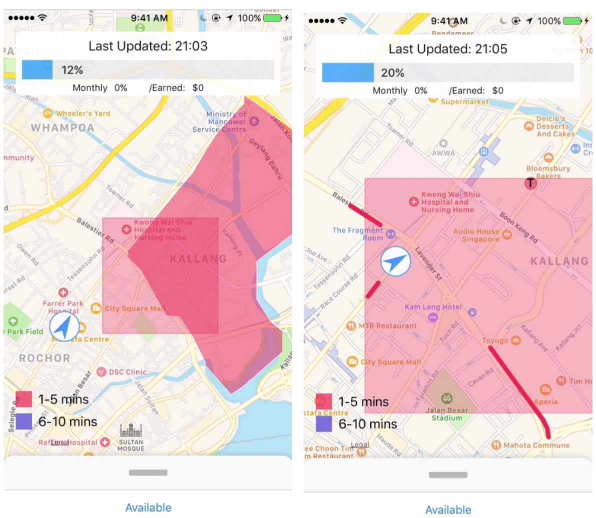
The objective function used in the recommendation engine is essentially a global function reflective of the *total* revenues earned by all drivers. The goal is to calculate an optimal matching for all drivers such that demand fulfillments and movement costs are balanced over the planning horizon. A practical concern with DGS is to explicitly consider its adoption by taxi drivers. Although the DGS is designed to serve all taxi drivers, a significant number of taxi drivers will not be using the system for various reasons. We handle such taxi drivers by simulating their behaviors based on historical cruising patterns. When performing demand assignments in our optimization model, we explicitly consider the part of demand fulfilled by drivers not using DGS.

To generate most up-to-date recommendations, the engine is loaded with latest information on demands and supplies and executed every minute. The planning horizon is 30 minutes.

3.4 Mobile Application

In order to deliver personalized recommendations to drivers, we have developed mobile phone Apps for both iOS and Android platforms. Versions for both platforms look identical and display recommendations as an overlay over the map of Singapore. The App is designed to work without any need for user interaction: the user just leaves the App active, and recommendations are streamed to the App based on user's current location. The App is designed to automatically adjust its zoom level and display different details in the following two modes:

- **Region mode** (see Figure 2a): The region-level recommendations are generated by the engine described above. The highlighted region provides the general direction in which the driver should move into.
- **Street mode** (see Figure 2b): The streets (or taxi stands) are highlighted probabilistically based on the likelihoods of demand generation derived from Equation (1).



(a) Region-level (b) Street-level

Figure 2: Mobile App UI displaying recommendations at different levels.

3.5 Evaluation

To evaluate the effectiveness of the DGS, we simulate the operation of the DGS using a realistic agent-based taxi fleet simulation platform called TaxiSim [3]. TaxiSim is calibrated using the actual taxi driver's movement traces we obtained from the LTA. To prepare the simulation environment, we populate the simulation with 24,000 taxis, and randomly generate passenger demands using the demand profiles chosen from a list of representative days (including both typical workdays and non-workdays). To evaluate the effectiveness and the scalability of the DGS, we execute the simulations with different percentages of taxi drivers using the DGS (from 5% to 100%).

The simulation is designed to provide streaming data feed exactly like the real API, and the DGS is attached to this emulated stream data and generates recommendations as if it is used in the real-world environment.

The average number of daily trips per driver is selected to be the key performance metric, and we vary the market share of guided taxis (DGS taxis) from 5% to 100%. The resulting performance is illustrated in Figure 3, and we can see that although the performance of DGS drivers slowly deteriorates as the market share increases, DGS drivers always outperform non-DGS drivers (even at 100% DGS market share).

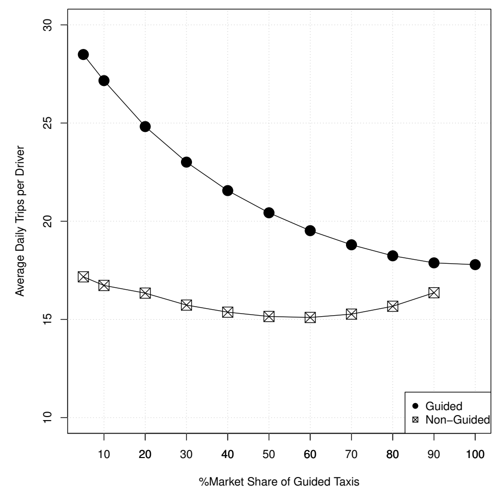


Figure 3: The performance of guided versus non-guided taxi drivers over the market share of the DGS users.

These simulation studies confirm that the DGS platform is scalable and ready for real-world deployment and testing.

4 THE DGS FIELD TRIAL

4.1 The Taxi Industry in Singapore

The taxi market in Singapore is made up of one big player owning over 50% of the market share and the rest of the market is split among 5 smaller players. Only Singapore citizens above 30 years old are eligible as taxi drivers. After obtaining a Taxi Vocational Driving License, the prospective driver has to rent a taxi from one of the operators (individual taxi ownership is not allowed). Once the taxi rental is paid, the driver can keep all the earned fares (while paying their own fuels).

The taxi market was growing steadily until 2015 (the industry as a whole had around 28,000 taxis at the end of 2015), after which it faced intensive competitions from a number of ride-hailing startups, most notably Uber and Grab (a major ride-hailing service provider in the Southeast Asia). The fleet size of the private-hire cars is now estimated to be 50% more than the taxi population, and due to this competition, the taxi population is shrinking and now stands at only 24,000 (at the end of September 2017).

Note that although we focus only on taxi drivers in our study, a significant number of taxi drivers also receive jobs from Uber and Grab. In particular, Grab has inked formal agreements with all taxi operators in Singapore (except Comfort and CityCab) to encourage the taxi drivers to use GrabTaxi as the preferred third-party taxi booking service. The integration went further in March 2017, when some taxi operators and Grab collaborated and were allowed (by LTA) to offer dynamic taxi fares booked via mobile application. Grab launched a new service called JustGrab[3], which essentially allows riders to book from both the participating chauffeured private hire cars and taxis. Riders pay the same fare regardless of the types of vehicle they are assigned to.

4.2 The Field Trial Setup

As described above, the taxi industry is facing intensive competition from the 40,000-strong chauffeured private hire cars operated under Uber and Grab. We believe that the implementation of the DGS is an important first step in addressing such challenge. To confirm the effectiveness of the DGS in the real-world environment, we have begun a series of field trials since the beginning of September, 2017[4]. Each batch of the field trial lasts one month, and drivers are invited to participate via open recruitment. To encourage the participation and active usage, we provide the following two types of incentives:

- **For participation**: the drivers are expected to install the App and accumulate a minimum of 4 hours of usage time during the one-month trial period. A S$100 incentive will be awarded for drivers meeting the 4-hour usage requirement.
- **For compliance**: we track how drivers are following the guidances shown. For drivers whose average compliance time per day meets the requirement in Table 1, additional incentives (on top of S$100) will be awarded (we set 2 hours of daily compliance to be the daily goal; therefore drivers need to achieve at least 50% of this goal to begin receiving the compliance incentives). To avoid binge accumulation of hours, drivers can accumulate at most 4 hours per day.

Table 1: Compliance-based incentives.

%	Avg. Compliance (min)	Var. Incentives
50%	60 – 71	S$50
60%	72 – 83	S$60
70%	84 – 95	S$70
80%	96 – 107	S$80
90%	≥ 108	S$100

To assist trial drivers in tracking their daily and monthly compliance progress, we display this information in the DGS App directly, as seen in Figure 4. The percentage bar indicates daily compliance as percentage of the 2-hour goal, which will be reset every day. The monthly compliance (as percentage of the 2-hour daily average goal) and the earned incentive are displayed below the daily compliance bar.

[3]https://www.grab.com/sg/justgrab/
[4]Approved under IRB-17-113-A099-M1(218).

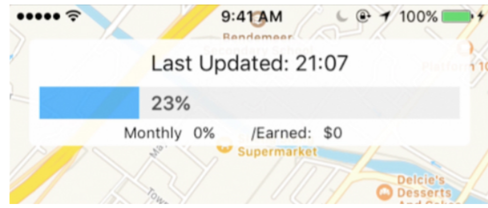


Figure 4: Compliance progress shown at the top of the App.

4.3 Collected Information

When drivers register, they are asked to provide their basic demographic information (age group), number of years driving professionally, whether they accept booking jobs from Uber or Grab, and if they do, their estimates on the number of Uber/Grab jobs accepted per day.

During the field trial, we have collected the following two types of information:

- **DGS App-related**: When the DGS App is in use, our backend server receives the status and the location of the device at one-minute interval. Based on these continuous updates, we can infer when the DGS App is used, and when the driver follows the DGS guidance. More specifically, two critical series are derived: 1) the usage episodes, and 2) the compliance episodes. For both series, they are recorded as a series of the following tuples: <taxi-id, start-time, end-time>. The usage episodes are straightforward to determine, as it simply means that the DGS App is currently being displayed. However, the compliance episodes are harder to determine. To be as objective as we can, we assume the driver is following the DGS guidance if the driver is: 1) currently inside the recommended region, or 2) moving *closer* to the recommended region.
- **Taxi trip-related**: Separately, we have obtained from the Land Transport Authority (LTA) of Singapore the comprehensive state updates of all participating taxis. This dataset allows us to infer the starting and ending time and location of all taxi trips. We can also infer how the trip is generated (i.e., pick up from the street or via booking)

5 FIELD TRIAL ANALYSIS

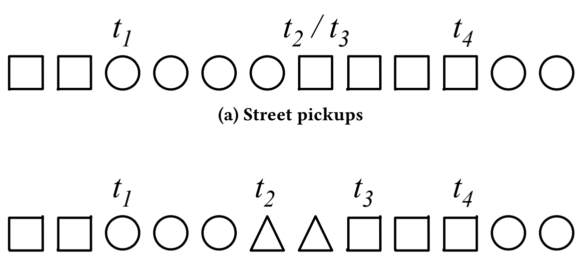
As noted earlier, our field trials have begun from September 2017, and have been running continuously since then. So far we have collected three consecutive batches of trial results, and they look promising. Before going into the performance analysis, we will first introduce our performance metric, and how we calculate it.

5.1 Performance Metric and Data Processing

From the driver's perspective, the most important measure of the performance would be his/her productivity, i.e., the number of trips per hour of working time. This is equivalent to the *average roaming time* required to obtain a trip. Given long enough observations, we expect that the average roaming time per trip should reflect the actual driving skill. In our analysis, we try to determine whether the use of the DGS technology could significantly affect the roaming time in acquiring the next passenger.

To measure this, we will need to provide two measurements for each trip: 1) the *estimated* roaming time it takes for the driver to acquire this trip, and 2) whether the acquisition of this trip can be attributed to the following of DGS guidance. To estimate the roaming time for each trip, we should observe the state transitions based on the data stream provided by the LTA. For most street pickups, we expect that the state transitions should be from Available to Hired. For booking trips via official channel, the state transitions are usually from Available to On-Call (the official state of responding to booking request), to Hired. For booking trips via Apps such as Uber or Grab, the state transitions are usually from Available to Busy (indicates that the taxi is now committed and unwilling to serve street pickups), to Hired.

Examples on how roaming time can be calculated for both street pickups and booking (either official or App-based) are illustrated in Figure 5. In Figure 5, we use circle, triangle, and square to represent the states of Available, On-Call (or Busy, if we are trying to detect App-based booking), and Hired respectively. According to the rules specified earlier, for each detected trip, let t_1 be the first Available state after the previous Hired state; this would be the beginning of the roaming. Let t_2 be the first state that is not Available after t_1; this will be the end of the roaming. Let t_3 be the first Hired state after t_1, and t_4 be the last Hired state after t_1. Once these four major state transition time points are identified, we can then calculate the following important durations: 1) the roaming time: $t_2 - t_1$, 2) the response time (to taxi booking): $t_3 - t_2$, and 3) the service time: $t_4 - t_3$.



(a) Street pickups

(b) Booking jobs

Figure 5: Examples on how roaming times are derived.

To determine whether the acquisition of a trip can be attributed to the following of the DGS guidances, we look at the fraction of the roaming time during which the driver is following the DGS guidances. We assume that if the fraction of the compliant period during the roaming time is higher than a pre-determined threshold, the resulting trip will then be labeled as resulting from following DGS (since the driver is sufficiently affected by the DGS). The illustration of how the DGS compliance ratio is calculated can be seen in Figure 6. In this example, we can see that there are two DGS compliant episodes plotted as black bars. The fragments of the roaming period that overlap with the DGS compliant episodes are shaded in gray. For this example, we can then conclude that the DGS compliance ratio for this trip is 2/3.

Figure 6: Determining DGS compliance ratio.

5.2 Empirical Results and Analyses

Table 2: The overall statistics of non-DGS vs DGS trips (for all drivers).

Time Period		Non-DGS	DGS
06-10	Avg. Roaming (min)	8.32	7.53
	Trip Count	97,336	1,769
10-17	Avg. Roaming (min)	7.72	6.34
	Trip Count	195,768	3,107
17-24	Avg. Roaming (min)	9.23	7.89
	Trip Count	188,000	1,755
24-06	Avg. Roaming (min)	15.61	10.67
	Trip Count	70,143	527
Overall	Avg. Roaming (min)	9.34	7.33
	Trip Count	551,247	7,158

Table 3: The overall statistics of non-DGS vs DGS trips (for drivers with at least 20-compliant DGS trips).

Time Period		Non-DGS	DGS
06-10	Avg. Roaming (min)	8.32	7.58
	Trip Count	11,787	1,685
10-17	Avg. Roaming (min)	7.90	6.30
	Trip Count	24,405	2,917
17-24	Avg. Roaming (min)	9.06	7.94
	Trip Count	24,511	1,663
24-06	Avg. Roaming (min)	17.06	10.68
	Trip Count	9,254	493
Overall	Avg. Roaming (min)	9.59	7.34
	Trip Count	69,957	6,758

The analyses presented in this section are based on our field trials from September 11, 2017 to January 31, 2018, participated by 361 drivers. The recruitment of drivers occurred in three batches: 26 drivers were recruited on September 11, 2017, another 31 drivers were recruited on October 16, 2017 while 304 drivers were recruited over two days on November 18 and 19, 2017. We set the threshold for DGS compliance ratio to be 60%. In other words, a trip will be labeled as being DGS compliant if the DGS compliance ratio is at least 60% or above. The grand overview of non-DGS trips versus DGS trips can be seen in Table 2. Note that as queueing time at airport is significantly longer and may skew our roaming time analysis, we have excluded airport trips from our analysis.

As can be seen from Table 2, we have divided all days into four time periods: 06-10 and 17-24 are morning and evening peak periods, while 10-17 and 24-06 are day-time and night-time non-peak periods. Overall, DGS trips experience 21.5% less roaming time than non-DGS trips. In fact, across all time periods, DGS trips experience shorter roaming time than non-DGS trips (all differences are significant with p-values close to 0).

We also conduct similar analysis for active users of the DGS (defined as users with at least 20 DGS-compliant trips, which include 44 drivers). The results are shown in Table 3. For this group of drivers, the saving in roaming time is slightly higher at 23.5%, yet their usage rates are much higher: the percentage of DGS-compliant trips over all trips is around 10%.

To further understand the impact of DGS compliance ratio on driver's roaming time performance, we also plot all roaming times against DGS compliance ratios for trips whose DGS compliance ratios are at least 0.1. By fitting these data points to a linear model, we have the following linear regression model:

$$\text{Roaming Time (min)} = -19.03 \text{ Compliance Ratio} + 24.73 + \epsilon, \quad (3)$$

where there are 11,800 observations, $R^2 = 0.1857$, and all parameters are significant with $p < 0.001$. Intuitively speaking, this implies that for every 10% increase in the DGS compliance ratio, the empty roaming time is expected to fall by 1.903 minutes (or close to 114 seconds).

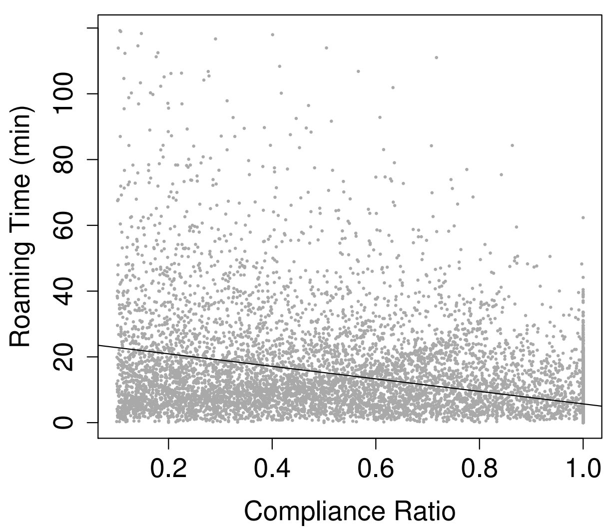


Figure 7: Roaming time against DGS compliance ratio for all trips.

5.3 Spatial-Temporal Breakdowns of DGS Performance

To investigate the impact of external factors on DGS performances, we extend the above analysis by breaking down the performance comparison spatially and temporally. More specifically, for temporal dimension, we compare performances on workdays versus non-workdays. For spatial dimension, we compare the performances within the Central Business District (CBD) and non-CBD. Note that as before, we have excluded the airport trips from our consideration. The location of CBD in Singapore is illustrated in Figure 8.

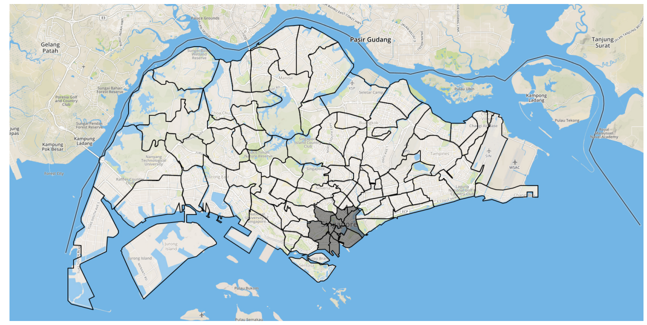


Figure 8: Singapore's Central Business District (CBD) is colored as the shaded region in the center.

The breakdowns of DGS performances on workday vs. non-workday and CBD vs. non-CBD can be found in Table 4. Our focus is on finding particular time periods and areas where we observe relatively good performance from the DGS-compliant trips. In summary, DGS guidance appears to work best for the following four major blocks of spatial-temporal combinations:

- Midnight hours (24-06) on all days, for all areas. The advantages of DGS over non-DGS trips range from 24% to 36%.
- Daytime non-peak hours (10-17) on all days for non-CBD areas and non-workday for CBD area. The advantages of DGS over non-DGS trips range from 17% to 19%.
- Evening peak hours (17-24) on workdays for all areas and on non-workday for non-CBD area. The advantages of DGS over non-DGS trips range from 15% to 17%.
- Morning rush hours (06-10) on non-workdays, for non-CBD areas. The advantage of DGS over non-DGS trips is 12%.

By examining these blocks, we can see that for most cases where DGS is most helpful, demands are sporadic and less predictable (e.g., non-peak hours and non-CBD areas).

6 CONCLUSIONS

In this paper, we have implemented and field tested a first-of-its-kind taxi driver guidance system (DGS) in Singapore. By incorporating live data feed of close to 20,000 taxis, we show that high-quality decision supports can be generated for taxi drivers by integrating a dynamic demand prediction engine and a multi-agent, multi-period optimization engine. Since September 2017, we have begun recruiting test drivers to test DGS, and our initial findings from the five-month trials (September 2017 to January 2018) are encouraging.

We measure taxi driver's performance by looking at the roaming time before each acquired trip, and we also identify if each trip is affected by the DGS guidance. For the latter identification, we measure whether the driver follows our DGS guidances sufficiently long (more than 60% of the roaming time for the trip in interest). For DGS-compliant trips, we show that the corresponding roaming

Table 4: Performance breakdown by workdays (WD) and locations (CBD and non-CBD).

		WD				Non-WD			
		Non-CBD		CBD		Non-CBD		CBD	
Time Period		Non-DGS	DGS	Non-DGS	DGS	Non-DGS	DGS	Non-DGS	DGS
06-10	Avg. Roaming Time (min)	8.54	7.83	8.00	7.61	7.98	7.01	7.75	4.16
	Trips	64,144	1,180	10,910	204	19,264	324	2,745	61
10-17	Avg. Roaming Time (min)	8.78	7.11	6.21	5.51	6.78	5.52	5.58	4.88
	Trips	107,609	1,665	32,340	481	44,565	795	10,768	166
17-24	Avg. Roaming Time (min)	9.33	7.79	10.50	8.93	7.86	6.50	8.92	8.68
	Trips	94,377	779	40,520	447	40,004	370	12,796	159
24-06	Avg. Roaming Time (min)	16.83	11.52	18.58	11.87	13.04	9.90	12.97	9.22
	Trips	29,413	231	12,849	52	17,317	144	10,466	100
Overall	Avg. Roaming Time (min)	**9.70**	**7.73**	**9.86**	**7.44**	**8.23**	**6.42**	**9.09**	**6.93**
	Trips	295,513	3,855	96,619	1,184	121,150	1,633	36,775	486

time can be reduced by 22% (when compared against non-DGS trips). After looking at the performance breakdown by time periods, workdays, and area of service, we further pinpoint the combinations of times and areas that the DGS is most useful for our test drivers.

To the best of our knowledge, this is the first operating guidance system for taxi drivers that combines real-time data analytics and high-performance multi-agent optimization. The field test of the guidance system is still ongoing, and we expect sign ups to reach 1,000 by the end of 2018. Further in the future, we aim to roll out the system for thousands of drivers in Singapore.

ACKNOWLEDGMENTS

This research is funded by the National Research Foundation Singapore under its Corp Lab @ University scheme and Fujitsu Limited as part of the A*STAR-Fujitsu-SMU Urban Computing and Engineering Centre of Excellence.

REFERENCES

[1] Colin F. Camerer, Linda Babcock, George Loewenstein, and Richard Thaler. 1997. Labor supply of New York City cabdrivers: One day at a time. *The Quarterly Journal of Economics* 112, 2 (1997), 407–441.
[2] Pablo Samuel Castro, Daqing Zhang, Chao Chen, Shijian Li, and Gang Pan. 2013. From taxi GPS traces to social and community dynamics: A survey. *Comput. Surveys* 46, 2 (2013).
[3] Shih-Fen Cheng and Thi Duong Nguyen. 2011. TaxiSim: A multiagent simulation platform for evaluating taxi fleet operations. In *2011 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology.* 14–21.
[4] Meng-Fen Chiang, Tuan-Anh Hoang, and Ee-Peng Lim. 2015. Where are the passengers?: A grid-based Gaussian mixture model for taxi bookings. In *Twenty-Third SIGSPATIAL International Conference on Advances in Geographic Information Systems.* 32.
[5] Yong Ge, Hui Xiong, Alexander Tuzhilin, Keli Xiao, Marco Gruteser, and Michael Pazzani. 2010. An energy-efficient mobile recommender system. In *Sixteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* 899–908.
[6] Shashi Shekhar Jha, Shih-Fen Cheng, Rishikeshan Rajendram, Meghna Lowalekar, Nicholas Wong, Trong Khiem Tran, Pradeep Varakantham, Truong Trong Nghia, and Firmansyah Bin Abd Rahman. 2018. Upping the Game of Taxi Driving in the Age of Uber. Thirtieth Annual Conference on Innovative Applications of Artificial Intelligence. (2018).
[7] Liang Liu, Clio Andris, Assaf Biderman, and Carlo Ratti. 2009. Uncovering taxi driver's mobility intelligence through his trace. *IEEE Pervasive Computing* 160 (2009), 1–17.
[8] Meghna Lowalekar, Pradeep Varakantham, and Patrick Jaillet. 2016. Online Spatio-temporal Matching in Stochastic and Dynamic Domains. In *Thirtieth AAAI Conference on Artificial Intelligence.* 3271–3277.
[9] Luis Moreira-Matias, Joao Gama, Michel Ferreira, Joao Mendes-Moreira, and Luis Damas. 2013. Predicting taxi–passenger demand using streaming data. *IEEE Transactions on Intelligent Transportation Systems* 14, 3 (2013), 1393–1402.
[10] Paul Newson and John Krumm. 2009. Hidden Markov map matching through noise and sparseness. In *Seventeenth ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems.* 336–343.
[11] Meng Qu, Hengshu Zhu, Junming Liu, Guannan Liu, and Hui Xiong. 2014. A cost-effective recommender system for taxi drivers. In *Twentieth ACM SIGKDD International conference on Knowledge discovery and Data mining.* 45–54.
[12] Dongxu Shao, Wei Wu, Shili Xiang, and Yu Lu. 2015. Estimating taxi demand-supply level using taxi trajectory data stream. In *IEEE International Conference on Data Mining Workshop.* 407–413.
[13] Jing Yuan, Yu Zheng, Chengyang Zhang, Wenlei Xie, Xing Xie, Guangzhong Sun, and Yan Huang. 2010. T-Drive: driving directions based on taxi trajectories. In *Eighteenth SIGSPATIAL International Conference on Advances in Geographic Information Systems.* 99–108.
[14] Nicholas Jing Yuan, Yu Zheng, Liuhang Zhang, and Xing Xie. 2013. T-finder: A recommender system for finding passengers and vacant taxis. *IEEE Transactions on Knowledge and Data Engineering* 25, 10 (2013), 2390–2403.
[15] Kai Zhang, Zhiyong Feng, Shizhan Chen, Keman Huang, and Guiling Wang. 2016. A Framework for Passengers Demand Prediction and Recommendation. In *IEEE International Conference on Services Computing.* 340–347.
[16] Yingjie Zhang, Beibei Li, Ramayya Krishnan, and Siyuan Liu. 2015. Learning from the Offline Trace: A Case Study of the Taxi Industry. In *Thirty-Sixth International Conference on Information Systems.*