



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

공학박사 학위논문

Matching and Pricing Algorithm in a Ride-hailing Service Based on Reinforcement Learning to Reduce Pickup Waiting Time

픽업 대기 시간의 감소를 위한 강화 학습 기반의
승차 공유 서비스의 매칭 및 요금 책정 알고리즘

2023년 8월

서울대학교 대학원
공과대학 건설환경공학부
정 다 운

Matching and Pricing Algorithm in a Ride-hailing Service Based on Reinforcement Learning to Reduce Pickup Waiting Time

지도교수 김 동 규

이 논문을 공학박사 학위논문으로 제출함
2023년 8월

서울대학교 대학원
공과대학 건설환경공학부
정 다 운

정다운의 박사 학위논문을 인준함
2023년 8월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

Abstract

With the popularization of smartphones, ride-hailing services mediate hailing requests from many users to drivers. They are expected to ease urban problems such as traffic congestion by replacing the demand for private cars. However, contrary to their intended purpose, there exists criticism that ride-hailing services exacerbate negative externalities such as traffic congestion and air pollution. Therefore, it is necessary to alleviate those negative externalities in ride-hailing services. We focus on reducing pickup waiting times, as this also helps decrease negative externalities by reducing the distance traveled by vehicles. This study develops a matching and pricing algorithm to reduce pickup waiting time based on reinforcement learning methods, namely contextual bandits and temporal difference learning. The algorithm iteratively assigns surge multipliers to individual requests and matches the requests to drivers. We apply the proposed algorithm to ride-hailing data in Singapore while investigating the effects of changes in price sensitivity and value term coefficient. The results show similar patterns to the historical matching rate during both peak hours and off-peak hours, demonstrating its applicability in large cities as big as Singapore. Simulations during the morning peak hour show that our proposed model reduces 20 kilograms of carbon dioxide emissions and decreases pickup waiting time by 15% with only a

2% reduction in revenue. By addressing matching, pricing, and reducing pickup waiting time together, this study estimates how much pickup waiting time can be potentially saved and contributes to improving ride-hailing services both in terms of user benefits and reducing externalities.

Keywords : Ride-hailing service, Dynamic pricing, Negative externalities, Pickup waiting time, Reinforcement learning, Simulation

Student Number : 2016-21269

Contents

Chapter 1. Introduction	1
1.1. Background	1
1.2. Research Purpose	7
Chapter 2. Literature Review	9
2.1. Concepts	9
2.2. Dynamic Pricing	12
2.3. Negative Externalities	14
2.4. Implications	17
Chapter 3. Methodology	19
3.1. Contextual Bandit	19
3.2. Temporal Difference Learning	23
3.3. Model Formulation	26
Chapter 4. Data Description	33
4.1. Data Tables	33
4.2. Descriptive statistics	38
4.3. Spatial Distribution	39
Chapter 5. Results	43
5.1. Base Model	43
5.2. Pickup Waiting Time Reducing Model	52
5.3. Effect of Changes in Price Sensitivity	56

5.4. Effect of Changes in Value Term Coefficient	61
5.5. Applicability of the Model	69
Chapter 6. Conclusion	73
6.1. Summary	73
6.2. Limitations and Future Research	74
References	75
국문초록	82

List of Tables

Table 1. Revenue and forecasts of ride-hailing and taxi in Singapore	2
Table 2. Comparison of matching time and pickup waiting time in historical data	6
Table 3. Elements of reward function	28
Table 4. Number of requests by type after preprocessing	35
Table 5. Columns of ride data table	36
Table 6. Columns of ping data table	37
Table 7. Descriptive statistics for Monday, November 16, 2020	38
Table 8. Comparing two terms of trial simulation	44
Table 9. Matching rate and error rate by travel distance	47
Table 10. Matching rate and error rate by time	48
Table 11. Distribution of surge multipliers (Base model)	49
Table 12. Value term and matching rate in base model	51
Table 13. Comparison of base model and pickup cost reducing model	54
Table 14. Comparison to TADA's pricing	55
Table 15. Matching rate according to travel distance with changing price sensitivity	57
Table 16. Matching rate according to time with changing	

price sensitivity	58
Table 17. Revenue and reward with changing price sensitivity	59
Table 18. Convergence of algorithm with changing price sensitivity	60
Table 19. Matching rate according to travel distance with changing value term coefficient	62
Table 20. Matching rate according to travel distance with changing value term coefficient	64
Table 21. Matching rate according to time with changing value term coefficient	65
Table 22. Revenue and reward with changing price sensitivity	66
Table 23. Convergence of algorithm with changing value term coefficient	68
Table 24. Matching rate by travel distance (off-peak)	70
Table 25. Comparison of base model and pickup cost reducing model (Off-peak hours)	71

List of Figures

Figure 1. Ride-hailing companies	2
Figure 2. Time intervals that constitute a single request ...	5
Figure 3. Structure of reinforcement learning	19
Figure 4. Difference between a full reinforcement learning problem and a contextual bandit	20
Figure 5. Update of value function	24
Figure 6. Bootstrapping structure of algorithm	30
Figure 7. An example illustrating that updating the algorithm results in new situations	31
Figure 8. Hourly hailing requests of November 16, 2020	33
Figure 9. Spatial scope of study	34
Figure 10. Spatial distribution of origins (left) and destination (right)	40
Figure 11. Spatial Distribution of Origins of Matched (Left) and Unmatched (Right) Requests	41
Figure 12. Spatial Distribution of Destinations of Matched (Left) and Unmatched (Right) Requests	41
Figure 13. The proportion of each surge multiplier converges as iteration progresses	45
Figure 14. The percentage of requests with altered surge multipliers in the subsequent iteration	46

Figure 15. Matching rate by travel distance	47
Figure 16. Matching rate by time	48
Figure 17. Spatial distribution of origins and requests with surge multipliers less than 1	50
Figure 18. Matching rate according to value term	51
Figure 19. Comparison of surge multiplier distribution with base model (Left)	53
Figure 20. Comparison of surge multiplier distribution with base model (Left)	63
Figure 21. Temporal matching rate of off-peak	70
Figure 22. Surge multiplier gap between base model and pickup cost reducing model	72

Chapter 1. Introduction

1.1. Background

1.1.1. Social Background

Ride-hailing services mediate hailing requests from many users (riders) to drivers. Unlike public transportation, users do not have to wait for a cramped vehicle that travels along a fixed route according to a predetermined timetable. Riders just set the origin and destination on a smartphone application, and the service will soon match them with the driver. Such personalized services were difficult to emerge before the development of information and communication technology. Ride-hailing services are expanding with the generalization of smartphones and the activation of the sharing economy. As each vehicle is shared by multiple people, ride-hailing services are expected to ease urban problems such as traffic congestion by replacing the demand for private cars. Some services are available in multiple countries and each country has a popular ride-hailing service. Uber of the United States, Grab and Tada of Singapore, and Cabify of Latin America, KakaoT of South Korea, and DiDi of China are well-known services, which are listed in Figure 1.

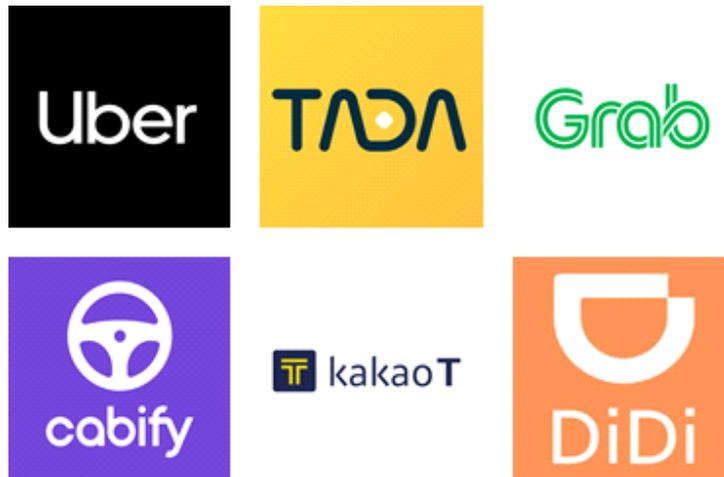


Figure 1. Ride-hailing companies

In Singapore, which is the spatial scope of this study, revenues of ride-hailing and taxi have been increasing (except for a decrease in 2020 due to COVID-19). As shown in Table 1, it is expected to have an annual growth rate of 4% from 2023 to 2027.

Table 1. Revenue and forecasts of ride-hailing and taxi in Singapore

	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027
Total	1.05	1.11	1.14	0.88	1.21	1.28	1.35	1.42	1.48	1.54	1.59
in billion USD, Most recent update: Dec 2022											
Notes: Data shown is using current exchange rates. Data shown does reflect market impacts of Russia-Ukraine war.											
Source: Statista Market Insights											

Despite the aforementioned advantages, ride-hailing services carry inherent issues that originate from their nature as platforms

that mediate between users and drivers. Ride-hailing is a two-sided market where demand (riders) and supply (drivers) match (Wang and Yang, 2019). Demand and supply of a ride-hailing service vary over time and space, and their spatiotemporal heterogeneity often leads to imbalances, resulting in a mismatch between demand and supply.

Contrary to its intended purpose, there exists criticism that ride-hailing service exacerbates negative externalities such as traffic congestion and air pollution. Despite the potential of ride-hailing services to reduce car ownership, ride-hailing vehicles continue to move on roads even when they are not carrying users. This incurs traffic and travel distance, which leads to congestion and air pollution. Ride-hailing companies should address these issues and respond to such concerns, as this is an important issue that can affect the quality of life of the public.

In Singapore, carbon tax is going to increase. The initial carbon tax rate was set at 5 SGD per ton of carbon dioxide equivalent from 2019 to 2023 (Source: National Environment Agency of Singapore). However, it will be raised to 50 SGD to 80 SGD per ton by 2030, which is Asia's highest level. The increasing carbon tax imposes a burden on ride-hailing platforms. Along with each platform's carbon neutral plan, it is necessary to analyze how much carbon footprint can be reduced. Therefore, it is necessary to alleviate those negative externalities as well as dealing with the mismatching problem.

From the social background, the problem to be resolved is set. As driver supply cannot be increased in the short term to address the imbalance problem, price adjustment is required. Simultaneously, it is necessary to find ways to alleviate negative externalities.

1.1.2. Academic Background

Dynamic pricing is an incentive for enhancing efficiency of resource since a price surge makes more drivers work and wait for potential requests. In other words, an increase in price leads to a decrease in demand. From the perspective of rider, a price surge basically acts as a burden on riders and lowers the demand for service. Therefore, changes in prices affect both drivers and riders. However, previous studies focused on increasing revenue of ride-hailing services through dynamic pricing while often overlooking riders' inconvenience.

Unlike street-hailing (traditional taxi), app-based ride-hailing involves waiting time after making a request, making unoccupied time (the time the vehicle moves without a rider: from t_5 to next t_4 in Figure 2) of a vehicle longer. Average unoccupied time between requests in Manhattan, New York is 8min for Yellow cab and 11min for Uber and Lyft (Feng et al., 2021). Figure 2 shows time intervals that constitute a single request. Duration of the request is expressed as $t_5 - t_1$. Subtracting travel time ($t_5 - t_4$) from the

duration is total waiting time of a rider. The total waiting time of a rider ($t_4 - t_1$) consists of matching time ($t_3 - t_1$) and pickup waiting time ($t_4 - t_3$).

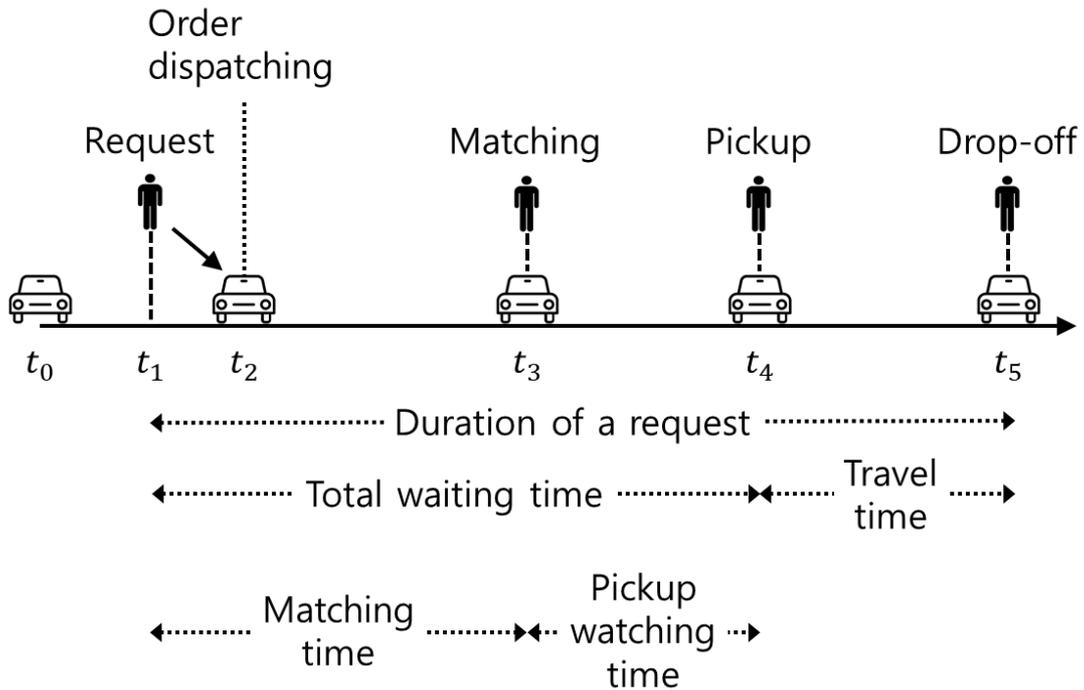


Figure 2. Time intervals that constitute a single request

Pickup waiting time is a key for this study. While many studies have focused on the matching time, such as the batching window for assignment, this study focuses on pickup waiting time. As for Singapore data which was used in this study, pickup waiting time occupies the majority of the total waiting time (Table 2). Therefore it is necessary to address pickup waiting time. Pickup waiting time is related to the negative externalities which are

stated in the previous part. If the pickup waiting time is decreased, it leads to shorter distances covered by the vehicle. This reduction in travel distance contributes to lowering the vehicle's carbon footprint, thus aiding in the reduction of negative externalities.

Table 2. Comparison of matching time and pickup waiting time in historical data

	Matching time (min)	Pickup waiting time (min)
Mean	0.26	6.20
STD	0.17	2.90

When pickup waiting time is reduced, riders can save time before boarding, and drivers can decrease non-profitable travel distance and save fuel. As for the platform company, the company can reduce carbon taxes without losing riders (revenue). Regarding society, if the pickup waiting time is reduced, the distance traveled by the vehicle decreases, resulting in a reduction in congestion and carbon footprint.

1.2. Research Purpose

This study aims to develop a matching and pricing algorithm in a ride-hailing service based on reinforcement learning to reduce pickup waiting time. This is achieved through several steps:

- 1) assigning a surge multiplier to each request;
- 2) matching riders and drivers while reflecting the spatio-temporal value of each grid;
- 3) developing a base model that utilizes historical data;
- 4) adding a term that reduces pickup waiting time to the base model;
- 5) investigating how the model benefits the platform, riders, drivers, and society.

After the base model is developed, the effect of changes in price sensitivity and value term coefficient is investigated to analyze how the model behaves with varying parameters. In addition, the applicability of the model is studied to check if it is compatible with different cities.

The contributions of this research are twofold. First, this study treats pricing and reducing pickup waiting time together, considering rider (passenger) convenience. Second, this study investigates how much potential pickup waiting time can be saved in a ride-hailing service. This information can be utilized as evidence of the effectiveness of ride-hailing services, helping them determine how much their service can potentially alleviate negative externalities.

The remainder of this paper is organized as follows. Chapter 2 provides the literature review on characteristics of ride-hailing

service, dynamic pricing, and negative externalities. Chapter 3 presents the methodology of this research including contextual bandit and temporal difference learning. Data description is presented in Chapter 4. This chapter presents the data tables employed in the study and examines the spatial distribution of the data. Chapter 5 explains the results and discussion. Lastly, concluding remarks and future research are presented in Chapter 6.

Chapter 2. Literature Review

2.1. Concepts

2.1.1. Characteristics of Ride-hailing Service

Academic interest in ride-hailing services have been increasing. To deal with ride-hailing service, identifying the characteristics of ride-hailing service is necessary. The characteristics are summarized into 5 types, 4 of which were summarized in Agatz et al. (2012).

First, ride-hailing services are dynamic: requests can occur anytime and anywhere. This demand-side fluctuation is the cause of an imbalance between demand and supply. Second, ride-hailing is not a recurring service at any given time, and each request is considered non-repeating. Third, ride-hailing is a two-sided market based on prearrangements, where service provision is made under the mutual consent of two sides. Fourth, when called through the app, ride-hailing requests are automatically matched with a driver by an internal algorithm. Lastly, there are no holding or backorder costs (Bertsimas and Perakis, 2006). In a ride-hailing service, a request is not a tangible product, unlike typical manufacturing.

2.1.2. Definition of Dynamic Pricing

Dynamic pricing can manipulate supply and demand: a higher price would attract more drivers while delaying requests from riders who are not in a hurry; and a lower price does the opposite. In most cases, dynamic prices are represented by price multipliers, and the fare of a trip is the normal price multiplied by the price multiplier. The fixed normal price is determined by travel time and distance, while the dynamic surge multiplier depends on conditions of supply and demand (Shen, 2021). Hence, dynamic pricing can be written as:

$$\text{Dynamic price} = \text{Surge multiplier} \times \text{Base price} \quad (1)$$

In Equation (1), the fixed normal price is referred to as the base price. Additionally, by analyzing Equation (1), it becomes evident that dynamic pricing affects the driver's revenue; specifically, the higher the dynamic pricing multiplier, the greater the driver's earnings.

2.1.3. Ride-hailing Service as a Sequential Decision Problem

The service process of ride-hailing can be expressed as a continuous decision-making problem. Markov decision process (MDP) is a useful analytical tool for solving ongoing decision-making problems under uncertainty (Alagoz et al., 2010).

In the MDP, the next state of the process only depends on the current state and the action, but not on any previous state or action (Rong et al., 2016). This can derive exact optimal policies in the long run in a stochastic context (Legros, 2019). For solution methods of the MDP, dynamic programming, evolutionary algorithms, and reinforcement learning are widely applied.

To address a sequential decision under uncertainty, reinforcement learning is a powerful technique to adaptively improve the policy (Qin et al., 2021). It is challenging to search for the optimal matching and pricing as they interact with each other, and especially matching is compounded with past actions. Reinforcement learning can efficiently solve these complex decision-making processes.

2.2. Dynamic Pricing

Many researchers have tried to improve dynamic pricing. Korolko et al. (2018) introduced the concept of dynamic waiting to the dynamic pricing. Vehicles are matched while changing the waiting time of the consumer in time and space. A steady-state model for pricing and matching in ride-hailing was studied. The platform determines the dynamic price as well as window of time to wait before dispatching drivers to incoming ride requests. The system equilibrium of supply and demand under different prices and waiting windows was characterized. This study showed that dynamic pricing and matching are critical to reduce waiting times for both riders and drivers. Using data from Uber, the steady-state model was calibrated. The study showed that price and its uncertainty are reduced by jointly optimizing price and waiting window.

Bimpikis et al. (2019) offered incentives to drivers. This research make drivers move to areas with high demand. Profits and consumer surplus are maximized when the demand pattern is balanced across the network's locations. In addition, profits and consumer surplus both increased monotonically with the balancedness of the demand pattern. Furthermore, if the demand pattern was not balanced, the platform could benefit substantially from pricing rides differently depending on the location they originate from. It is meaningful that this is the first study to

differentiate prices based on passenger location. New strategy showed better supplier profit and consumer surplus.

He and Shin (2019) balanced between zones. Research started from the problem that existing pricing policies only respond to short-term demand fluctuations without accurate trip forecast and spatial demand-supply balancing, thus mismatching drivers to riders and resulting in loss of profit. A new spatio-temporal deep capsule network (STCapsNet) that accurately predicts ride demands and driver supplies with vectorized neuron capsules while accounting for comprehensive spatio-temporal and external factors was used. Given accurate perception of zone-to-zone traffic flows in a city, developed adaptive pricing scheme formulated a joint optimization problem by considering spatial equilibrium to balance the platform, providing drivers and riders with proactive pricing signals.

2.3. Negative Externalities

Studies on negative externalities are also being actively conducted. Agarwal et al. (2023) reported a discernible drop in travel time during periods of ride-hailing unavailability due to driver strikes in India. This research commenced study from that early research has documented significant growth in ride-hailing services worldwide and allied benefits. However, growing evidence of ride-hailing services negative externalities were leading to significant policy scrutiny. Despite demonstrated socioeconomic benefits and consumer surplus worth billions of dollars, cities were choosing to curb these services in a bid to mitigate first order urban mobility problems. Existing studies on the congestion effects of ride-hailing were limited, reported mixed evidence, and exclusively focused on the United States, where the supply consists primarily of part-time drivers. This research studied how the absence of ride-hailing services affects congestion levels in three major cities in India, a market where most ride-hailing drivers participate as full time. Using rich real-time traffic and route trajectory data from Google Maps, they showed that in, all three cities, periods of ride-hailing unavailability due to driver strikes saw a discernible drop in travel time. The effects were largest for the most congested regions during the busiest hours, which saw 10.1%-14.8% reduction in travel times. Additionally, suggestive evidence for some of the mechanisms behind the

observed effects, including deadheading elimination, substitution with public transit, and opening up of shorter alternative routes were provided. These results suggest that despite their paltry modal share, ride-hailing vehicles are substituting more sustainable means of transport and are contributing significantly to congestion.

Beojone and Geroliminis (2021) investigated how shared rides and dynamic congestion are affected by riders' willingness to share. Study started from the problem that although ride-hailing is a notorious service, little is known about to what degree its operations can interfere in traffic conditions, while replacing other transportation modes, or when a large number of idle vehicles is cruising for passengers. The efficiency of ride-hailing services using taxi trip data from a Chinese megacity and an agent-based simulation with a trip-based macroscopic fundamental diagram (MFD) model for determining the speed were experimentally analyzed. The effect of expanding fleet sizes, passengers' inclination towards sharing rides, and strategies to alleviate urban congestion was investigated. The research observed that, although a larger fleet size reduces waiting time, it also intensifies congestion, increasing the total travel time. Such congestion effect was so significant that it is nearly insensitive to passengers' willingness to share and flexible supply. Finally, parking management strategies could prevent idle vehicles from cruising without assigned passengers, mitigating the negative impacts of

ride-sourcing over congestion, and improving the service quality.

Pollution is the other part of negative externalities. Barnes et al. (2020) investigated the effects of introducing ride-hailing services on pollution in a Chinese city. This study used a difference-in-difference approach to longitudinally examine cities that introduced the Didi Chuxing service. The results suggest that the announcement of the ride-hailing service led to a decrease in PM2.5 pollution during the month before the service launch and the early months of the introduction. However, this effect is temporary, as the growth in ride-hailing vehicles and a rapid increase in the number of overall trips intensify vehicle pollution.

Tirachini and Gomez-Lobo (2020) have pointed out the importance of determining the impact of ride-hailing on vehicle kilometers traveled (VKT), and thus on transport externalities like congestion. Survey results on Uber use by residents of Santiago, Chile were used. They also used information from other studies to parameterize a model to determine whether the advent of ride-hailing applications increases or decreases the number of VKT. Given the intrinsic uncertainty on the value of some model parameters, a Monte Carlo simulation for a range of possible parameter values was used. Results indicated that ride-hailing increases vehicle kilometers traveled (VKT) unless ride-hailing applications substantially increase the average occupancy rate of trips and induce shared or pooled ride-hailing.

2.4. Implications

From the previous studies, concepts about ride-hailing and dynamic pricing are covered. In addition, studies to improve dynamic pricing are reviewed. Papers which observed or quantified the negative externalities incurred by ride-hailing service are also reviewed. There have been many attempts to reduce match waiting time in a dynamic pricing scheme, but there have been few studies to reduce pickup waiting time.

Reducing pickup waiting time is beneficial for both stakeholders of ride-hailing service. Unlike taxis that used to operate by wandering around, vehicles of ride-hailing services need to move to pick up riders once matched with them. This additional time, which did not exist before, is the pickup waiting time for riders, and the deadheading time for drivers. Generally, riders do not prefer to wait long to board, so the shorter the time, the better the quality of the ride-hailing service. Reducing pickup waiting time is also beneficial to drivers, as it is not profitable while imposing drivers' time and fuel costs. Moreover, decreasing pickup waiting time is also helpful in reducing negative externalities by decreasing the distance traveled by vehicles. Therefore, this study aims to develop a matching and pricing model to reduce pickup waiting time.

Figure 4 shows the difference between a full reinforcement learning problem and a contextual bandit. Contextual bandit algorithm is a simplified reinforcement learning approach. Unlike a full reinforcement learning, states of contextual bandit are not determined by the previous states or actions. Selecting a surge multiplier for a request can be modeled as selecting an arm at slot machines.

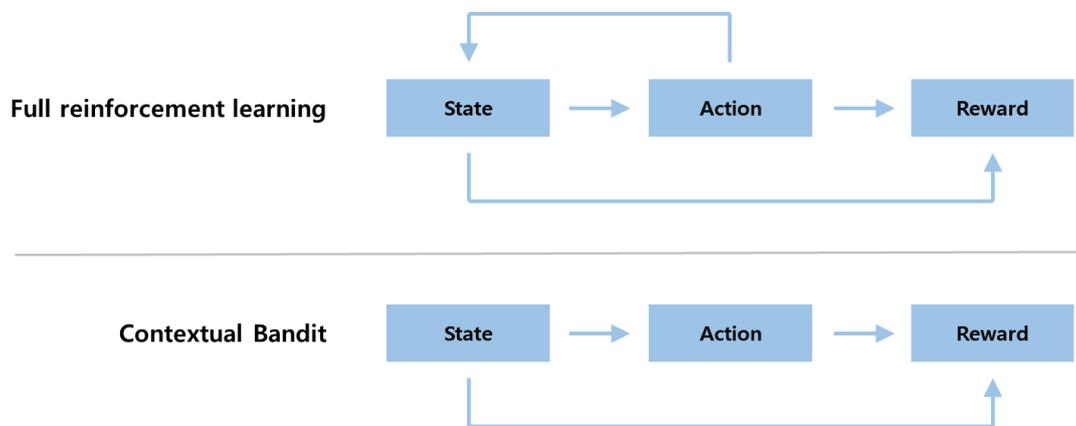


Figure 4. Difference between a full reinforcement learning problem and a contextual bandit

Source: <https://medium.com/@awjuliani/simple-reinforcement-learning-with-tensorflow-part-1-5-contextual-bandits-bff01d1aad9c>

Contextual bandit needs a context vector, and the i^{th} request is characterized by a 5-dimensional context vector x_i .

$$x_i = [t_i, l_i, l'_i, p_i, d_i]$$

t_i : The time request occurs

l_i : Pickup location

l'_i : Drop-off location (2)

p_i : The estimated base price determined by the estimated trip distance and time

d_i : The estimated distance from pickup location to drop-off location in meter

When the surge multiplier a is applied, the expected payoff u is expressed by the linear product of (1×5) context vector x and (5×1) parameter θ_a .

$$E\{u(x, a)|x\} = x^T \theta_a \tag{3}$$

θ_a can be estimated using the ridge regression method where D_a represents a $(m \times 5)$ matrix which includes m training inputs (contexts) observed before the current trial, c_a is the corresponding payoff vector, and I_d is a (5×5) identity matrix.

$$\hat{\theta}_a = (D_a^T D_a + I_d)^{-1} D_a^T c_a \tag{4}$$

To select a_i , one popular algorithm for solving contextual bandit problems is the LinUCB algorithm, which is an extension of the upper confidence bound (UCB) algorithm for the multi-armed bandit problem (Li et al., 2010). The algorithm works by maintaining an estimate of the expected reward for each action, conditioned on the available context. The algorithm then chooses the action with the highest upper confidence bound based on the estimated reward and the uncertainty associated with that estimate.

The key advantage of LinUCB is its use of a linear model to estimate the expected reward function, which allows it to efficiently handle large-scale contextual bandit problems. The algorithm also uses a Bayesian framework to update its estimates and confidence intervals based on the observed rewards and contexts, which allows it to handle non-stationary environments and adapt to changes over time.

In addition, LinUCB uses a regularization term to prevent overfitting to the training data and to encourage exploration of new actions. The regularization term is based on the L2 norm of the model parameters, which helps to control the complexity of the model and prevent overfitting.

Overall, LinUCB is a powerful and flexible algorithm for solving contextual bandit problems, with many practical applications in areas such as online advertising, recommendation systems, and personalized medicine. Its use of a linear model, Bayesian

framework, and regularization term allow it to efficiently handle large-scale contextual bandit problems, while balancing the trade-off between exploration and exploitation.

3.2. Temporal Difference Learning

Temporal Difference (TD) learning is a type of reinforcement learning algorithm that is used in artificial intelligence and machine learning. TD learning is a model-free approach to reinforcement learning, which means that it does not require a complete model of the environment. Instead, it uses a combination of prediction and control techniques to learn from experience and optimize decision making. In particular, it estimates the value function, which represents the value of a state. The value function is essential for TD learning because it helps the agent to determine which actions will lead to a higher reward in the long run.

The value function is typically represented as a function of the state, denoted as $V(s)$, where s is a particular state. The value of a state is the expected cumulative reward that the agent will receive starting from the state and following a particular policy. This means that the value function is defined recursively, as the value of a state depends on the values of the neighboring states.

The TD learning algorithm works as follows. At each time step, the agent observes the current state of the environment, selects an action to take based on its current estimate of the value

function, and then receives a reward and transitions to a new state. The agent then updates its estimate of the value function using the temporal difference error, which is the difference between the predicted value of the current state and the actual value obtained from the reward and the next state. The value function is updated by adding a fraction (learning rate) of the TD error (TD target minus the current estimate of the value) to the current estimate of the value of the state. This procedure is summarized as Figure 5. This process is repeated for each state visited during the agent's interactions with the environment. By updating the value function, the agent can learn to maximize its rewards by selecting actions that lead to higher values. Over time, the agent becomes more accurate in estimating the values of the states, which leads to better decisions and improved performance.

$$\begin{array}{c}
 \text{Learning} \\
 \text{rate} \\
 \hline
 \underline{V(S_t)} \leftarrow \underline{V(S_t)} + \alpha [\underline{R_{t+1}} + \underline{\gamma V(S_{t+1})} - \underline{V(S_t)}] \\
 \begin{array}{l}
 \text{New value} \\
 \text{of state } t
 \end{array}
 \quad
 \begin{array}{l}
 \text{Former} \\
 \text{estimation of} \\
 \text{value of state } t
 \end{array}
 \quad
 \begin{array}{l}
 \text{Reward}
 \end{array}
 \quad
 \begin{array}{l}
 \text{Discounted value of} \\
 \text{next state}
 \end{array}
 \quad
 \begin{array}{l}
 \text{TD target}
 \end{array}
 \end{array}$$

Figure 5. Update of value function

One of the advantages of TD learning is that it allows the agent

to learn from incomplete and noisy information. It also allows the agent to learn online, meaning that it can update its estimate of the value function as it receives new information and feedback from the environment. This makes it well-suited to applications in which the environment is complex and dynamic, such as game-playing, robotics, and autonomous systems.

In summary, TD learning is a powerful and effective technique for learning and decision-making in reinforcement learning. Its ability to learn from incomplete and noisy information and to update its estimates online makes it well-suited to complex and dynamic environments.. In this study, driver' s spatio-temporal value function of each grid is updated when total pricing and assignment is completed. Value term, the difference of driver' s values between the destination and the current location is used for reward.

3.3. Model Formulation

3.3.1. Semi-MDP (Markov Decision Problem)

Let (l_j, t_j) be the state of a driver j , and assume that the driver accepts a request i . If the rider's destination is l'_i and the sum of pickup waiting time and travel time is T_i , the state transition of the vehicle is expressed by Equation (5).

$$s_j = (l_j, t_j) \rightarrow (l'_i, t_j + T_i) \quad (5)$$

Location l is discretized using a hexagonal grid. Time t is also discretized using a fixed batching window, making the state discrete as well. There exists two kinds of actions in this study. The first action is the surge multiplier, a_i . It takes one number from set $\{0.8, 0.9, 1.0, 1.1, 1.2, 1.3\}$. The authority in Singapore (Public Transport Council) does not set any price ceiling, rather leave fare to be determined by market forces. Although there is no price ceiling regulation, the set was formed considering realistic surge multipliers. The second action is a binary vector $b_j = \langle b_{ji} \rangle$ representing whether the request is accepted. This vector appears in the integer programming (Equation (6)), which maximizes the total reward for all requests subject to each request or driver matched up to one counterpart. This optimization problem is

solved using Kuhn–Munkres algorithm.

$$\begin{aligned}
& \max_b \sum_{j \in \mathcal{J}} \sum_{i \in \mathcal{J}} v_\pi(i, j) b_{ji} \\
& \sum_{j \in \mathcal{J}_i} b_{ji} \leq 1, \forall i \in \mathcal{J} \\
& \text{s.t.} \quad \sum_{i \in \mathcal{J}_j} b_{ji} \leq 1, \forall j \in \mathcal{J} \\
& \quad \quad b_{ji} \in \{0, 1\}
\end{aligned} \tag{6}$$

Reward of the MDP is defined as Equation (7). This is the case of base model and the model reducing the pickup waiting time will be explained later. Reward of each request is summation of driver's revenue and value term. Basically this reward leads assignment and pricing to revenue maximization like previous studies and ride-hailing services seek. Since value function reflects the potential of the location, higher value term means the driver has a potential to make more revenue. It is consistent with the possibility that the following request may occur at the destination. Elements of reward function are listed in Table 3.

$$\begin{aligned}
v_\pi(x_i, a_i) = & \sum_j f(x_i, a_i) b_{ji} \times \\
& [p_i a_i + \gamma (V_\pi(l'_i, t_j + T_i) - V_\pi(l_j, t_j))]
\end{aligned} \tag{7}$$

Table 3. Elements of reward function

Element	Definition	Element	Definition
i	Individual Hailing Request	j	Individual Driver
b_{ji}	Dispatch Action of a Driver j	x_i	Context of Request i
p_i	Base Price of Request i	a_i	Surge Multiplier (0.8, 0.9, 1.0, 1.1, 1.2, 1.3)
l'_i	Destination of Request i	T_i	Time Consumed to Complete Request i
s_j	$s_j = (l_j, t_j)$: State of Driver j	$f(x_i, a_i)$	Conversion Rate Function
π	Policy (Surge Multiplier and Dispatch)	$V_\pi(s_j)$	Value Function of Driver j at State s under Policy π

f refers to the conversion rate function, indicating the rate at which riders actually make requests based on the surge multiplier. Chen et al. (2019) employed a linear function that increases as the surge multiplier increases. However, this approach has the limitation of focusing solely on the driver's perspective, as they are more inclined to accept requests with higher prices. In real-life service usage scenarios, when fares rise, users may feel burdened by the increased cost of the request. Therefore, a new term that decreases as surge multiplier increases is introduced into the conversion rate function, as depicted in Equation (8). For initial implementation, $f_0 = 0.5$ and $\alpha = 1.5$ were applied and then

adjusted.

$$f = (f_0 + \alpha \times (a - 1)) \times (f_0 - \alpha \times (a - 1)) \quad (8)$$

3.3.2. Bootstrapping Structure

As previously mentioned, decision variable a_i of each request is calculated using contextual bandit algorithm. When total pricing and assignment is completed, drivers value function is updated using temporal difference learning. Updated value function is followed by calculating decision variable b_{ji} using Kuhn–Munkres algorithm. As this procedure is repeated, decision variables are converged. This iterative training shows that both the immediate and future rewards of a pricing decision variable a_i are closely dependent on the dispatch decision variable b_j on the current time period. Inter-dependency between pricing and dispatch exists. This bootstrapping structure is illustrated in Figure 6. Figure 7 shows updating the algorithm creates new situations, thus matching and pricing affects on the next phase.

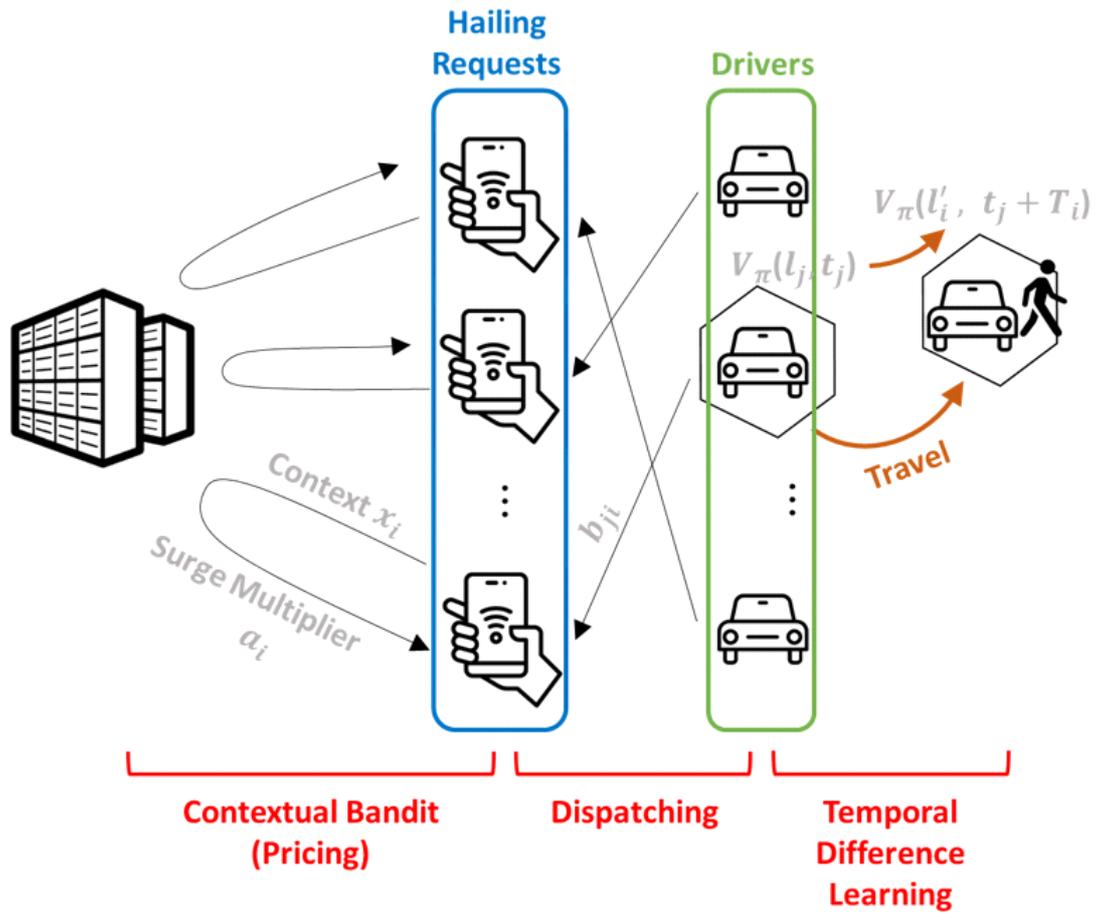


Figure 6. Bootstrapping structure of algorithm

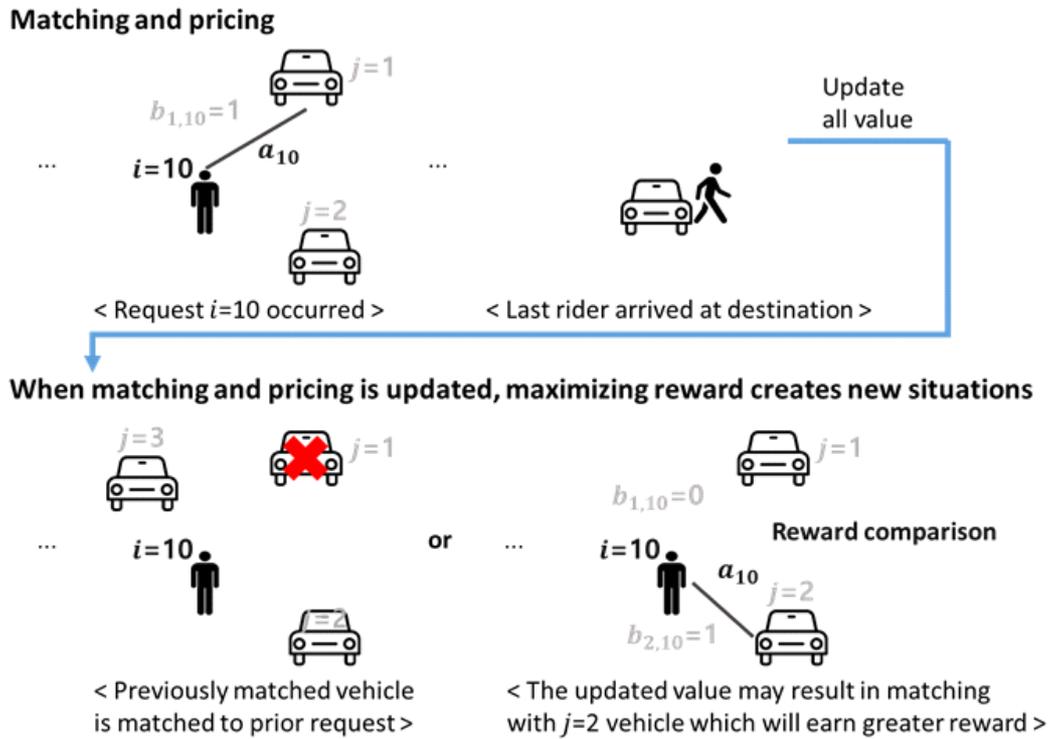


Figure 7. An example illustrating that updating the algorithm results in new situations

Chapter 4. Data Description

4.1. Data Tables

This research utilized data collected from the TADA service in Singapore. Utilization of this dataset allowed for a comprehensive analysis of the service's operational dynamics. The dispersion of data across space and time aided in defining the study's scope. The temporal scope of this study is established during the morning peak from 7:00 to 8:30. This determination is grounded in the hourly distribution of daily requests on Monday, November 16, 2020 (Figure 8).

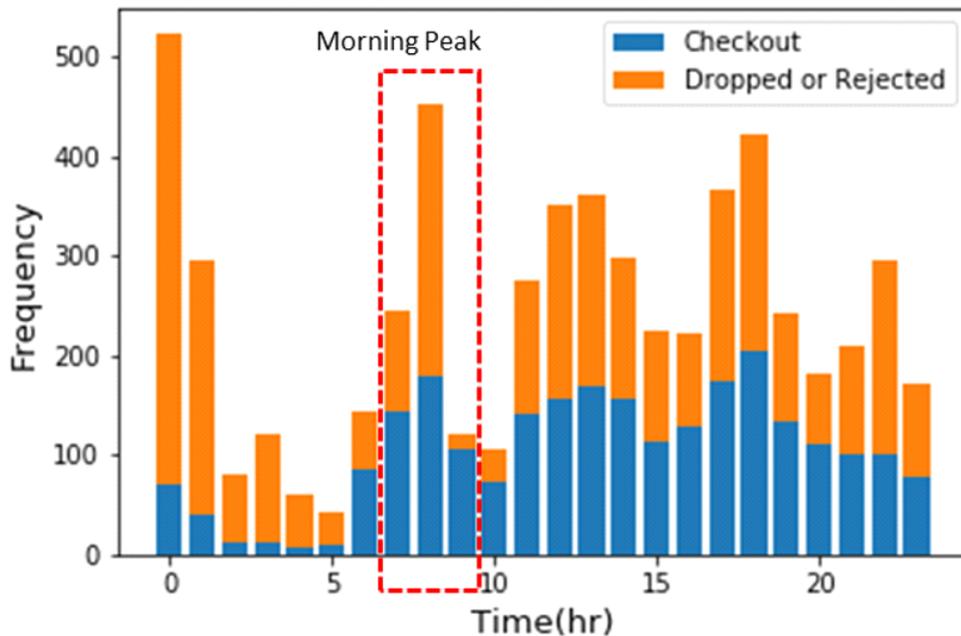


Figure 8. Hourly hailing requests of November 16, 2020

The spatial scope is defined within the eastern part of Bukit Panjang Station (Figure 9), encompassing the region where over 75% of all requests originate and conclude. This consideration takes into account the distribution of road networks, downtown, and residential areas. This scope also encompasses Marina Bay, which serves as the central business district of Singapore.

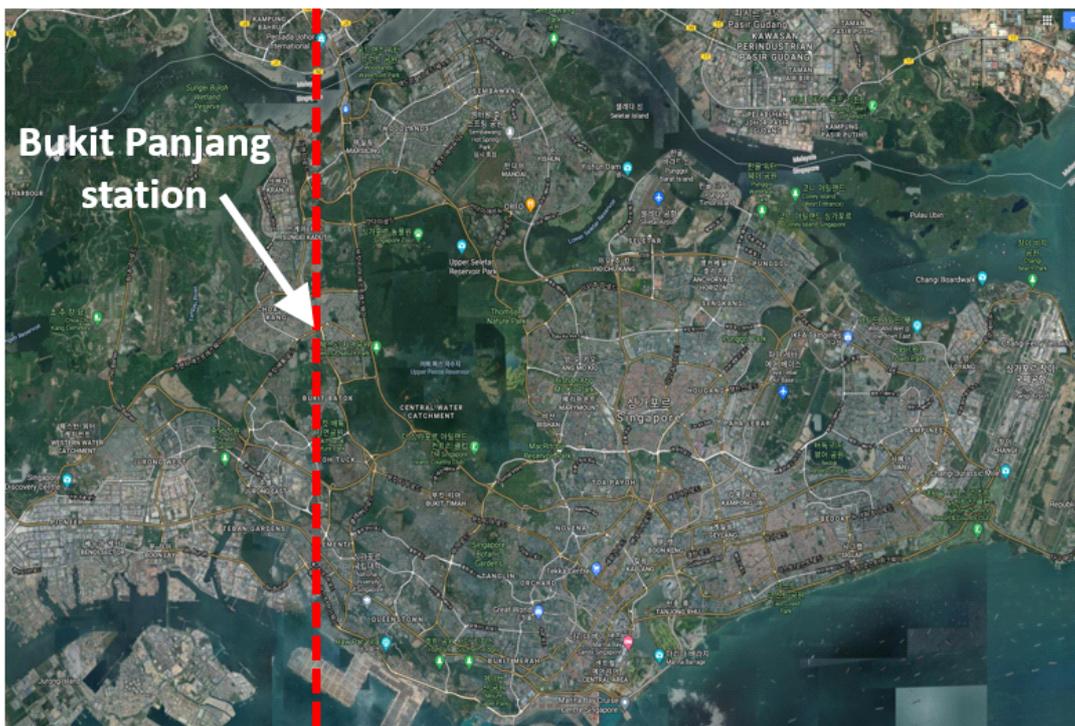


Figure 9. Spatial scope of study

Within the array of data tables, the ride data encompasses information generated when riders make requests in the TADA app. This dataset undergoes preprocessing, involving tasks such as cleaning up missing values. Notably, the preprocessing stage

yielded some findings. Rows with blank driver IDs are indicative of cancellations, which can be initiated by either the rider or the driver. Moreover, all checkout completed requests consistently included driver IDs. However, it's important to highlight that the pickup distance is absent as a distinct column in this dataset. The detailed list of columns are in Table 5.

The ping data, on the other hand, consists of information transmitted from requests to nearby drivers. After preprocessing over a span of approximately two months, this dataset comprises an impressive 18,685,392 rows of data. A comprehensive breakdown of the columns comprising the ping data table is available in Table 6.

Upon merging the pickup distance of the checkout request with the ride data, a significant insight emerges: the total number of checkouts reached a noteworthy 247,233, as elucidated in Table 4. Through the integration of these diverse data sources, a comprehensive understanding of the dynamics that define TADA services can be achieved.

Table 4. Number of requests by type after preprocessing

Type	Frequency	Daily average
Checkout (Matched)	247,233	4,053
Dropped+Rejected	493,507	8,090
Total	740,740	12,143

Table 5. Columns of ride data table

Column name	Type
ride_id	object
driver_uuid	object
RecordTime	datetime64
PickupArrivalTime	datetime64
TripStartTime	datetime64
FinishTime	datetime64
RecordDate	object
Status	int64
car_type	int64
dest_lat	float64
dest_long	float64
start_lat	float64
start_long	float64
estimated_distance	int64
surge_fixed	float64
surge_multiplier	float64
total_price	float64
smart_call_fee	float64
base_price	float64
dest_h3_id_r8	object
origin_h3_id_r8	object
dest_h3_id_r9	object
origin_h3_id_r9	object

Table 6. Columns of ping data table

Column name	Type
ride_id	object
driver_uuid	object
pickup_distance	float64
driver_location_latitude	float64
driver_location_longitude	float64
driver_h3_id_r8	object
driver_h3_id_r9	object
pickup_location_latitude	float64
pickup_location_longitude	float64
pickup_h3_id_r8	object
pickup_h3_id_r9	object
round	int64
rejected	bool
accepted	bool
created_at	datetime64
updated_at	datetime64
timedelta	timedelta64
mdt	bool
ride_type	float64

4.2. Descriptive statistics

Descriptive statistics for Monday, November 16, 2020, are presented in Table 7. On this day, there were 2,730 checkout (completed) requests, contributing to a total of 5,956 requests encompassing completed checkouts, dropped, and rejected requests. An analysis of the descriptive statistics reveals that drivers who fulfilled TADA app requests throughout the day averaged 2.2 requests each. Furthermore, the application of a T-test yielded that, the examination of travel distance and fare revealed no significant statistical difference between completed checkout requests and the combined set of dropped and rejected requests (D+R).

Table 7. Descriptive statistics for Monday, November 16, 2020

	Target	Min	Max	Average	Standard Deviation	p-value
Requests per vehicle	Checkout	1	25	2.22	2.11	–
Pickup distance (m)	Checkout	1,000	10,000	1,656	1,006	–
Travel distance (m)	Checkout	639	49,767	10,594	7,706	0.65
	D+R	11	46,803	10,506	7,165	
	Total	11	49,767	10,546	7,417	–
Revenue (SGD)	Checkout	5.00	46.80	11.31	5.18	0.71
	D+R	5.00	34.00	11.26	4.82	
	Total	5.00	46.80	11.28	4.99	–

4.3. Spatial Distribution

In order to investigate the spatial distribution of requests, we confined our focus to the predefined spatial and temporal boundaries of the study. During the designated time frame of 7:00 to 8:30, a total of 549 requests were generated. Remarkably, out of this total, 384 requests both originated and terminated within the specified spatial range, a representation that can be observed in Figure 10.

To provide a visual representation, the origin and destination of each request were aggregated utilizing the H3 resolution 7, where each hexagon's side spans a length of 1.22 km. Notably, the distribution of requests per grid is far from uniform, with certain hexagons registering as many as 28 requests for both origin and destination points.

The origins of these requests predominantly align with residential areas concentrated in the northern region of Singapore. In contrast, the destinations are primarily centered around the bustling downtown region encompassing Marina Bay in the southern part of Singapore. This contrast in spatial distribution underscores the non-homogeneous characteristics of the demand and supply dynamics at play within this context.

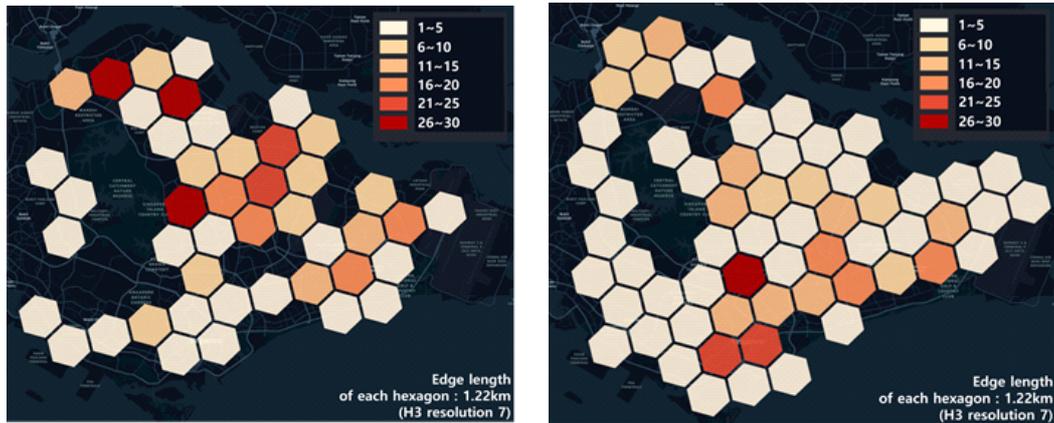


Figure 10. Spatial distribution of origins (left) and destination (right)

Spatial distribution is categorized into matched requests and unmatched requests. Interestingly, when examining both origins and destinations, the regions with the highest request frequencies varied between matched and unmatched requests. Regarding origins specifically, Figure 11 illustrates that the areas with the most matched requests were Woodlands, recording 23 cases, and Yishun, with 18 cases. On the other hand, the highest number of unmatched requests, totaling 19 cases, was observed in Ang Mo Kio.

Regarding the destinations, Figure 12 highlights that Yishun and Siglap both exhibited the highest number of matched requests, each totaling 13 cases. In contrast, Novena recorded the highest number of unmatched requests, with a count of 20 cases.

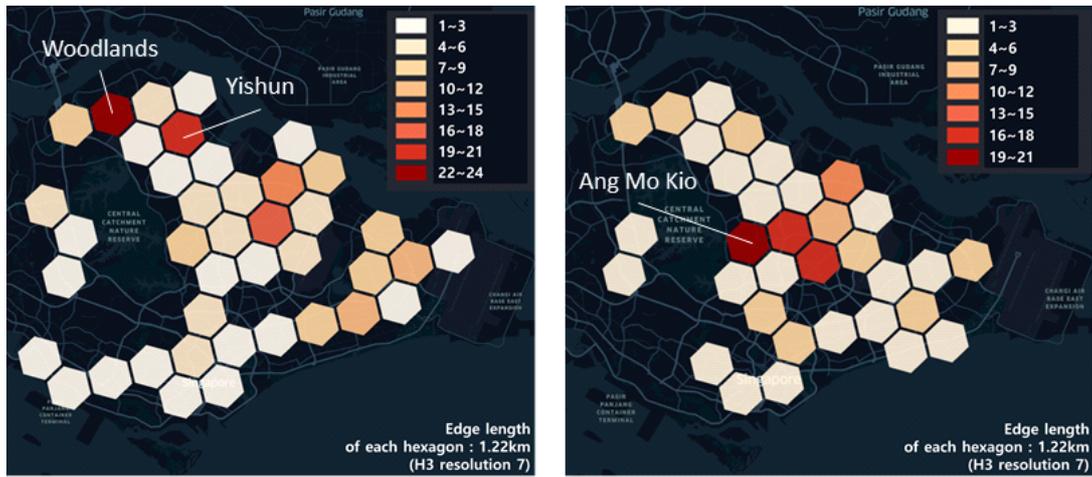


Figure 11. Spatial Distribution of Origins of Matched (Left) and Unmatched (Right) Requests

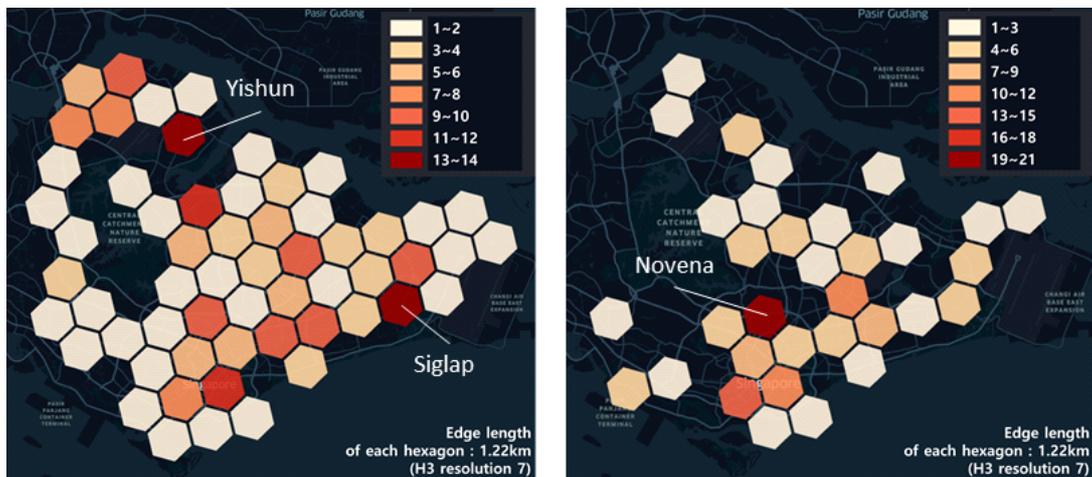


Figure 12. Spatial Distribution of Destinations of Matched (Left) and Unmatched (Right) Requests

Chapter 5. Results

5.1. Base Model

In this study, base model refers to a model that has been detailed in Chapter 3. This model is compared with historical data and will subsequently incorporate a pickup cost term in Section 5.2.

The experimental setup for the base model includes the following components:

- Only one market player exists, so this study does not take into account other market participants
- 384 requests from 7:00 to 8:30 which is morning peak.
- When matching requests with drivers, an exploration is conducted within three hexagon rings (H3 resolution 7, with an edge length of 1.22km, resulting in a radius of 5.9km). This setup closely emulates real-world ride-hailing services.
- Implementation of a 5-minute batching window, which aggregates requests for assignment.
- Surge multipliers encompassing values {0.8, 0.9, 1.0, 1.1, 1.2, 1.3} are considered, mirroring conditions found in actual environments. Subsequently, the study proceeds to test surge multipliers with higher magnitudes in Section 5.5.
- Application of the conversion function (Equation 8) with an

initial simulation incorporating values for $f_0=0.5$ and $\alpha=1.5$, followed by calibration to achieve a value of $\alpha=1.55$. Notably, the exploration of α is conducted to assess the implications of alterations in price sensitivity, as detailed in Section 5.3.

– A multiplication of 20 ($\beta=20$) is applied to the value function term, a decision arrived at based on the outcomes of trial simulations. From reward function, two terms (revenue term and value term) are compared at trial simulation. From Table 8, since maximum of revenue term is 20 times greater than the value term, $\beta=20$ is selected and tested to investigate the effect of changes in value term coefficient (Section 5.4).

Table 8. Comparing two terms of trial simulation

	Revenue Term	Value Term
Mean	10.7	0.0
SD	5.1	0.2
Min	3.2	-2.2
25%	6.9	-0.1
50%	9.7	-0.0
75%	13.7	0.1
Max	39.1	2.2

The matching rate of the base model is calculated as 204 out of 384, resulting in a percentage of 53.1%. In contrast, the historical data exhibits a matching rate of 228 out of 384, corresponding to a percentage of 59.4%. To observe the algorithm's convergence, the trends of the proportions for each surge multiplier are

visualized in Figure 13. The number of occurrences for each surge multiplier tends to stabilize and remain constant as iterations progress.

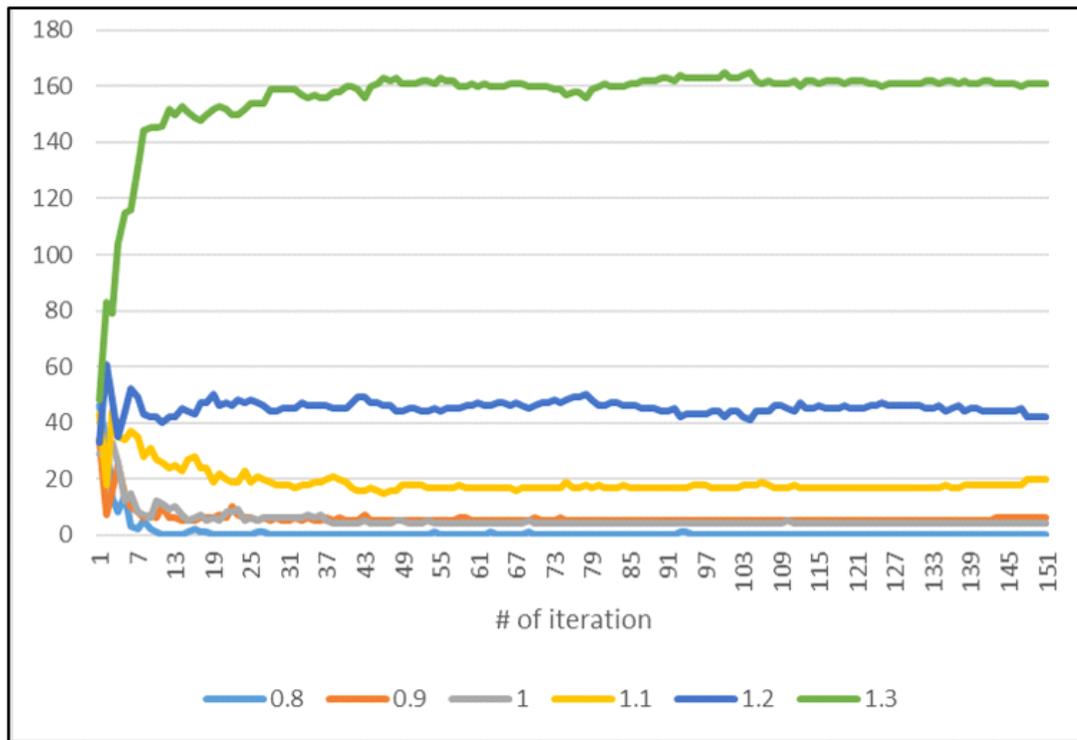


Figure 13. The proportion of each surge multiplier converges as iteration progresses

To assess convergence, Figure 13 has been transformed into Figure 14. The percentage of requests with altered surge multipliers in the subsequent iteration, relative to the total number of requests, stabilizes at 0% by the 30th iteration.

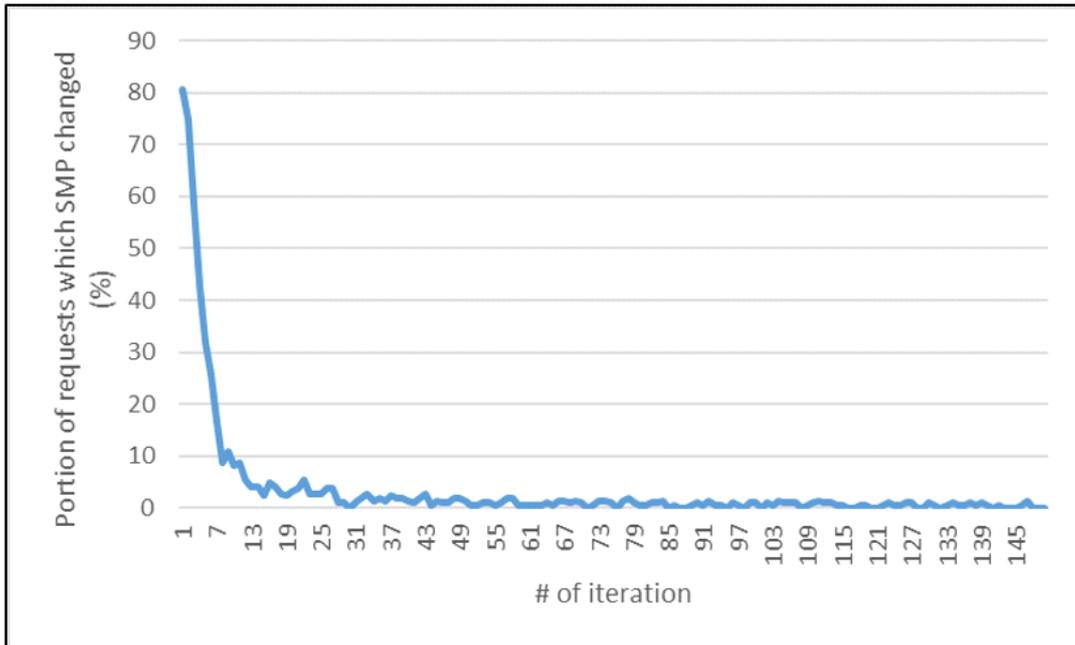


Figure 14. The percentage of requests with altered surge multipliers in the subsequent iteration

For a comprehensive comparison with historical data, the matching rate based on travel distance has been graphed (Figure 15) and is also provided in Table 9. The error rate typically remains within 15% when aggregated in 2 km units. The simulation results closely align with the historical matching rate.

The matching rate based on time has been depicted in a plot (Figure 16) and is presented in Table 10. The error rate reaches its peak at -42.9% within the initial 10 minutes. This highest error rate can be attributed to the relatively low volume of requests during this time frame. In a broader perspective, the error rates remain within the range of 20% across all intervals.

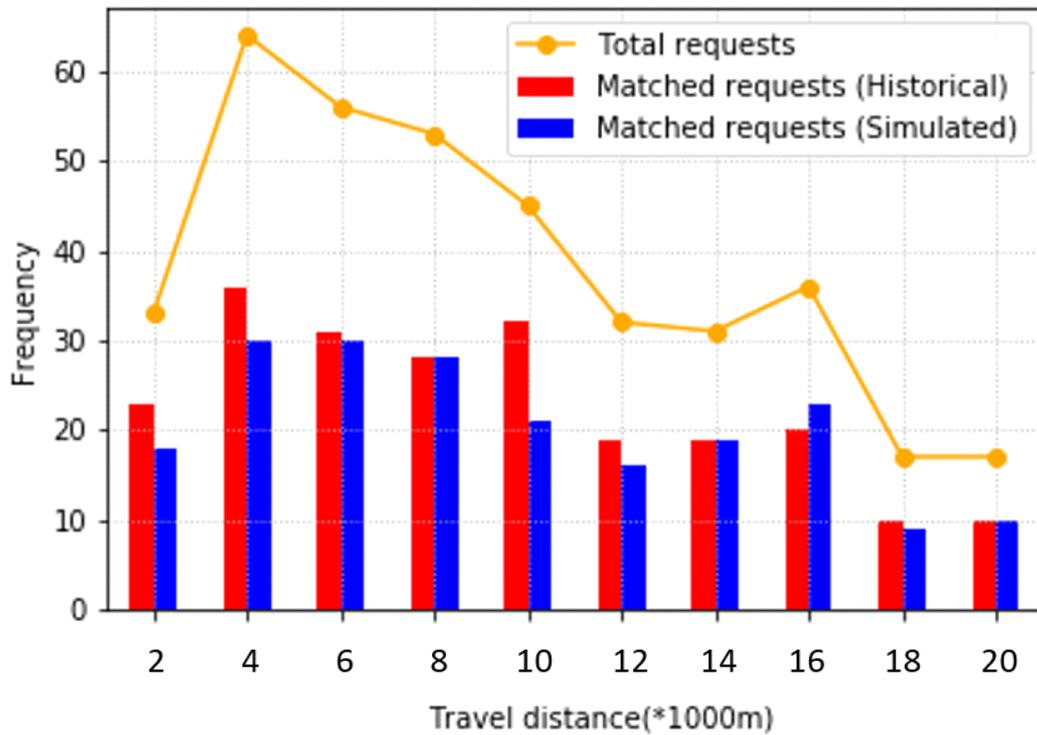


Figure 15. Matching rate by travel distance

Table 9. Matching rate and error rate by travel distance

Travel distance (km)	Requests	Matched requests	Matching rate (A)	Sim_matched requests	Sim_matching_rate (B)	Error_rate (B-A) (%)
1~3	33	23	69.7	18	54.5	-15.2
3~5	64	36	56.3	30	46.9	-9.4
5~7	56	31	55.4	30	53.6	-1.8
7~9	53	28	52.8	28	52.8	0.0
9~11	45	32	71.1	21	46.7	-24.4
11~13	32	19	59.4	16	50.0	-9.4
13~15	31	19	61.3	19	61.3	0.0
15~17	36	20	55.6	23	63.9	8.3
17~19	17	10	58.8	9	52.9	-5.9
19~21	17	10	58.8	10	58.8	0.0
Total			59.4	-	53.1	-6.3

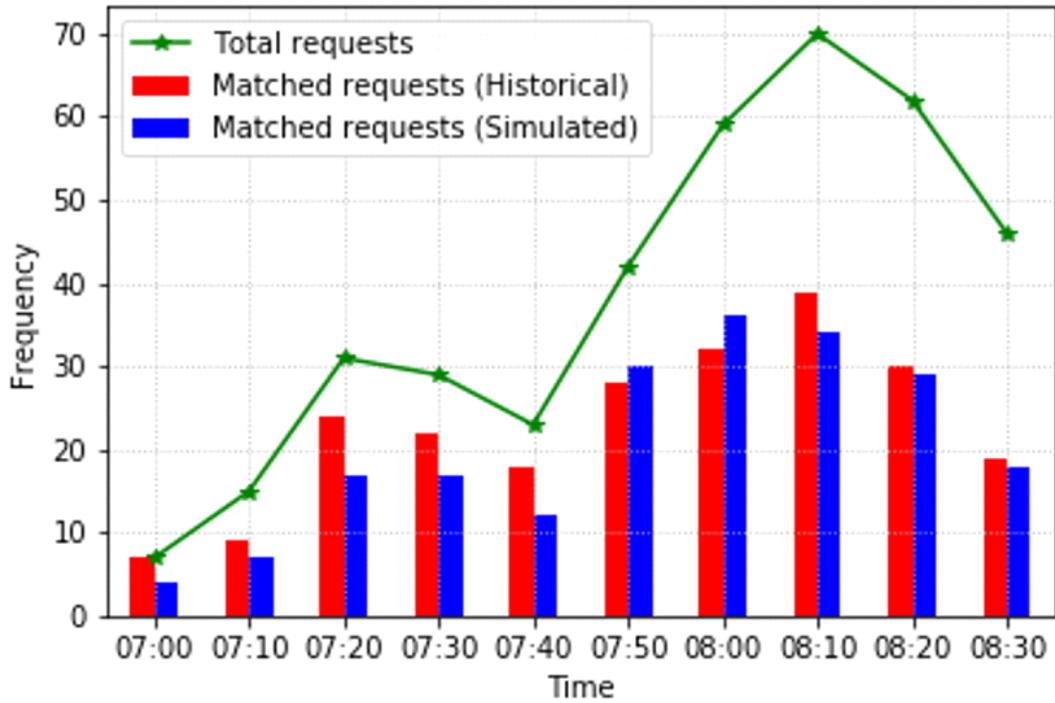


Figure 16. Matching rate by time

Table 10. Matching rate and error rate by time

Time	Requests	Matched requests	Matching_rate (A)	Sim_matched requests	Sim_matching_rate (B)	Error_rate (B-A) (%)
7:00 ~ 7:10	7	7	100.0	4	57.1	-42.9
7:10 ~ 7:20	15	9	60.0	6	40.0	-20.0
7:20 ~ 7:30	31	24	77.4	18	58.1	-19.3
7:30 ~ 7:40	29	22	75.9	17	58.6	-17.2
7:40 ~ 7:50	31	19	61.3	17	54.8	-6.5
7:50 ~ 8:00	48	33	68.8	31	64.6	-4.2
8:00 ~ 8:10	64	34	53.1	38	59.4	6.3
8:10 ~ 8:20	73	40	54.8	36	49.3	-5.5
8:20 ~ 8:30	64	31	48.4	30	46.9	-1.6
8:30 ~ 8:35	48	19	39.6	19	39.6	0.0

The distribution of surge multipliers is presented in Table 11. While the majority of requests have been assigned surge multipliers above 1, instances with values below 1 also exist.

Table 11. Distribution of surge multipliers (Base model)

Surge multipliers	Frequency	Unmatched requests	Matched requests	Matching rate (%)
0.8	6	1	5	83.3
0.9	4	1	3	75.0
1.0	33	18	15	45.5
1.1	33	16	17	51.5
1.2	79	38	41	51.9
1.3	255	120	135	52.9

There is a small number of requests with surge multipliers below 1. These requests are predominantly characterized by relatively short distances traveled, as they cover less than 7 km, which contrasts with the maximum travel distance of 20 km observed across all requests. Notably, these specific requests have destinations indicated by blue circles in Figure 17, primarily located within high-demand areas. Furthermore, a noteworthy observation is that certain vehicles, having fulfilled these requests, subsequently arrived at the grid marked in red and proceeded to accept the subsequent request.

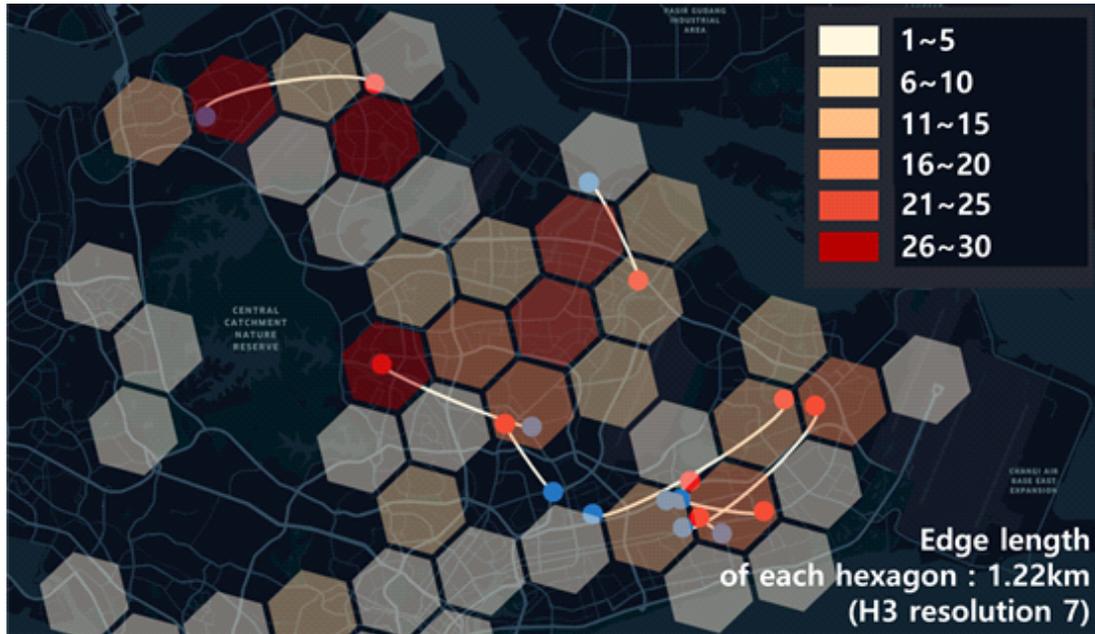


Figure 17. Spatial distribution of origins and requests with surge multipliers less than 1

The matching rate based on the value term is explored in Figure 18. The anticipation was that as the value term increases, both the reward and matching rate would also increase. While the relationship isn't strictly monotonically increasing, as indicated in Table 12, the matching rate tends to be relatively high when considering higher value terms. Interestingly, certain requests are still matched even when the value term holds a negative value. These instances predominantly involve long-distance requests that are economically viable and lucrative enough to offset the negative value term.

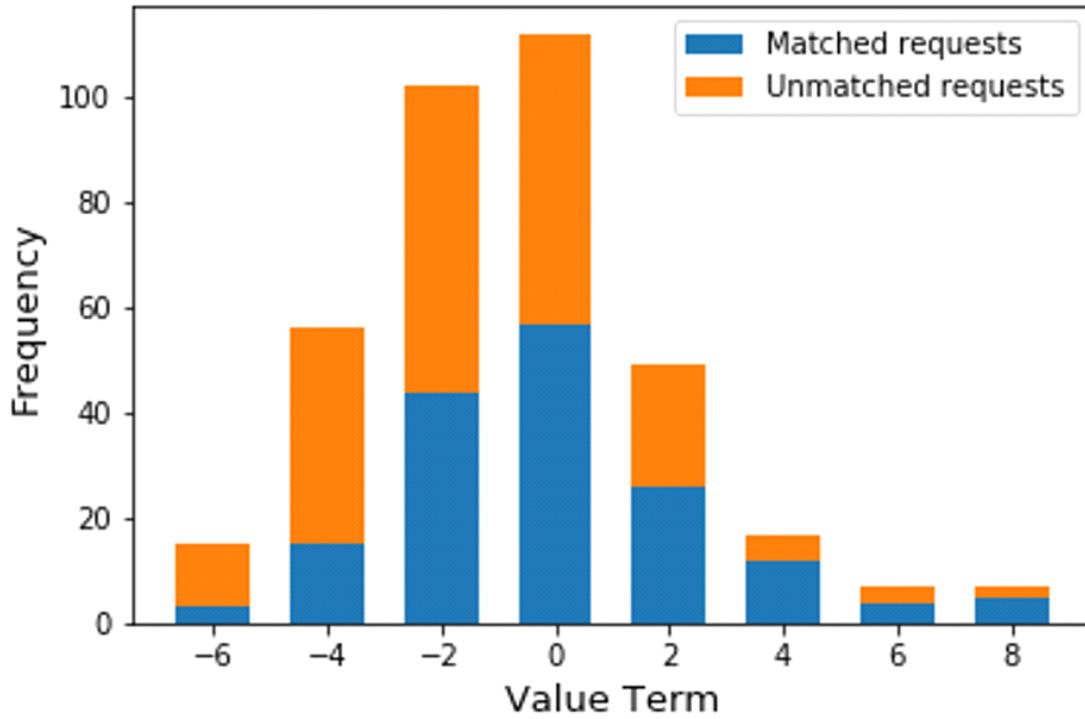


Figure 18. Matching rate according to value term

Table 12. Value term and matching rate in base model

Value Term	Requests	Unmatched requests	Matched requests	Matching rate (%)
-7~-5	15	12	3	20
-5~-3	56	41	15	26.8
-3~-1	102	58	44	43.1
-1~1	112	55	57	50.9
1~3	49	23	26	53.1
3~5	17	5	12	70.6
5~7	7	3	4	57.1
7~9	7	2	5	71.4

5.2. Pickup Waiting Time Reducing Model

To construct the pickup waiting time reducing model, a term indicating cost of pickup waiting time is added (Equation (9)).

$$\begin{aligned}
 & v_{\pi}(x_i, a_i) \\
 &= \sum_j f(x_i, a_i) b_{ji} \left[p_i a_i + \gamma \{V_{\pi}(l_i', t_j + T_i) - V_{\pi}(l_j, t_j)\} \right. \\
 & \quad \left. - \frac{d_{pick,i}}{e} p_{oil} r \right] \tag{9}
 \end{aligned}$$

= Immediate reward (revenue term) + Future reward (value term) - Cost of pickup waiting time (pickup cost)

As discussed in the introduction, platforms bear the responsibility of reducing carbon footprint. Emissions during the time when riders are being transported and generating revenue cannot be reduced. However, platforms can address this issue by adjusting matching and pricing to reduce the pickup waiting time, which accounts for a significant portion of the overall waiting time. In this model, cost of pickup waiting time is subtracted from the reward to get a matching and pricing result which reduce pickup waiting time. Pickup waiting time cost is expressed as emission cost during pickup.

To compare surge multiplier distribution of pickup waiting time reducing model with base model, Figure 19 shows that the

distribution of surge multipliers appears similar. It is judged that the dispatch was made in the direction of reducing pickup time compared to the basic model.

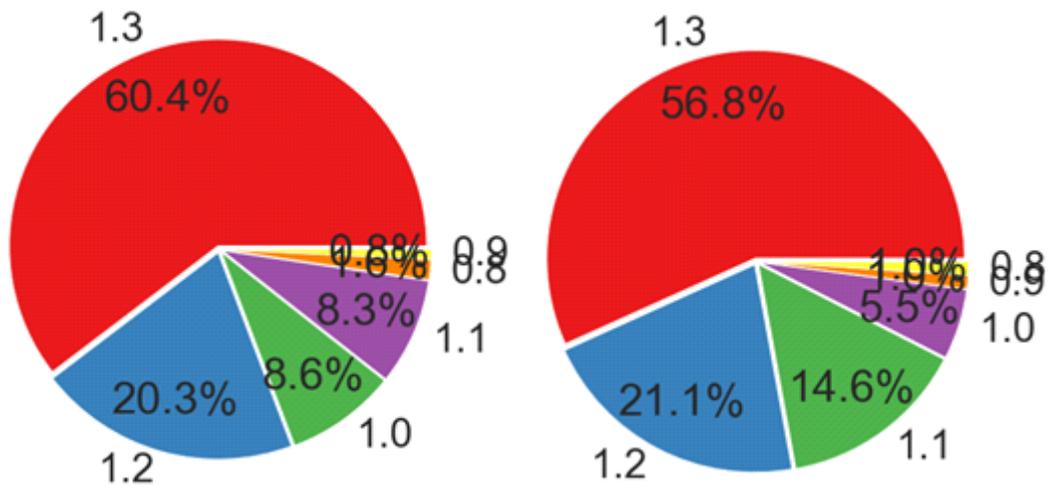


Figure 19. Comparison of surge multiplier distribution with base model (Left)

To investigate if pickup waiting time is reduced when pickup cost is reflected, items of Table 13 are listed. Note that for this part, models take temporal scope from 7:00 to 8:30. Pickup time is reduced by 15%, operating costs are expected to be reduced, and revenue reduction is only 2%. 20.4kg of carbon dioxide emission can be reduced due to reduced pickup time of 10,200 seconds. For reference, an average person breathes out around 1kg of CO₂ in one day. Matched riders experience an average of 46 seconds of waiting time savings, which is 15% less than the base model.

Table 13. Comparison of base model and pickup cost reducing model

* Result of 7:00 ~ 8:30		Base model	Pickup cost reducing model
Matching rate		204/384	201/384
Matched drivers		161	153
Total (In-service) drivers		1,205	1,206
Revenue (SGD)		2,571	2,505
Total pickup waiting time (sec)		69,300	59,100
Pickup waiting time (min)	Mean	5.67	4.90
	STD	4.23	4.32

Table 14 compares pickup waiting time reducing model to TADA' s pricing. Developed models show higher average revenue than TADA' s historical data. Due to variations in matching rates, the revenue per matched request was compared. Pickup waiting time reducing model only resulted in a decrease of 0.14 SGD per request compared to base model.

Table 14. Comparison to TADA's pricing

	TADA' s pricing	Base model	Pickup cost reducing model
Matching rate	228/384	204/384	201/384
Revenue (SGD)	2,601	2,571	2,505
Revenue per matched request (SGD)	11.41	12.60	12.46

Applying the pickup waiting time reducing model benefits platform, rider and driver, and society. For platform, 20kg of carbon emissions during pickup is reduced while maintaining a similar matching rate and revenue. Also, requests can be served with fewer vehicles (161→153), although further analysis will be required. For rider, waiting time (5.67→4.90sec) per request before boarding is reduced by 15% compared to the base model. For driver, fuel is saved during non-profitable travel (The percentage saved depends on the total non profitable travel time). For society, by reducing carbon emissions by 20kg compared to the base model, it helps alleviate negative externalities.

5.3. Effect of Changes in Price Sensitivity

The objective of this section is to examine how alterations in price sensitivity for both stakeholders (rider and driver) influence the matching rate concerning travel distance and time. From Table 15, compared to the historical data (with a matching rate of 228/384, as seen in Table 14), the matching rate is most similar at $\alpha = 1.45$, but matching rate of longer travel distance is relatively high. Considering the matching rate error by travel distance, a value of $\alpha = 1.55$ aligns most closely with the historical data.

The impact of adjustments in price sensitivity for both riders and drivers on the matching rate based on time is consolidated in Table 16. As for the upper two cases in Table 16, the matching rates surpass those of historical data during periods of high demand, such as around 8:10. A relatively lower sensitivity prompts both sides to readily accept the price, consequently contributing to an elevated matching rate. Evaluating the matching rate error with respect to time, a value of $\alpha = 1.55$ demonstrates the closest resemblance to the historical data.

Table 15. Matching rate according to travel distance with changing price sensitivity

	$\alpha = 1.40$				$\alpha = 1.45$			
Portion of Surge Multipliers	0.8	1.82%	0.9	–	0.8	–	0.9	–
	1.0	3.12%	1.1	5.19%	1.0	4.68%	1.1	3.38%
	1.2	27.53%	1.3	62.34%	1.2	14.55%	1.3	77.40%
Matching rate	298/384				227/384			
	$\alpha = 1.50$				$\alpha = 1.55$			
Portion of Surge Multipliers	0.8	1.04%	0.9	2.60%	0.8	1.56%	0.9	1.04%
	1.0	3.38%	1.1	5.97%	1.0	8.57%	1.1	8.31%
	1.2	19.74%	1.3	67.27%	1.2	20.26%	1.3	60.26%
Matching rate	177/384				204/384			

Table 16. Matching rate according to time with changing price sensitivity

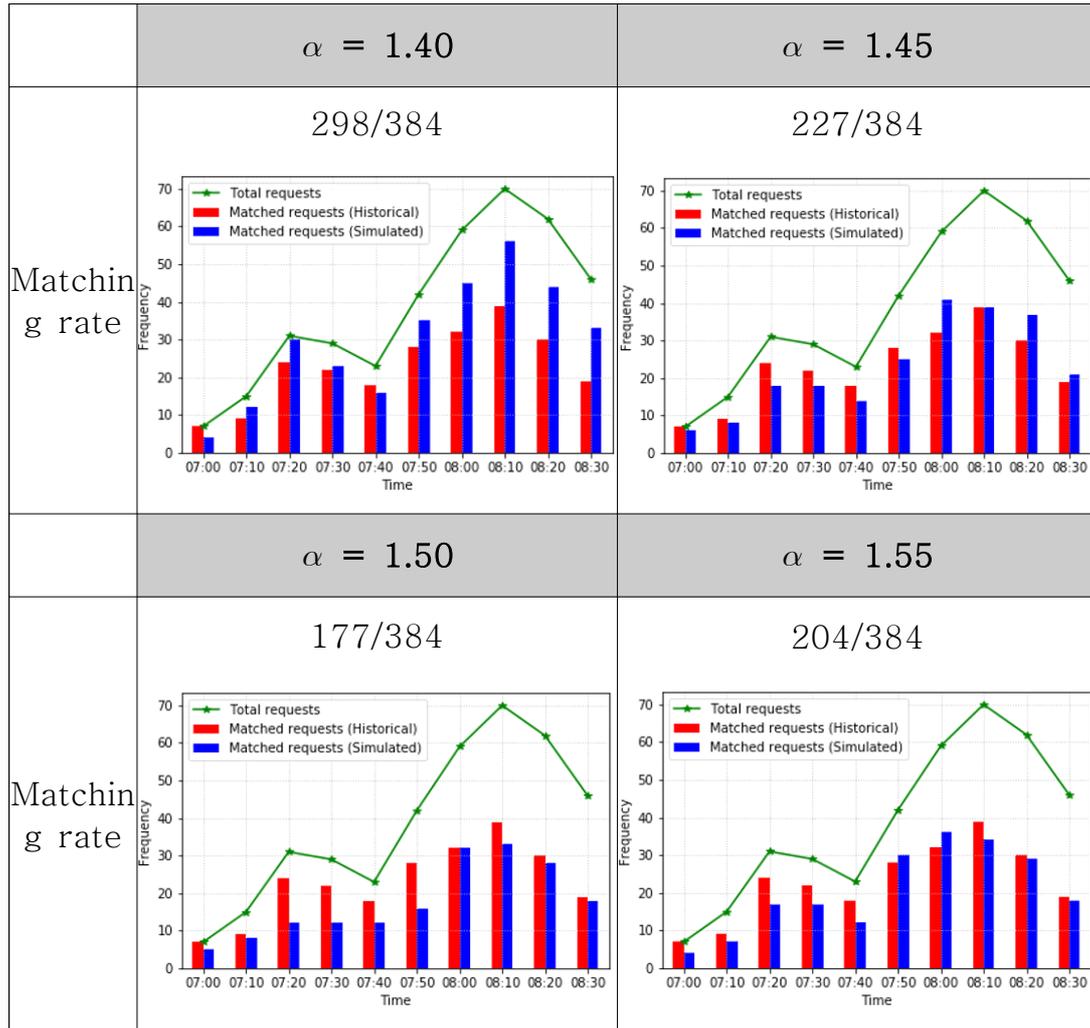


Table 17 shows revenue and reward according to price sensitivity.

Table 17. Revenue and reward with changing price sensitivity

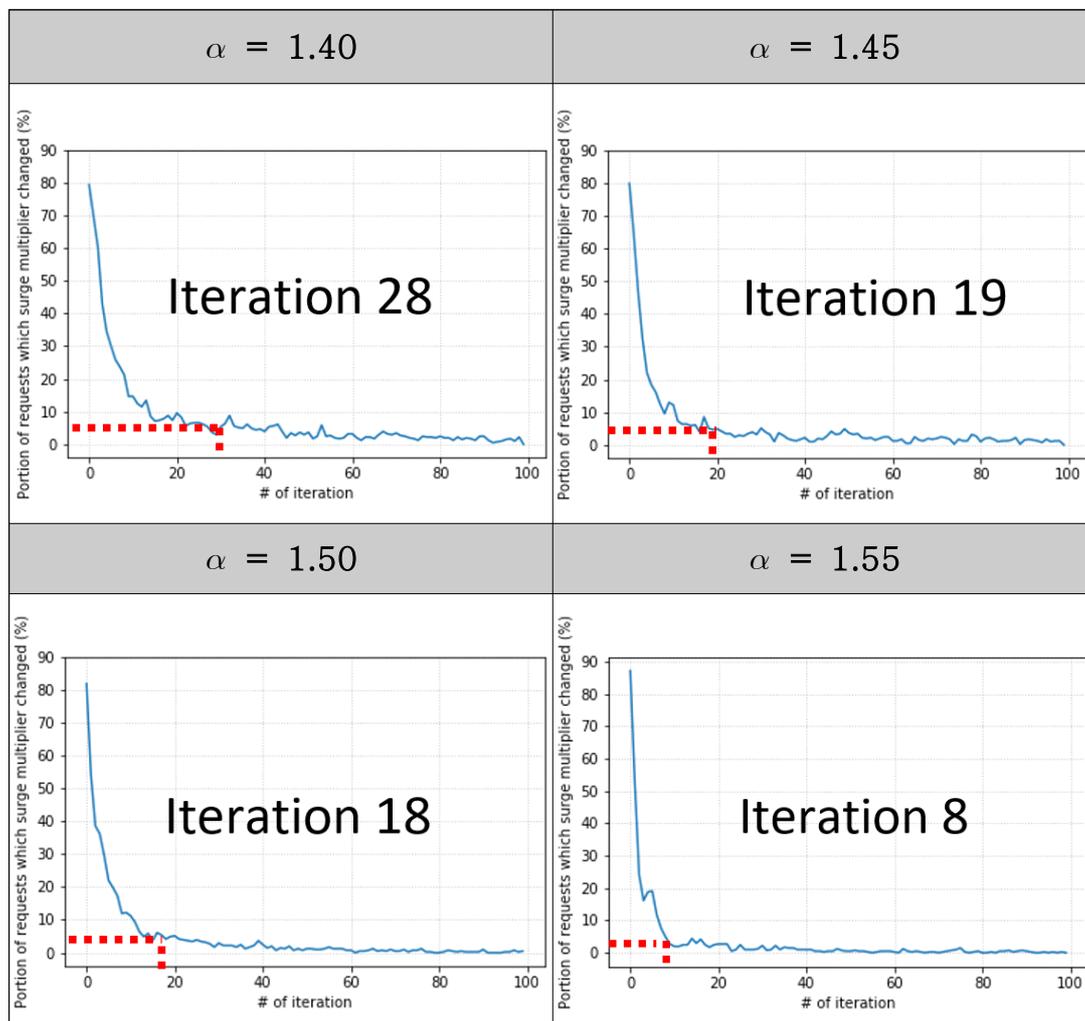
$\alpha = 1.40$ (298/384)			$\alpha = 1.45$ (227/384)		
	Total	Per matched request		Total	Per matched request
Revenue (SGD)	4043.06	13.57	Revenue (SGD)	3185.56	14.03
Reward (SGD)	402.12	1.35	Reward (SGD)	225.85	0.99
$\alpha = 1.50$ (177/384)			$\alpha = 1.55$ (204/384)		
	Total	Per matched request		Total	Per matched request
Revenue (SGD)	2361.45	13.34	Revenue (SGD)	2570.67	12.60
Reward (SGD)	170.85	0.97	Reward (SGD)	175.64	0.86

The upper two cases in Table 17 exhibit a relatively higher matching rate for requests with longer distances, consequently resulting in a heightened revenue per matched request when compared to the two cases listed below. The lowest revenue per matched request occurs at $\alpha = 1.55$, yet this value remains higher than the historical data, which reported a total revenue of 2309.45 SGD and a revenue per matched request of 6.77 SGD.

Furthermore, Table 18 presents a comparison of the number of

iterations required for the algorithm to converge. The criteria for convergence entail the occurrence of two consecutive percentages below 5%. Notably, a higher α value leads to swifter convergence. The last case, which most accurately simulates the actual data, demonstrates the fastest convergence.

Table 18. Convergence of algorithm with changing price sensitivity



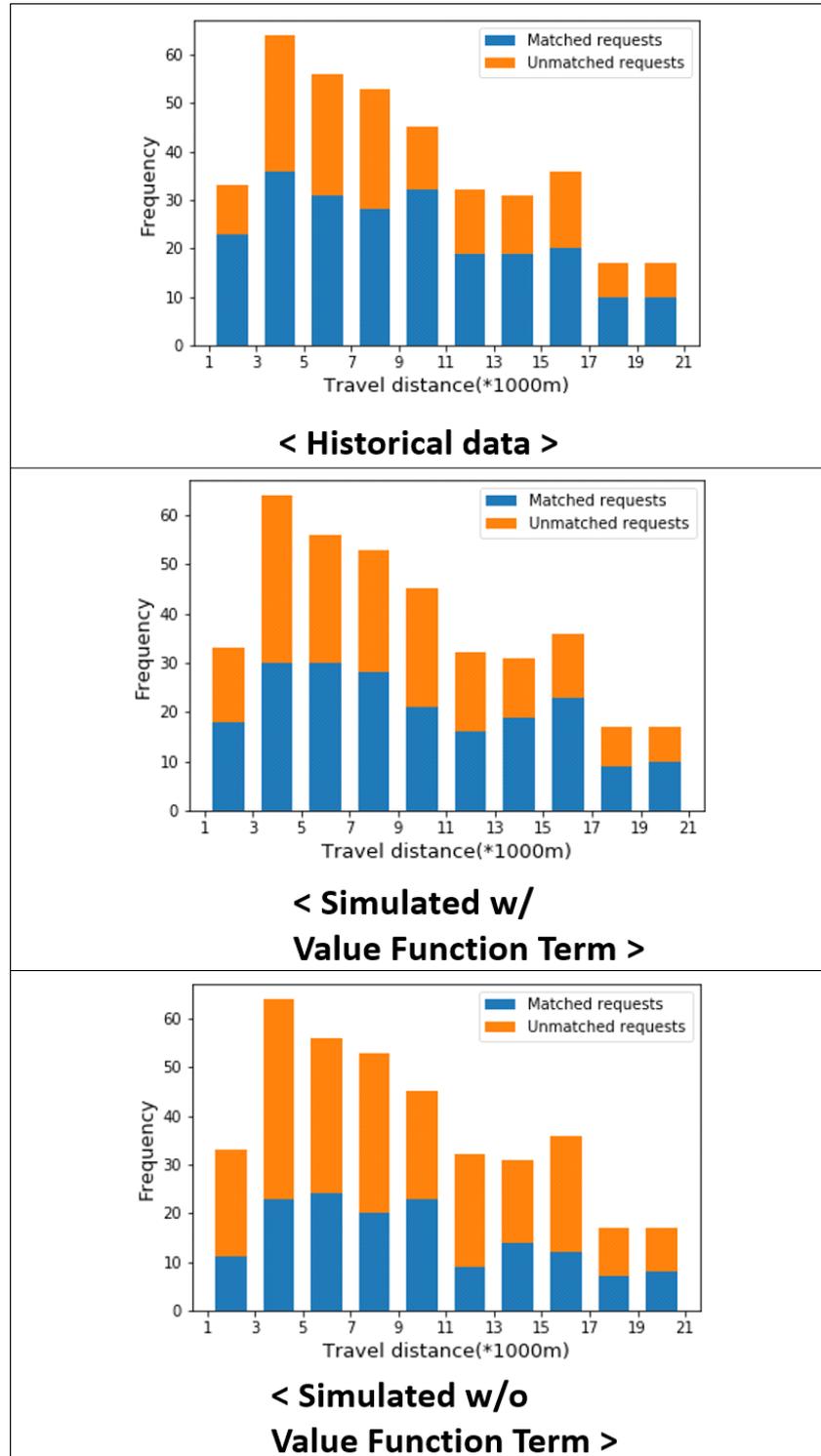
5.4. Effect of Changes in Value Term Coefficient

In order to investigate the impact of changes in the value term coefficient β , the study initially applies $\beta=0$. This adjustment can be conceptualized as "excluding potential requests." The matching rates corresponding to historical data ($\beta=20$), and $\beta=0$ are plotted in Table 19 based on travel distance.

In historical data, the matching rate for long-distance requests is comparatively lower than that for short-distance requests. However, within the base model, the matching rate for long-distance requests exhibits an increase compared to historical data. Notably, upon removal of the value term, the matching rate for long-distance requests regresses to a level lower than that observed in historical data, leading to a decline from the base model's matching rate of 204/384 to 151/384.

The presence of the value term in the reward function signifies the significance (potential) of the destination. The removal of the value term diminishes the allure of long-distance requests, as the potential requests associated with these requests is no longer factored in. Consequently, these potential requests are not considered, contributing to the observed decrease in matching rate.

Table 19. Matching rate according to travel distance with changing value term coefficient



Comparison of surge multiplier distribution between $\beta=20$, and $\beta=0$ is in Figure 20. The portion of the maximum rate decreased by 8%p as value term vanishes.

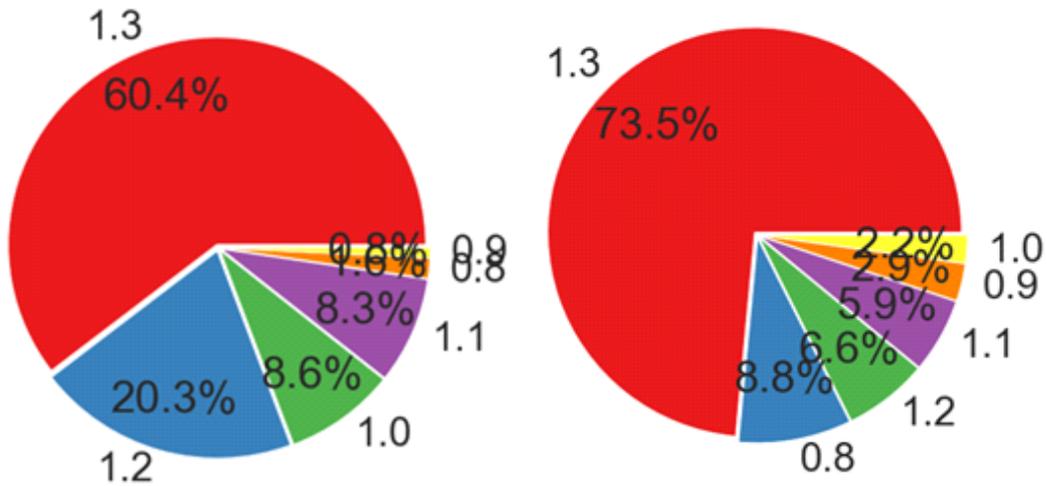


Figure 20. Comparison of surge multiplier distribution with base model (Left)

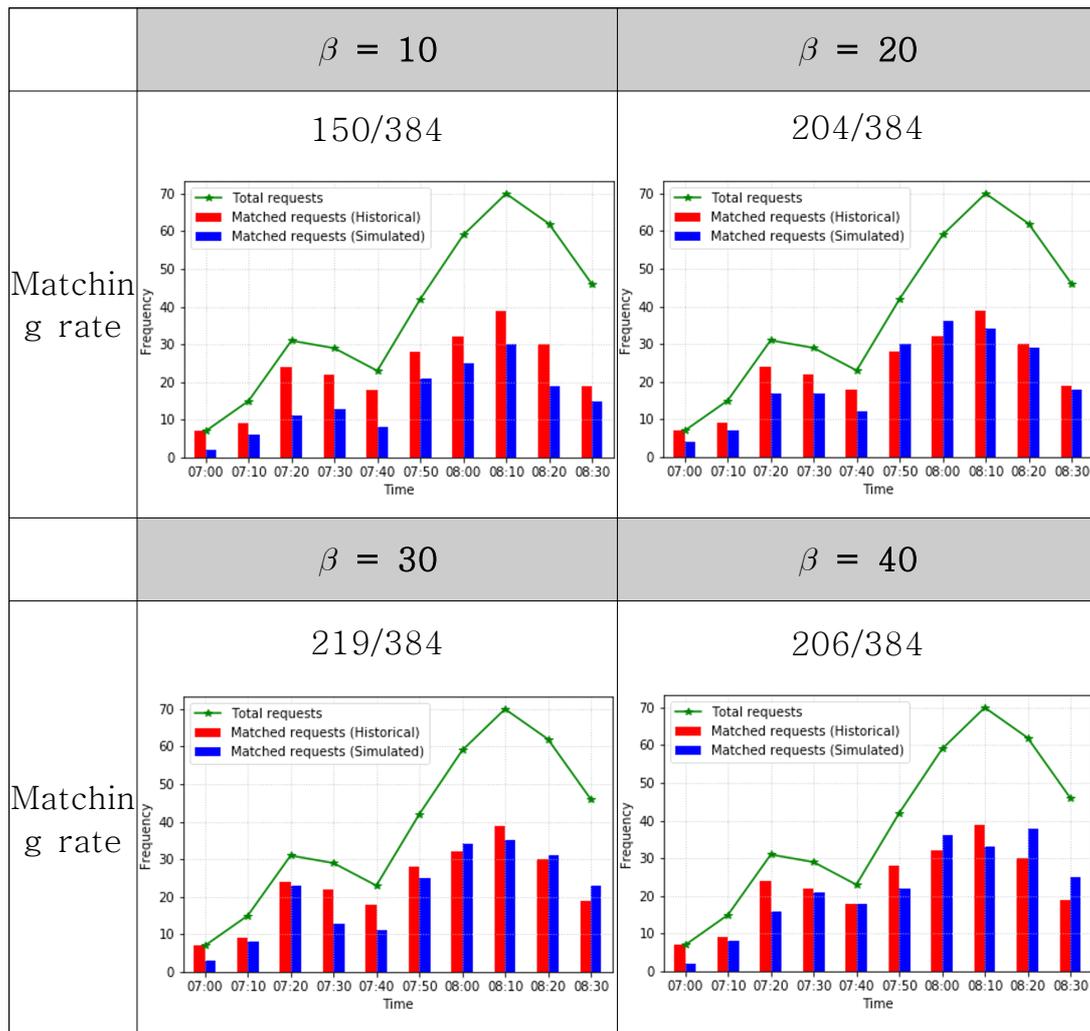
Next, potential requests are considered at base model with varying β . Table 20 shows the effect of changes in value term coefficient on the matching rate according to travel distance. Compared to the historical data (228/384), the matching rate is most similar at $\beta=30$, but error rate is greater. Considering the matching rate error by travel distance, $\beta=20$ case is most similar with historical data.

Table 20. Matching rate according to travel distance with changing value term coefficient

	$\beta = 10$				$\beta = 20$			
Portion of Surge Multipliers	0.8	14.84%	0.9	3.39%	0.8	1.56%	0.9	1.04%
	1.0	2.60%	1.1	5.73%	1.0	8.57%	1.1	8.31%
	1.2	14.84%	1.3	58.59%	1.2	20.26%	1.3	60.26%
Matching rate	<p>150/384</p>				<p>204/384</p>			
	$\beta = 30$				$\beta = 40$			
Portion of Surge Multipliers	0.8	0.52%	0.9	5.99%	0.8	0.52%	0.9	2.86%
	1.0	8.07%	1.1	12.50%	1.0	3.91%	1.1	14.58%
	1.2	20.83%	1.3	52.08%	1.2	20.05%	1.3	58.07%
Matching rate	<p>219/384</p>				<p>206/384</p>			

The effect of changes in value term coefficient on the matching rate according to travel distance is summarized in Table 21.

Table 21. Matching rate according to time with changing value term coefficient



In contrast to α , β did not exhibit a discernible temporal trend in terms of its magnitude. Matching rates demonstrated greater similarity at $\beta = 30$ or 40 when compared to the $\beta = 20$ case,

albeit with a higher error rate. Evaluating the matching rate error with respect to time, the $\beta = 20$ case exhibits the closest resemblance to historical data.

Table 22 presents a comparison of revenue and reward according to β .

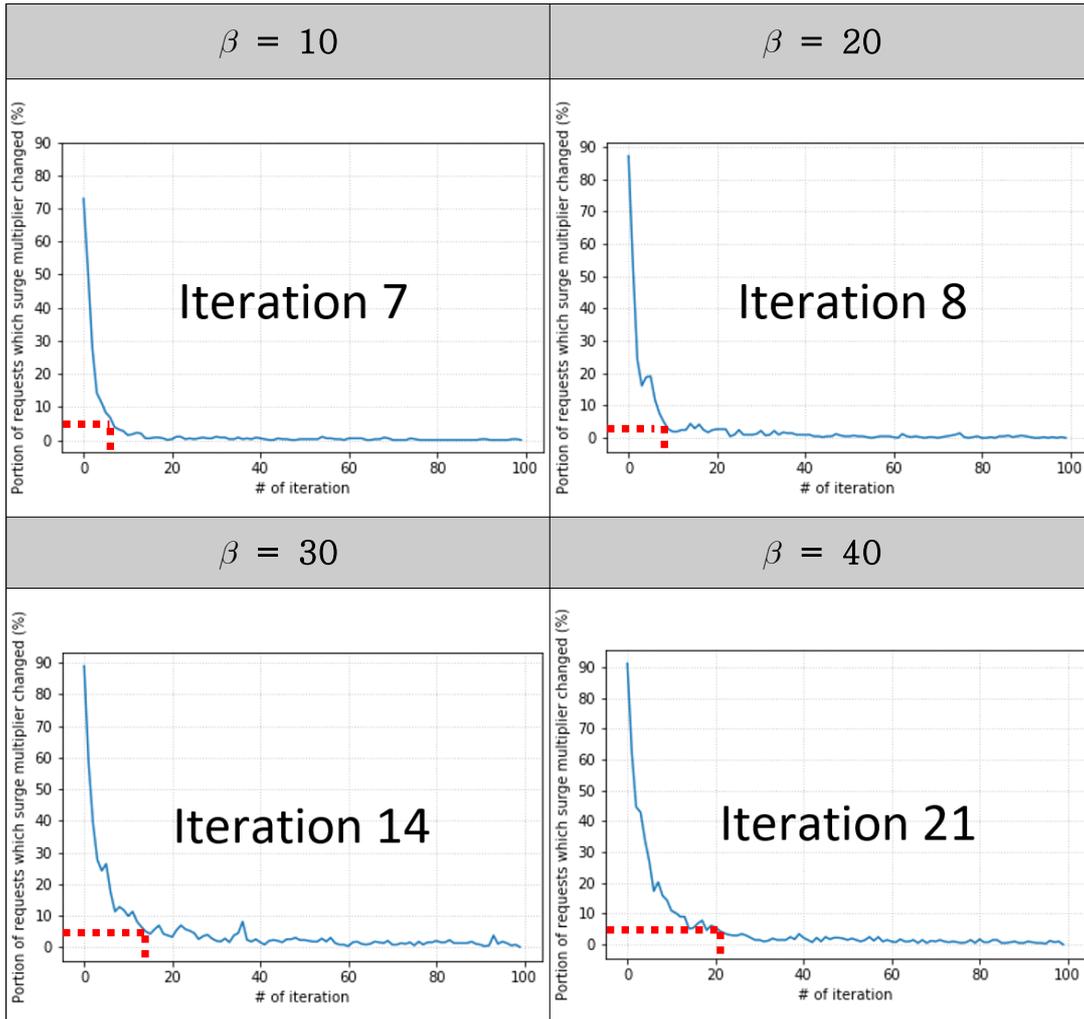
Table 22. Revenue and reward with changing price sensitivity

$\beta = 10$ (150/384)			$\beta = 20$ (204/384)		
	Total	Per matched request		Total	Per matched request
Revenue (SGD)	1765.36	11.77	Revenue (SGD)	2570.67	12.60
Reward (SGD)	133.30	0.89	Reward (SGD)	175.64	0.86
$\beta = 30$ (219/384)			$\beta = 40$ (206/384)		
	Total	Per matched request		Total	Per matched request
Revenue (SGD)	2607.16	11.90	Revenue (SGD)	2691.98	13.07
Reward (SGD)	184.99	0.84	Reward (SGD)	222.54	1.08

The $\beta = 30$ case exhibited a relatively higher matching rate for requests with shorter distances, consequently leading to a decreased revenue per matched request when compared to other cases. On the other hand, while the $\beta = 20$ case did not exhibit the highest revenue per matched request, it still outperformed the historical data, which reported a total revenue of 2,601 SGD and a revenue per matched request of 11.41 SGD.

Moreover, Table 23 offers a comparison of the number of iterations required for algorithm convergence. The convergence criterion involves the presence of two consecutive percentages falling below 5%. Interestingly, higher β values result in slower convergence. This can be attributed to the fact that a larger β value corresponds to a greater value term. Consequently, it can be inferred that the convergence of temporal difference learning is prolonged with larger β values.

Table 23. Convergence of algorithm with changing value term coefficient



5.5. Applicability of the Model

5.5.1. Upper Limit of Surge Multiplier

When the maximum grantable surge multiplier is increased (1.3 → 1.9), 1.5 or higher does not appear. As depicted in the conversion rate function (Equation (8)), an excessively high surge multiplier results in rider's rejection and a decrease in the reward, consequently leading to a determination that the request will not be matched. In reality, even without the imposition of a price ceiling, the market tends to establish an appropriate surge multiplier. Therefore, the model reflects reality well and holds the potential for application in other cities of similar scale to Singapore.

5.5.2. Applicability to Off-peak Hours

To test the applicability in urban areas, the model was applied not only to peak hours but also to off-peak hours. Applicability in off-peak hours (14:00~15:00) was tested. When the matching rate of historical data was 130/220, the simulation result was 122/220. Figure 21 and Table 24 shows temporal matching pattern, which is similar to historical data.

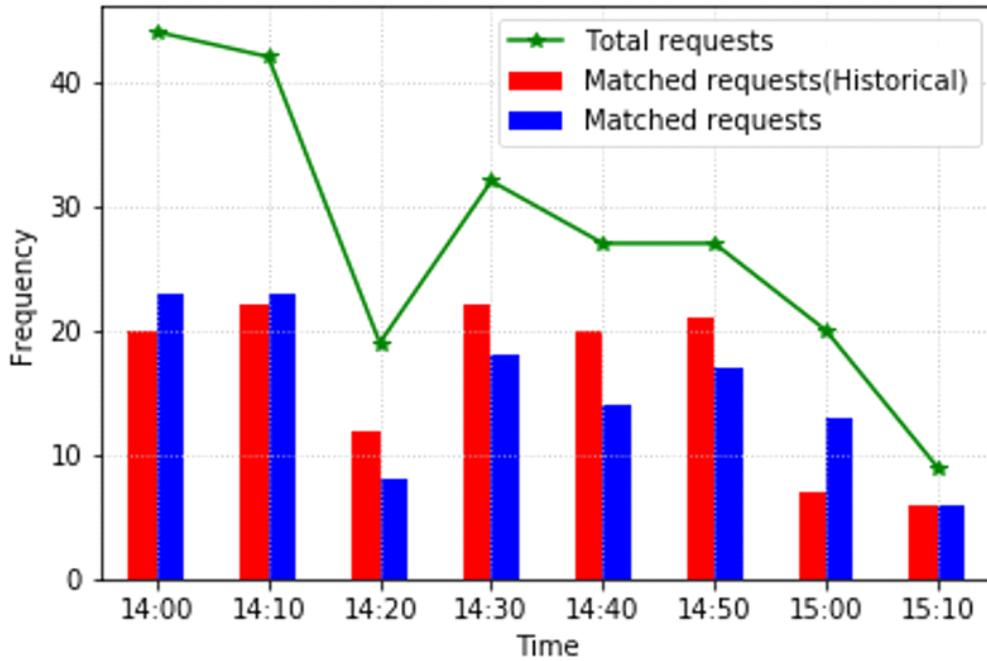


Figure 21. Temporal matching rate of off-peak

Table 24. Matching rate by travel distance (off-peak)

Travel distance (km)	Requests	Matched requests	Matching rate (A)	Sim_matched requests	Sim_matching_rate (B)	Error_rate (B-A) (%)
1~3	27	18	66.7	14	51.9	-14.8
3~5	75	31	41.3	39	52.0	10.7
5~7	35	18	51.4	17	48.6	-2.9
7~9	16	11	68.8	12	75.0	6.3
9~11	13	11	84.6	10	76.9	-7.7
11~13	13	10	76.9	6	46.2	-30.8
13~15	17	11	64.7	8	47.1	-17.6
15~17	8	8	100.0	5	62.5	-37.5
17~19	8	5	62.5	3	37.5	-25.0
19~21	7	4	57.1	5	71.4	14.3
Total			58.0	-	54.3	-3.6

Pickup waiting time reducing model is applied. From Table 25, Total pickup time is reduced by 27%, operating costs are expected to be reduced, and revenue reduction is only 3%. Matched riders experience an average of 87 seconds of waiting time savings, which is 25% less than the base model. Large reduction in pickup waiting time (27%) compared to the morning peak (15%) is observed. During off-peak hours, the number of vehicles is small due to low demand, but compared to the morning peak, supply is less than demand (Morning peak $204/161=1.27$ requests/vehicle (See 5.2.), Off-peak $122/89=1.37$ requests/vehicle).

Table 25. Comparison of base model and pickup cost reducing model (Off-peak hours)

		Base model	Pickup cost reducing model
Matching rate		122/220	119/220
Matched drivers		89	85
Revenue (SGD)		1,265	1,223
Total pickup waiting time (sec)		42,300	30,900
Pickup waiting time (min)	Mean	5.78	4.33
	STD	4.02	4.22

5.5.3. Formulate Policies to Reduce Pickup Cost

As a third and final part concerning the applicability of the model, potential to formulate policies to reduce pickup cost is addressed. The model can help formulate policies by comparing base model and pickup cost reducing model. Figure 22 shows the gap of average surge multiplier (Pickup cost reducing model – Base model). Regions with a deep red color are relatively concentrated in the south. This contrasts with the typical pattern, wherein requests frequently originate from residential zones in the north and are directed towards business areas in the south. Thus, setting prices in a manner that mitigates the imbalance in vehicle numbers between regions, which could arise due to the handling of disparate demand, becomes a viable policy for the platform company.

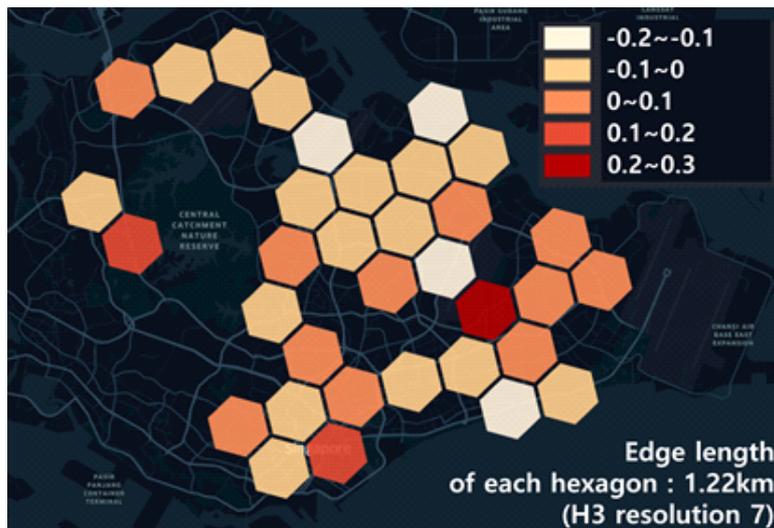


Figure 22. Surge multiplier gap between base model and pickup cost reducing model

Chapter 6. Conclusion

6.1. Summary

This study formulated a model for matching and pricing in a ride-hailing service to reduce pickup waiting time. The model decided surge multipliers and dispatching using contextual bandit and temporal difference learning. For the contextual bandit, a request was characterized by a 5-dimensional vector of contextual features. When the value term is eliminated, the potential requests that a driver could receive are not considered, leading to a decrease in the matching rate. When pickup cost is reduced, simulations during the morning peak hour in Singapore showed a reduction of 20 kilograms of carbon dioxide emissions and a 15% decrease in pickup waiting time with only a 2% revenue reduction.

The model's applicability to metropolitan areas as big as Singapore was also explored. In the reward function, the revenue term and the value term are large and small, respectively. However, when the value term was multiplied by the coefficient, it showed a matching rate similar to that of historical data. The actual surge multiplier had an upper limit even if larger surge multiplier candidates were included in the model. It demonstrated similar patterns to the historical matching rate not only during peak hours but also during off-peak hours.

6.2. Limitations and Future Research

The model shows a matching rate similar to the historical data but assumes the price sensitivity of riders (passengers) and drivers as a linear function. Another limitation is that this algorithm is not applied in a real-time situation.

For further research, a simulation environment that reflects competition with other companies can be developed. Some countries have a duopoly ride-hailing market, so studying this kind of market is a meaningful endeavor. From this limitation, representing the price sensitivity of riders and drivers in a form other than a linear function is also feasible.

References

- Agarwal, S., Mani, D., & Telang, R. (2019). The Impact of Ride-hailing Services on Congestion: Evidence from Indian Cities. Available at SSRN 341062.
- Agatz, N., Erera, A., Savelsbergh, M., & Wang, X. (2012). Optimization for dynamic ride-sharing: A review. *European Journal of Operational Research*, 223(2), 295–303. <https://doi.org/10.1016/j.ejor.2012.05.028>
- Alagoz, O., Hsu, H., Schaefer, A. J., & Roberts, M. S. (2010). Markov decision processes: a tool for sequential decision making under uncertainty. *Medical Decision Making*, 30(4), 474–483.
- Bailey, N. (2022). Dynamic Ridesharing under Travel Time Uncertainty: Passenger Preference and Optimal Assignment Methods. PhD thesis, MIT.
- Barnes, S.J., Guo, Y., Borgo, R., (2020). Sharing the air: Transient impacts of ride-hailing introduction on pollution in China. *Transportation Research Part D: Transp. Environ.* 86, 102434.

- Beojone, C.V., Geroliminis, N., (2021). On the inefficiency of ride-sourcing services towards urban congestion. *Transp. Res. Ck* 124, 102890.
- Bertsimas, D., Perakis, G. (2006). Dynamic Pricing: A Learning Approach. *Mathematical and Computational Models for Congestion Charging*, 45–80.
- Bimpikis, K., Candogan, O., & Saban, D. (2019). Spatial pricing in ride-sharing networks. *Operations Research* . In Press <https://www.gsb.stanford.edu/faculty-research/working-papers/spatial-pricing-ride-sharing-networks>.
- Burq, M. (2019). *Dynamic Matching Algorithms*. PhD thesis, MIT.
- Chen, H., Jiao, Y., Qin, Z., Tang, X., Li, H., An, B., Zhu, H., Ye, J. (2019). InBEDE: Integrating contextual bandit with td learning for joint pricing and dispatch of ride-hailing platforms. *Proc. – IEEE Int. Conf. Data Mining, ICDM 2019–November*, 61–70. <https://doi.org/10.1109/ICDM.2019.00016>
- Feng, G., Guangwen K., Zizhuo W. (2021). We are on the way: Analysis of on-demand ride-hailing systems. *Manufacturing*

& Service Operations Management, 23(5), 1237–1256.

Hall, J., Cory K., Chris N. (2015). The effects of Uber's surge pricing: A case study. The University of Chicago Booth School of Business URL http://1gluem2nc4jy1gzhn943ro0gz50.wpengine.netdna-cdn.com/wp-content/uploads/2016/01/effects_of_uber_surge_pricing.pdf.

He, S., & Shin, K. G. (2019). Spatio-temporal adaptive pricing for balancing mobility-on-demand networks. ACM Transactions on Intelligent Systems and Technology, 10(4). <https://doi.org/10.1145/3331450>

Jang, K., Song, M.K., Choi, K., Kim, D.K. (2014). A bi-level framework for pricing of high-occupancy toll lanes. Transport, 29, 317–325. <https://doi.org/10.3846/16484142.2014.952248>

Jehl, T. (2020). Data-Driven Decision Making Algorithms For Internet Platforms. PhD thesis, University of California, Berkeley

Kang, S., & Kim, T. J. (2009). Dynamic programming approach for a concierge service problem in location-based services. 88th

Annual Meeting of Transportation Research Board, January.

Kang, S., & Ouyang, Y. (2011). The traveling purchaser problem with stochastic prices: Exact and approximate algorithms. *European Journal of Operational Research*, 209(3), 265–272. <https://doi.org/10.1016/j.ejor.2010.09.012>

Korolko, N., Yan, C., Zhu, H., & Woodard, D. (2019). Dynamic pricing and matching in ride-hailing platforms. *Naval Research Logistics*. <https://doi.org/10.1002/nav.21872>

Legros, B. (2019). Dynamic repositioning strategy in a bike sharing system; how to prioritize and how to rebalance a bike station. *European Journal of Operational Research*, 272(2), 740–753.

Li, L., Chu, W., Langford, J., & Schapire, R. E. (2010). A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World Wide web*, 661–670.

Lyft. (2016). Matchmaking in Lyft line: Part 3. URL <https://eng.lyft.com/matchmaking-in-lyft-line-part-3-d8f9497c0e51>.

- Meghjani, M., & Marczuk, K. (2017). A hybrid approach to matching taxis and customers. *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, 167–169. <https://doi.org/10.1109/TENCON.2016.7847982>
- Ozkan, E., & Ward, AR. (2016). Dynamic matching for real-time ridesharing. Working Paper URL <https://ssrn.com/abstract=2844451>.
- Qin, Z. (Tony), Tang, X., Jiao, Y., Zhang, F., Xu, Z., Zhu, H., & Ye, J. (2020). Ride-Hailing Order Dispatching at DiDi via Reinforcement Learning. *INFORMS J. Appl. Anal.* 50, 272–286. <https://doi.org/10.1287/inte.2020.1047>.
- Rong, H., Zhou, X., Yang, C., Shafiq, Z., Liu, A. (2016). The rich and the poor: a Markov decision process approach to optimizing taxi driver revenue efficiency. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. ACM, New York, NY, USA, pp. 2329–2334.
- Shah, S., Lowalekar, M., Varakantham, P. (2022). Joint Pricing and Matching for City-Scale Ride-Pooling. *Proceedings of the International Conference on Automated Planning and Scheduling*, 32(1), 499–507.

- Shen, Q. (2021). Seeking for Passenger under Dynamic Prices: A Markov Decision Process Approach. *Journal of Computer and Communications*, 9, 80–97.
- Shou, Z. (2021). Harnessing Big Data for the Sharing Economy in Smart Cities. PhD thesis, Columbia University.
- Sun, L., Teunter, R. H., Babai, M. Z., & Hua, G. (2019). Optimal pricing for ride-sourcing platforms. *European Journal of Operational Research*, 278(3), 783–795. <https://doi.org/10.1016/j.ejor.2019.04.044>
- Tirachini, A., & Gomez-Lobo, A. (2020). Does ride-hailing increase or decrease vehicle kilometers traveled (VKT)? A simulation approach for Santiago de Chile. *Int. J. Sustain. Transp.*, 14, 187–204.
- Wang, H., & Yang, H. (2019). Ridesourcing systems: A framework and review. *Transportation Research Part B: Methodological*, 129, 122–155.
- Wang, S. (2019). Dynamic Pricing and Demand Shaping: Theory and Applications in Online Assortments, Ride Sharing and Smart Grids. PhD thesis, Columbia University.

Wu, T., Joseph, A., Russell, S. (2016). Automated Pricing Agents in the On-Demand Economy. Technical Report No. UCB/EECS-2016-57.

Zhang, L., Tao H., Yue M., Guobin W., Junying Z., Pengcheng F., Pinghua G., Jieping Ye. (2017). A taxi order dispatch model based on combinatorial optimization. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2151-2159.

국문초록

스마트폰의 대중화와 함께, 승차 공유 서비스는 많은 이용자들의 차량 호출을 운전자에게 중개해주고 있다. 이는 자가용 수요를 대체해 교통혼잡 등 도시문제를 완화할 것으로 기대된다. 그러나 애초 취지와 달리 승차 공유 서비스가 교통 혼잡과 대기오염 등 부정적인 외부효과를 악화시킨다는 비판이 제기되고 있다. 따라서, 승차 공유 서비스에서 이러한 부정적인 외부효과를 저감할 필요가 있다. 픽업 대기 시간이 줄어들면 차량이 이동하는 거리가 줄어들어 부정적 외부 효과를 줄일 수 있으므로, 본 연구는 픽업 대기 시간의 감소에 초점을 맞춘다. 본 연구에서는 픽업 대기 시간의 감소를 위하여 강화 학습을 기반으로 한 매칭 및 요금 책정 알고리즘을 개발하였다. 컨텍스츄얼 밴딧과 시간차 학습을 이용하여 개별 호출에 효율을 부여하고 운전자에게 할당하는 작업을 반복적으로 수행하는 알고리즘이다. 개발된 알고리즘은 싱가포르의 승차 공유 서비스 데이터에 적용되었다. 가격 민감도 변화의 영향과 가치항 계수 변화의 영향도 조사하였다. 첨두시간대 뿐만 아니라 비첨두 시간대에도 실제 매칭률과 유사한 매칭률 패턴을 보여 싱가포르와 같은 대도시에서의 적용 가능성을 연구하였다. 오전 첨두 시간에서의 시뮬레이션 결과, 픽업 대기 시간 단축 모델은 픽업 시간을 15% 단축하고, 이산화탄소 배출량을 20kg 절감하며, 수익은 2% 감소하는 데 그쳤다. 본 연구를 통해 동적 요금과 매칭, 픽업 대기 시간 단축을 함께 다뤄 승차 공유 서비스에서 잠재적인 픽업 대기 시간을 얼마나 절약할 수 있는지를 추정하고, 이용자 편익 측면에서 승차 호출 서비스를 개선하는 데 기여하였다.

주요어 : 승차 공유 서비스, 동적 요금, 부정적 외부효과, 픽업 대기 시간, 강화 학습, 시뮬레이션

학 번 : 2016-21269