Article

An Application of a Deep Q-Network Based Dynamic Fare Bidding System to Improve the Use of Taxi Services during Off-Peak Hours in Seoul

Yunji Cho [1], Jaein Song [2], Minhee Kang [3] and Keeyeon Hwang [4],*

1 National Transport Safety and Disaster Prevention Research Center, Korea Transport Institute, Sejong 30147, Korea; cyj9943@koti.re.kr
2 Research Institute of Science and Technology, Hongik University, Seoul 04066, Korea; wodlsthd@nate.com
3 Department of Smart City, Hongik University, Seoul 04066, Korea; speakingbee@hanmail.net
4 Department of Urban Planning, Hongik University, Seoul 04066, Korea
* Correspondence: keith@hongik.ac.kr

Abstract: The problem of structural imbalance in terms of supply and demand due to changes in traffic patterns by time zone has been continuously raised in the mobility market. In Korea, unlike large overseas cities, the waiting time tolerance increases during the daytime when supply far exceeds demand, resulting in a large loss of operating profit. The purpose of this study is to increase taxi demand and further improve driver's profits through real-time fare discounts during off-peak daytime hours in Seoul, Korea. To this end, we propose a real-time fare bidding system among taxi drivers based on a dynamic pricing scheme and simulate the appropriate fare discount level for each regional time zone. The driver-to-driver fare competition system consists of simulating fare competition based on the multi-agent Deep Q-Network method after developing a fare discount index that reflects the supply and demand level of each region in 25 districts in Seoul. According to the optimal fare discount level analysis in the off-peak hours, the lower the OI Index, which means the level of demand relative to supply, the higher the fare discount rate. In addition, an analysis of drivers' profits and matching rates according to the distance between the origin and destination of each region showed up to 89% and 65% of drivers who actively offered discounts on fares. The results of this study in the future can serve as the foundation of a fare adjustment system for varying demand and supply situations in the Korean mobility market.

1. Introduction

By the fourth Industrial Revolution, the development of mobile and ICT (Information and Communication Technology) opened a new era in platform-based online businesses. In addition, automakers and IT (Information Technology) companies have recently expanded their entry into the mobility market [1]. This led to the provision of the new concept of mobile services that encompass various areas from digital infrastructure to platform services. Korea's mobility market has also launched a variety of mobility services, including car subscription and sharing, beyond the production and sale of vehicles, which were traditional areas of automobile companies. Significantly, the platform-based mobility on demand (MOD) service, which includes ride-hailing and ride-sharing services, has grown significantly. This includes car-hailing and ride-sharing services, and according to the Ministry of Land, Infrastructure and Transport of Korea, the number of brand taxis combined with platforms surpassed 30,000 [2]. This is attracting attention as a way to improve user mobility and accessibility [3].

Meanwhile, in the field of transportation research, structural imbalance between supply and demand has been continuously raised due to changes in traffic patterns by time

zone [4–6]. For example, urban areas in the late-night hours as well as in the peak hours of weekday mornings and evenings, where commuting is essential, are in explosive demand compared to supply. This resulted in increased waiting times for passengers and problems such as refusal of passengers. As a result, this problem has also had a direct adverse effect on the service level. Therefore, the development of supply-inducing measures to deal with excess demand has become a key resolution task for mobility service companies.

As a way to improve this, artificial intelligence and big data-based prediction and optimization techniques have recently been introduced. These provide real-time preference-based services by analyzing data obtained from platform service users. In addition, it has become possible to predict the expected route and arrival time according to real-time market conditions based on GPS (Global Positioning System) information, as well as to present optimal prices. Specifically, dynamic pricing became popular as it was applied to ride-sharing services including Uber and Lyft [7,8]. It is drawing attention as a new fare policy that can increase service usage, as it can yield optimal fare levels for each situation.

In general, ride-sharing services surge fares as a supply inducement during late night hours when demand is far higher than supply. However, there are insufficient measures to induce demand for the use of services during the off-peak hours, for which demand is lower than supply. Especially in Korea's metropolitan area, the decrease in demand for off peak hours is greater than that of overseas large cities [9]. In addition, demand is rapidly decreasing while supply is maintained, which leads to an empty vehicle problem, and the resulting congestion in waiting areas. At the same time, due to the aging of drivers, the working hours will be concentrated during the day, which will worsen the oversupply. Therefore, in the future, it is assumed that these phenomena will lead to increased waiting time and a decrease in profits for empty vehicles. Therefore, it is deemed necessary to study ways to induce demand through fare discounts, which are the expansion of a dynamic pricing strategy during the day.

In addition, the fare discount is also necessary from an environmental perspective. Seoul, the capital of Korea, has one of the highest carbon footprints of 13,000 cities in the world [10], with pollutants that have a significant impact on global warming, with more than 3 million vehicles registered in 2019 [11]. Also looking at the modal share in Seoul, taxis account for 6.7 percent, while passenger cars amount to 20.4 percent [12]. In particular, because the urgency of time is relatively low during off-peak hours, cars are often favored over the expensive taxis. Given this situation, environmental issues are raised because the level of emissions discharged from passenger cars is about 1.7 times higher than that of taxis [12]. Therefore, discouraging car uses by discounting taxi fares can not only help ease road congestion but also lowering air pollutant emission caused by more frequent uses of private cars. It also contributes to reducing $CO2$ emissions, which could be a solution for the transportation sector to the global warming problem facing Seoul.

The goal of this study is to increase not only the demand for taxi uses but driver's revenue through real-time fare discounts in each region of Seoul. Thus, this study aims to derive the discount level of taxi fares for Seoul city by introducing a real-time fare bidding system based on deep reinforcement learning methodology. It can also reduce the number of vehicles waiting empty and the waiting time of drivers, reducing air pollutant emissions.

The procedure of this study is as follows. In Section 2, we conduct theoretical considerations on the application of Dynamic Pricing and the reinforcement learning methodology in the transportation sector. Next, Section 3 presents a methodology for the definition and construction of real-time fare bidding simulation environment. In Section 4, we design a simulator of DQN-based fare competition between drivers. In Section 5, we then compare profit and matching rates among drivers on regional levels of fare discount based on simulation results. For this purpose, three scenarios representing regional characteristics were established and explored. Finally, Section 6 draws conclusions and implications of the study.

2. Literature Review

In this chapter, a fare-adjusting mechanism called dynamic pricing is considered, and we study the application of reinforcement learning (RL), an artificial intelligence-based learning methodology to implement it. Based on the review, we would like to derive the feasibility and differentiation of this study.

2.1. Study on Dynamic Pricing in Transportation Field

Dynamic pricing considers a variety of factors that affect market conditions, such as customer behavior characteristics, level of awareness of products, market structure, and level of product supply and demand. Based on these, dynamic pricing refers to a strategy that varies rates for the same product and service depending on the situation. Prior to the 2000s, a number of studies were conducted in terms of revenue management measures [13,14], but later expanded to various industries and attracted attention as a way to induce various demand matching through optimal rates for each situation [15–18]. Above all, the energy supply sector, which is key to optimizing its supply by dynamically adjusting prices according to consumption, is one of the areas where the dynamic pricing strategy is most actively applied [17,19–23].

In the transportation sector, the development of ITS (Intelligent Transportation System) and the emergence of platform-based mobility services have made it possible to identify real-time boarding demand. Based on this, studies on dynamic pricing have been conducted to maximize the efficient supply of services and the resulting benefits [24,25]. Among them, many platform-based ride-sharing service studies were conducted as a fare-setting model to reduce traffic demand through fare adjustment to induce waiting time and travel time, or to induce supply by improving driver profitability [26–28]. Furthermore, operational data from real-world mobility services enabled verification of the dynamic pricing strategy's effectiveness [29–31].

In Korea's mobility sector, demand-response mobility services by region and means are also spreading, while the Passenger Transport Service Act is being revised, and related regulations are being eased. As a result, various fare strategies are being tried out and introduced along with the existing fixed fare system. In Korea, taxis are a major means of transportation, along with public transportation, railways, buses, and cars, with a transportation share ratio of 2.9 percent as of 2016 [32]. Recently, regarding the taxi service sector, one of the most commonly used transportation systems in Korea, measures to improve the fare system have been proposed in preparation for changes in the market structure, including the dynamic pricing strategy, reverse auction fare strategy, and fare system for joint rides [33].

2.2. Study on Dynamic Pricing Based on Reinforcement Learning in Transportation Field

Reinforcement learning in the transportation sector has been widely used as a learning methodology for the optimization of road systems and the provision of mobile services. Significantly, reinforcement learning can address microscopic behaviors such as signals or individual vehicles, as it can assign different policies to each learning object, known as an agent. It is also applicable as a methodology of Markov chain attributes that reflect supply and demand characteristics in the transportation sector with a short cycle of change. In addition, reinforcement learning has been widely used as a learning methodology for signal system control policies for traffic congestion mitigation in urban road network environments such as intersections [34–36].

The ride-sharing service sector adopted reinforcement learning to establish a vehicle management plan to allocate supply in response to fluctuating demand in real time. To optimize matching between drivers and passengers, simulations have defined state variables, including supply and demand situations and the context of road surroundings. Based on this, mobility service operators can establish optimal vehicle deployment strategies [37,38].

Furthermore, the introduction of deep neural networks due to methodological advances has made learning environments more complicated and expanded their scope,

enabling real-world data-driven learning simulations [39,40]. Based on deep reinforcement learning algorithms such as DQN, A2C (Advantage Actor-Critic), and A3C (Asynchronous Advantage Actor-Critic), distribution mechanism models have been proposed that can reduce idle vehicles and waiting time for passengers in ride-sharing services [41,42].

Meanwhile, reinforcement learning has been used as a learning methodology in a number of areas where dynamic pricing is applied as a fare strategy. Among them, in the transportation sector, it was used as a fare for charging electric vehicles. In addition, the ride-sharing service market has also been studied to mitigate the imbalance in supply and demand and to optimize revenue [43–45]. In particular, in the ride-sharing service market, MDPs were constructed for simulating the introduction of dynamic pricing, and optimal fare policies were derived through reinforcement learning [46–49].

2.3. Summary of the Review and Differentiation of This Study

As a result of analyzing prior research on dynamic pricing, we can report that it was mainly studied in terms of revenue management measures before the early 2000s [13–23]. Since then, the scope of application has expanded with the development of ICT and the growth of the platform-based service industry, and application research has also begun in the transportation sector. Above all, a representative field that applied dynamic pricing within the transportation sector was platform-based mobility on demand services. Many studies have reduced passenger's waiting time by reducing traffic demand through surcharges in high-demand situations. At the same time, it was used as a fare-setting model to induce supply by improving driver profitability [24–33].

Next, for studies in the transportation sector applying reinforcement learning methodology, a number of studies have used reinforcement learning to alleviate traffic congestion in road network environments such as intersections [34–36]. Notably, in the field of ride-sharing services, studies were conducted as a vehicle management measure to optimize the distribution of supply with the aim of optimizing profits and easing imbalance between supply and demand [37,38].

Therefore, this study aims to derive application measures for fare discounts for non-peak hours, which are relatively low in demand, in consideration of the traffic characteristics of each time zone in Korea. In particular, while a number of studies have addressed surcharges to induce supply [24–29,46–49], this study is differentiated in that it provides discount measures to induce demand from the opposite perspective. Furthermore, in the simulation process, we consider interaction between drivers with micro-approaches using multi agent-based reinforcement learning methodologies.

3. Methodology

This chapter describes the centrality analysis methodology for calculating regional fare discount indices for off-peak hours and the methodology of reinforcement learning required for the definition and composition of simulations.

3.1. Degree Centrality Analysis

Social Network Analysis is a methodology for analyzing the identified relationships among the objects that construct the data. It is possible to assess the influence between components, analyze connection patterns, and interpret relationships. Centrality analysis is a type of network analysis methodology that indexes the actors that make up a network and determines the most important actor among them. This methodology is used in various fields such as anthropology [50], economics [51], and urban development [52].

The degree centrality analysis is measured based on the degree of connection between the actors, and the number of different actors is quantified. The degree centrality of an actor is to determine how many different actors are associated with a particular actor. The higher the value, the more other actors connected to the actor. Since the actor's degree centrality calculation is affected by the size of the network, the effect of the size of the

network on the degree centrality should be eliminated. This is called standardized degree centrality (Equation (1)), as described below [53].

$$
\begin{aligned}
C'_D(N_i) &= \frac{C_D(N_i)}{g-1} \\
C'_D(N_i) &: standardized\ actor\ degree\ centrality\ of\ actor\ i \\
C_D(N_i) &: actor\ degree\ centrality\ of\ actor\ i \\
g &: the\ number\ of\ actors
\end{aligned}
\tag{1}
$$

On the other hand, if there is directionality in the relationship between actors that make up the network, centrality should be measured taking that direction into consideration. Depending on the direction of the relationship, it is divided into In-Degree Centrality, which means the number that a particular actor receives from another actor, and Out-Degree Centrality, which means the number that a particular actor connects to another actor. In-Degree Centrality can be interpreted as power or popularity, and Out-Degree Centrality can be interpreted as activity [54].
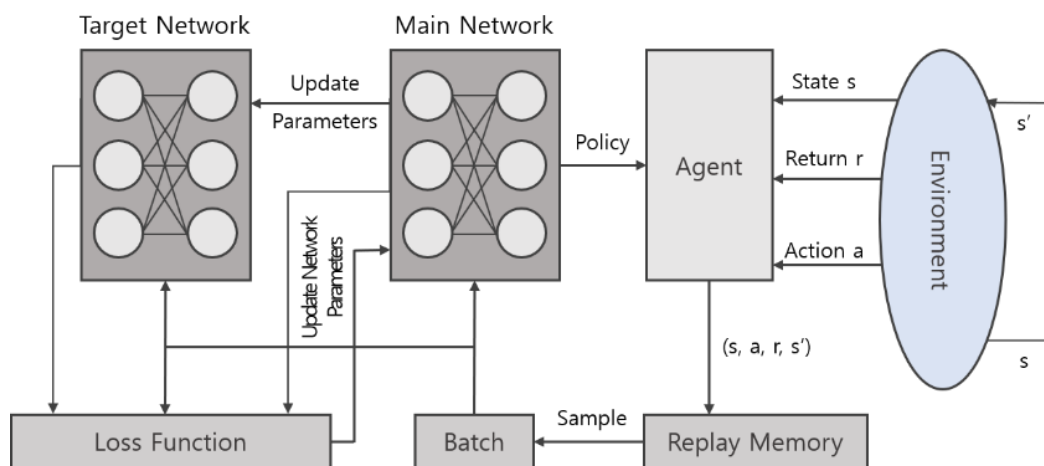
3.2. Deep Reinforcement Learning

Reinforcement learning is a machine learning method in which a defined agent recognizes the current state based on its interaction with the environment and learns action and policy that maximizes reward among the selectable behaviors. To solve the sequential behavioral decision problem within the learning environment where reinforcement learning is conducted, it must be mathematically defined by the Markov Decision Process (MDP). The MDP is represented by a $< S,\ A,\ P,\ R,\ \gamma >$ tuple consisting of state, action, reward function, state transition probability, and the Discount Factor.

Fundamental algorithms in reinforcement learning, including Monte Carlo, Sarsa, and Q Learning, target relatively simple problems with small state spaces and simple environments. However, reinforcement learning to solve problems in the more complex real world has resulted in limit points, including the finite nature of space, and the complexity of computation. Thus, as a way to address this, a deep Q-Network (DQN) algorithm, which parametrizes and approximates Q functions using artificial neural networks starts to be used. DQN is a reinforcement learning algorithm that constructs the Q function as a deep neural network, containing a key concept called Experience Replay, a Target Network to supplement the inefficient learning [55,56].

Experience Replay is a concept designed to prevent overfitting problems between episodes in the learning process of Naive DQN. The agent stores the sample values $(s,\ a,\ r,\ s')$ obtained from exploring the environment in Replay Memory and randomly extracts multiple samples from Replay Memory during training to update the neural network for those samples.

Furthermore, it is problematic that Naive DQN is unstable, and the learning value does not converge because the output value continues to change during the update of weights using a single network. To complement this, two networks have been applied as the Target Network and Main Network. During the learning process, prediction values were calculated for the Target Network and current values for the Main Network. In particular, the Main Network updates the weights according to the output value of the neural network for each learning event. On the other hand, Target Network is not replaced every time, but as a parameter of the Main Network, it operates by copying the weights of the Main Network after a certain number of sufficient learning processes. Figure 1 below represents the key components of the DQN algorithm and their relationships.

Figure 1. Overall learning structure of DQN algorithm based on core components of reinforcement learning.

DQN is trained for agents to behave optimally; thus, as the learning progresses, the Q-value for the agent's actions gradually converges. To confirm this, we derive the Mean Q-values for the last 5000 episodes. We also construct two scenarios to see how the components of the reward function affect learning. Scenario (1) meets zero-sum conditions, and scenario (2) is the case of a reward function that takes into account the driver's waiting time. The results of comparing the Q-values between scenarios are presented in the results section.

4. Materials, Data and Case Study

Simulations for real-time fare bidding systems for off-peak hours are conducted in the following order. First of all, we refined the taxi operation data in Seoul by Gu, which is the district unit. Among them, we calculated the amount of departure and arrival in Gu units, and the basic fares between origin and destination. Next, we were conducted a degree centrality analysis using the previously refined data and calculated the OI Index, a regional fare discount indicator. Finally, to establish a DQN-based fare bidding simulation environment, the origins and destinations are matrixed. Also, the MDP and reward function are defined.

4.1. Data and Refining Methods

This study used Seoul's taxi driving analysis data to reflect the current status of mobility services in Korea. Seoul is the capital city of Korea with a population of about 9.7 million [57], with the largest number of taxis operating [58]. The data used in the analysis consists of the number of rides and drop-offs per hour by major road links in Seoul in 30 min increments [59]. It was based on April 2017, and the scope was targeted for four hours from 10 a.m. to 4 p.m., during the off-peak hours on weekdays, in line with the time range of this study. Administrative district GIS files were used to calculate traffic volume values by origin and destination (OD) in Gu district units. The departure and arrival quantity values for each Gu node are calculated. Next, to calculate the basic fare and distance between origin and destination, the base fares for proportion of distance were calculated using the Network Analysis Tool in ArcGIS. As of November 2020, the basic taxi fare was 3800 won within 2 km, with an additional 132 won per 100 m. Figure 2 below is a map of 25 districts of Seoul, which are subject to analysis [60].
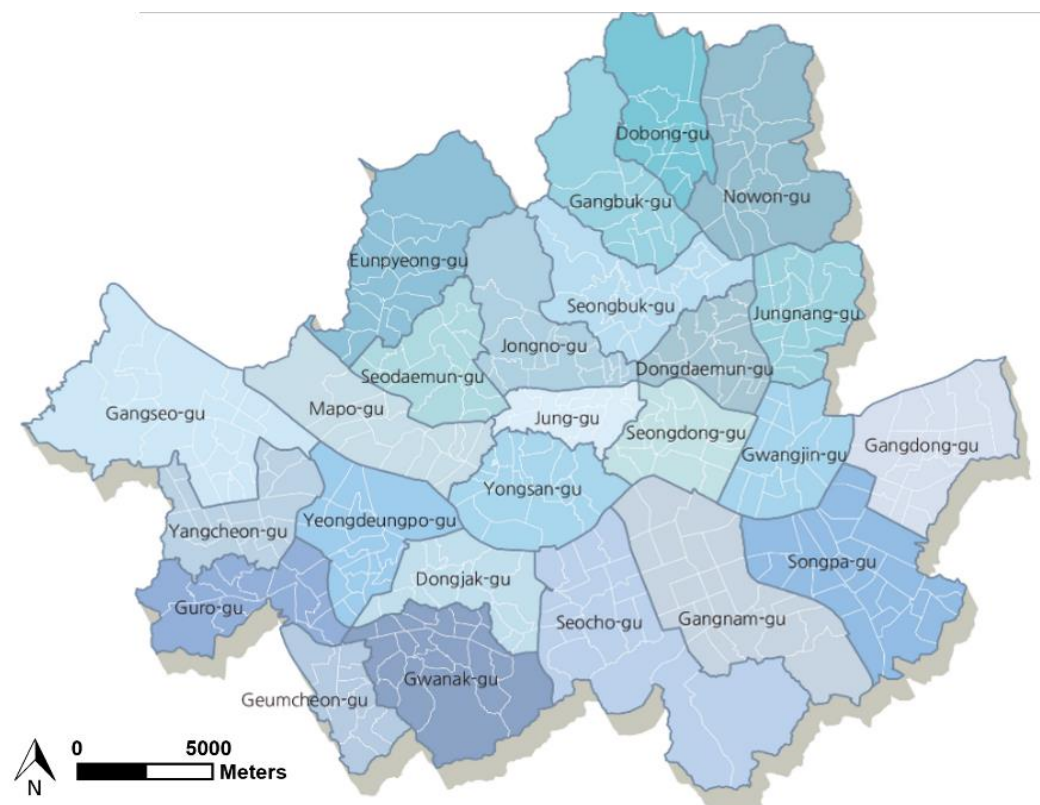
Figure 2. Administrative Map of Seoul.

4.2. Calculation of Fare Discount Indicators by Regions

Regional fares by time zone in the real-time fare bidding system are flexibly adjusted according to the degree of fare discounts offered by drivers. Drivers present the level of fare discount considering the possibility of operation, which is the possibility of matching with passengers, and the supply and demand situation. Therefore, prior to the establishment of a real-time fare competition system, it is intended to calculate indicators to supplement the discount level of fares presented by drivers by reflecting the amount of demand and supply in Gu units.

The regional fare discount indicator is the ratio of the In-degree centrality to the Out-degree centrality, which is defined as the OI Index (Out/In Index). In-degree centrality is an indicator of the amount of supply that is intended to enter a particular destination Gu node. Out-degree centrality can be interpreted as an indicator related to the demand to depart from a particular Gu node to another destination Gu node. Therefore, the smaller the OI index, the higher the supply compared to demand. Drivers compete for lower fares to attract customers in high supply Gu nodes. In other words, it is possible to offer a more drastic fare discount.

The calculated regional OI index values are as shown in Figure 3. The Gu nodes, such as Jungnang-Gu, Eunpyeong-Gu, and Gwanak-Gu, where residential areas are concentrated, showed high OI Index values from 10 a.m. to 11 a.m. due to lingering demand during the previous morning peak hours. However, it went down over time. On the other hand, Jongno-Gu, Jung-Gu, and Gangnam-Gu are Gu nodes, which have a large amount of arrival due to dense business areas. Namely, it was confirmed that low OI Index values were found at the time adjacent to the peak time in the morning. The OI Index was used to supplement the discount level of the regional fares. The basic fare discount option in the fare bidding system was lowered by 0.1 units from ×1.0 to ×0.6, with a total of five. Also, each basic fare discount level is multiplied by the OI index values by Gu nodes' time zones. The high-supply Gu nodes have many competitive drivers in the region, so they can offer daring fare discounts. On the other hand, Gu nodes, which are in high demand compared

to supply, mean that there are already enough opportunities to pick up customers, so the fare discount will not be high. The discount level of the fares corrected by the OI Index is up to ×0.6 times. The minimum fare discount level was set to ×1.0 times because it was only targeted at discounting fares.

Time Unit / Zone	T01 10:00–10:30	T02 10:30–11:00	T03 11:00–11:30	T04 11:30–12:00	T05 12:00–12:30	T06 12:30–13:00	T07 13:00–13:30	T08 13:30–14:00	T09 14:00–14:30	T10 14:30–15:00	T11 15:00–15:30	T12 15:30–16:00
Jongno-gu	0.66	0.72	0.75	0.81	0.88	0.86	0.89	0.97	1.03	1.07	1.10	1.15
Jung-gu	0.69	0.71	0.77	0.89	0.87	0.85	0.85	0.90	0.86	0.89	0.91	0.94
Yongsan-gu	1.26	1.29	1.22	1.19	1.11	1.20	1.16	1.23	1.05	1.07	1.15	1.12
Seoungdong-gu	1.29	1.21	1.24	1.18	1.08	1.08	1.06	1.06	1.03	1.03	0.97	0.92
Gwangjin-gu	1.24	1.24	1.16	1.08	1.07	1.07	1.03	1.03	1.00	0.98	0.95	0.98
Dongdaemun-gu	1.07	1.05	1.02	1.02	1.02	1.05	1.05	1.00	1.03	1.00	1.05	1.02
Jungnang-gu	1.60	1.50	1.45	1.35	1.26	1.32	1.30	1.26	1.19	1.19	1.04	1.04
Seongbuk-gu	1.05	1.05	1.03	0.97	0.98	0.98	0.97	0.94	0.92	0.95	0.90	0.90
Gangbuk-gu	1.17	1.17	1.13	1.14	1.08	1.17	1.14	1.15	1.05	1.00	1.00	1.00
Dobong-gu	1.33	1.40	1.42	1.32	1.29	1.24	1.21	1.21	1.15	1.10	1.05	1.09
Nowon-gu	0.91	0.88	0.91	0.90	0.94	0.88	0.90	0.86	0.87	0.90	0.93	0.90
Eunpyeong-gu	1.86	1.64	1.57	1.33	1.31	1.40	1.36	1.29	1.20	1.06	1.00	1.00
Seodaemun-gu	1.02	1.08	1.13	1.10	1.12	1.07	1.03	0.95	1.05	1.08	1.05	1.11
Mapo-gu	0.97	0.92	0.86	0.83	0.88	0.94	0.98	0.92	0.95	0.92	0.90	0.87
Yangcheon-gu	1.19	1.14	1.17	1.09	1.05	0.97	0.94	0.92	0.94	0.94	0.97	0.97
Gangseo-gu	1.36	1.30	1.29	1.26	1.20	1.29	1.18	1.11	1.11	1.11	1.14	1.07
Guro-gu	0.85	0.85	0.88	0.92	0.93	0.95	0.97	0.94	0.92	0.92	0.95	0.95
Geumcheon-gu	1.05	1.05	1.05	1.00	1.00	1.00	1.05	1.00	1.00	1.05	1.05	1.00
Yeongdeungpo-gu	0.81	0.85	0.89	0.97	0.92	0.85	0.93	0.97	1.03	1.01	1.01	1.04
Dongjak-gu	1.09	1.13	1.09	1.06	1.03	1.00	0.97	1.00	0.94	0.97	0.97	0.94
Gwanak-gu	1.52	1.41	1.36	1.31	1.28	1.33	1.36	1.33	1.27	1.23	1.15	1.11
Seocho-gu	1.01	1.00	0.97	0.94	0.94	0.94	0.90	0.88	0.88	0.90	0.87	0.88
Gangnam-gu	0.69	0.73	0.82	0.85	0.90	0.89	0.96	0.95	1.03	1.03	1.06	1.06
Songpa-gu	1.28	1.19	1.11	1.07	1.06	1.06	1.07	1.07	1.00	1.05	1.02	1.02
Gangdong-gu	1.25	1.26	1.26	1.17	1.21	1.26	1.23	1.14	1.04	1.05	1.04	0.96

Figure 3. OI Index by time zones and Gu nodes.

OI Index is also an indicator set on the basis of Gu Node's level of demand and supply, which should differ in the distribution of regional drivers' profits and matching rates when simulated to derive discount levels of fares. Therefore, in order to explore this, we would like to analyze the profit and matching rate of drivers who competed for fares in Gu Nodes with low, average, and high level of OI Index values. Thus, in Section 5, the results of the simulation are identified for the representative Gu Node, Gangnam-gu, Dongjak-gu, and Eunpyeong-gu.

4.3. Definition of MDP and Environment

Within the real-time fare bidding system, drivers compete for the amount of traffic to the destinations from the origins by time zones. Driver agents who offer cheaper fares are matched with the passengers. Therefore, each agent presents a discount level of the appropriate fare with the goal of matching more passengers. A real-time fare bidding system for simulation is defined as a tuple $G = < S, A, P, R, \gamma >$ in Markov game format for N agents. The environmental components and definitions are as follows.

1. Environment and Agent: The agent, the subject of learning, proceeds through interaction with the environment in an environment composed of all the rules set for learning. The environment of this simulation consists of 625 pairs, with 25 Gu regions set as departure and arrival nodes. The value included the case for traffic inside the Gu nodes. Next, the agent was set up as drivers competing for fares at each OD pairs. Thus, the relationship between agents is defined as competitive. Agents were also distinguished by their operational characteristics, called Waiting Time Value. Waiting Time Value (WTV) is the application of the concept of time value for waiting time. WTV refers to the psychological sacrifices and blows that drivers feel about waiting time [58]. Therefore, there are two types of drivers for each OD pairs: driver B with a high fare discount because of the relatively large loss of waiting time, and driver A with a relatively low fare discount level even if waiting for a long time. The number of agents per OD pair is fixed at $N = 2$, but N_t differs by t, the basic unit of time for fare competition.

2   State: The state of the environment consists of a global state of s_t and an agent specific state of s_t^i. s_t means the operation status between OD pairs using one-hot encoding. In each OD pair, agents include p_t, a discounted fare offered in time t. Agents located in the same OD pair share the global state and configure their own state values.

3   Action: The agent's action space is the discount level of the fare that will be multiplied by the basic fares between each OD pair. According to time unit t, agent i's action a_t^i is a total of five discretized actions ranging from $\times 1.0$ times to $\times 0.6$ times in units of 0.1. For example, $a_t = 1$ means that the agent offers a $\times 1.0$ discount on the basic fare at the time unit, and $a_t = 5$ gives a $\times 0.6$ discount level. This is the basic fare discount level set by referring to Uber's fare discount case [61]. Multiply the OI index to apply the corrected fare discount level for each regional time zone.

4   Reward: Each driver agent is given a reward value according to the reward function R_i, and different reward functions are applied depending on the outcome of the fare competition. Drivers who offer the highest discount level in the fare bidding for OD pairs are given positive rewards by multiplying the distance proportional basic fare (ODP) and the discount level (DC_t^i), which means the driver's profit when matched with passengers. On the other hand, drivers who have presented a lower discount level in fare bidding than other agents are considered to have failed to match in that time unit and are given negative rewards. At this time, the negative reward function should take into account the driver's waiting time (WT_i) and time value (TV_i) that occur if the matching fails. Therefore, the negative reward function was established by referring to the components of Song. J. et al.'s [49] negative reward values and the waiting time data of taxis in regions of Seoul [62]. Driver A was given 7 min, the average minimum waiting times for each region, and Driver B was given 9 min, the average maximum waiting times for each region. The reward function for the success and failure of the fare bidding consists of the following formula.

$$R_i = \begin{cases} DC_t^i \times ODP, & |Matched \\ -DC_t^i \times WT_i \times TV_i, & |Unmatched \end{cases} \quad (2)$$

5   State Transition Probability: In this simulation, the origins and destinations were set to node in advance and matrixed, targeting deterministic traffic environments. On the other hand, the lowest fares for each OD pair are calculated according to the discount level of the fares presented by the driver agents. This determines the stochastic number of operations per agent.

4.4. Designing of Fare Bidding Simulator between Drivers

Next, the fare bidding process between driver agents is described step by step according to the time unit of the simulation. When the simulation begins, the driver agents choose a discount level that will multiply the base fare of each OD pair for every round. If the discount level presented by the agents are the same, they are randomly matched.

Based on 30 min, which is the basic time unit of taxi operation data, the first round is set and it consists of 12 rounds (R). The OI Index values for OD pairs change for each round. At each round, the driver agent has a total of three time steps (t). The time step is the basic unit of time in the fare bidding system. This aims to vary the number of fare biddings in each round because the actual travel time between OD pairs is different. Therefore, three time steps per round can be applied. Depending on the actual travel time between origins and destinations, the fare bidding value was used three times for less than 15 min, two times for less than 30 min, and one time for more than 30 min. Optimization of the DNN network for DQN is constructed by referring to Mnih, V. et al. [56], Oroojlooyjadid, A. et al. [63] and is as shown in Algorithm 1 (DQN Algorithm Pseudo-code).

Algorithm 1. DQN pseudo-code for drivers' fare bidding system.

Procedure DQN
  for *Episode* $= 1 : n$ do
    Initialize Experience Replay Memory, $E_i = [\ ]$, $\forall^i$
    Reset OI Index, ODP, WT, TV for each agent
  For $r = 1 : R$ do
      For $t = 1 : T$ do
        For $i = 1 : 2$ do
          With probability ε take random action a_t
          otherwise set $a_t = argminQ(s_t, a, \theta)$
          Execute action a_t, observe reward r_t, and state s_{t+1}
          Add $(s_t^1, a_t^1, r_t^1, s_{t+1}^1)$ into the E_i
          Get a mini-batch of experiences (s_j, a_j, r_j, s_{j+1}) from E_i
          Set $y_i = \begin{cases} r_j & \text{if is ther terminal state} \\ r_j + minQ(s, a, \theta^-) & \text{otherwise} \end{cases}$
          Run one forward and one backward step on the DNN with loss function
          $(y_j - Q(s_j, a_j, \theta))^2$
          Every C iterations, set $\theta^- = \theta$
        end for
      end for
    end for
    Run feedback scheme, update experience replay of each agent
  End for
End procedure

The types of Hyperparameters used for learning and their values are shown in Table 1. Values have been applied by referring to DQN cases in Open ai Gym [64]. We use Python 3 as an analytical environment for experimental and model generation, and use Tensorflow 1.12 and Keras 2.1.5 libraries for neural network learning

Table 1. Type and value of hyper-parameters used in analysis.

Hyperparameter	Value
Minibatch Size	32
Replay Buffer Size	10,000
Init Exp	1.0
Final Exp	0.1
Discount Factor	0.99
Learning Rate	0.001
Target Update Rate	0.01
Annealing Step	10,000

5. Results

In this chapter, we conduct a DQN-based simulation for fare bidding in the off-peak hours and analyze the results. To verify the performance of learning before identifying the results for the level of discounts by regional time zones derived from the bidding simulation of driver agents, we checked the cumulative reward values and the mean Q-values by agents during 50,000 episodes.

5.1. Evaluation of the Learning Simulations

As shown in Figure 4, the average Q-value values for each episode were derived from the 45,000th to the last, 50,000th, episode. The average Q-values of the derived fully competitive scenario (1) were gradually reduced and converged. On the other hand, for Scenario (2) considering waiting time values, the Q-values converged to a higher level than Scenario (1). Therefore, we apply a reward function that takes into account the value of waiting time in learning the fare bidding simulations between agents.

(**a**)   (**b**)
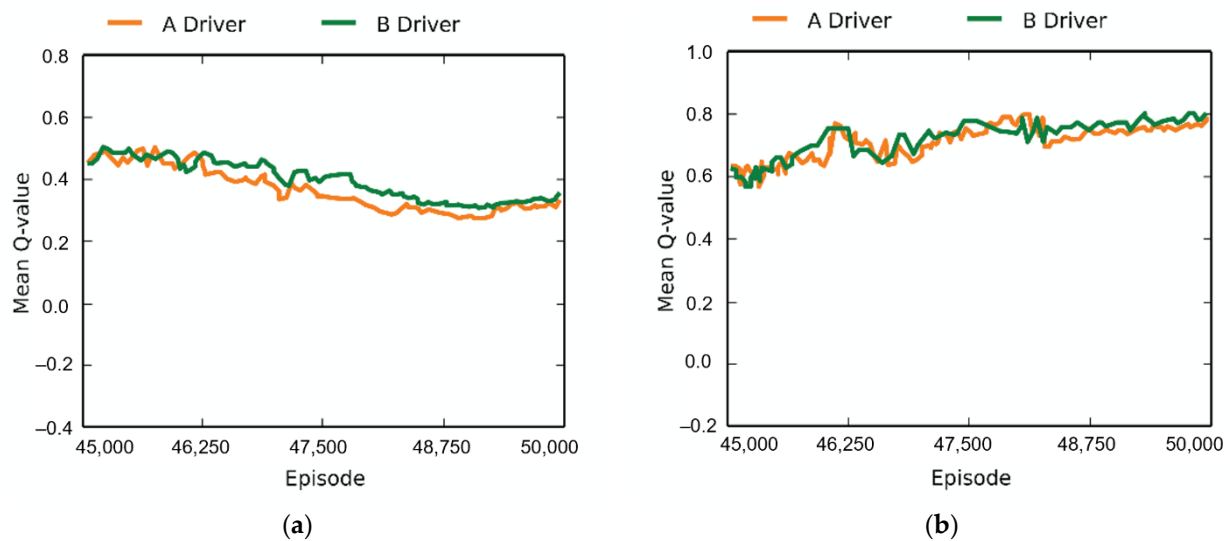
Figure 4. Overall learning structure of DQN algorithm based on core components of reinforcement learning as: (**a**) Scenario 1; (**b**) Scenario2.

Next, the average cumulative reward values for each agent according to the learning of 50,000 episodes were described when the reward function considering the driver's waiting time was applied. To verify, the average reward values of all OD pairs were calculated for every 1000 episodes. Figure 5 is the average cumulative reward values for driver A and B agents. The average cumulative reward values are the accumulated reward obtained under the agent's policy, which should be increased as the learning progresses. As a result of comparison of cumulative reward values between agents, when driver B with large waiting time values was compared to driver A, the agent obtained a higher reward value on average.
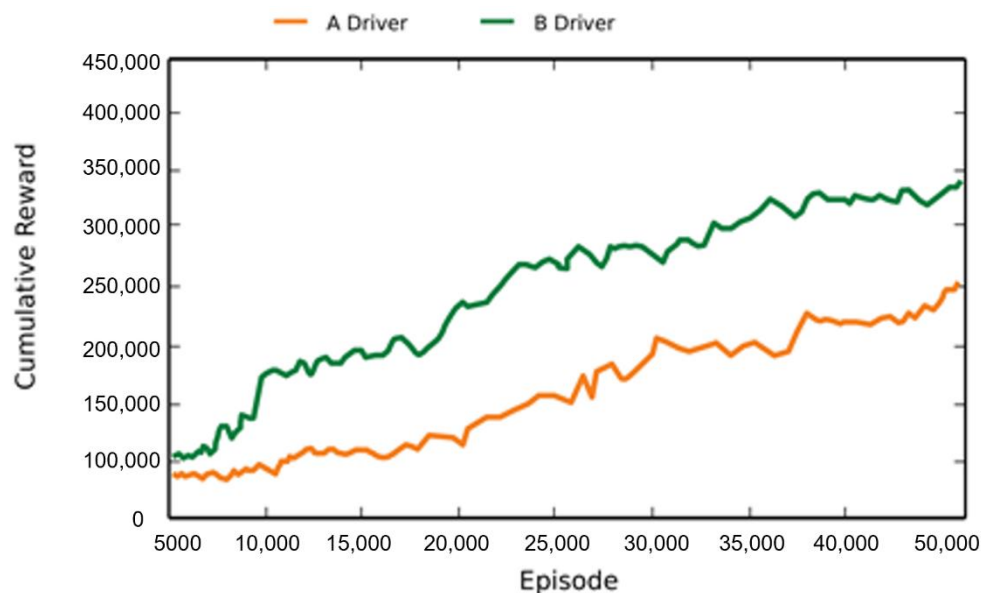
Figure 5. Average cumulative reward value of A driver and B driver by Episodes.

5.2. Analysis of Driver's the Profit and Matching Rates by Gu Nodes

The OI Index by Gu nodes is an indicator calculated to reflect regional demand and supply characteristics when selecting the discount level for driver agents. Thus, the Gu nodes classified according to the characteristics of supply and demand to determine the discount level of regional fares derived from this simulation. The discount level of

fares for each origin and destination was compared with the driver's profit and matching rates. Profit ratio means revenue plus waiting losses divided by revenue. This confirms the actual operating profit on the winning of the competition. The matching rate refers to the percentage of drivers' wins among the total number of matches between origin and destination.

According to the analysis of drivers' profits and matching rates, separated by regional demand and supply characteristics, the level of fare discounts that can improve the operational success rates varies depending on the average OI index level by region. In areas with high OI Index, the higher the demand, the higher the success rate of drivers who offered low discount levels. However, in areas with smaller OI indexes, the higher the demand, the higher the success rate of drivers who offered higher discount levels. In addition, regions with average OI Index values had high operational success rates for drivers who offered high discount levels in all time zones.

In order to closely examine the results of learning, drivers' profit and matching rates were identified in areas representing low, average, and high OI Index values. The representative areas are Gangnam-Gu, Dongjak-Gu, and Eunpyeong-Gu. In addition, the driver's profit and number of trips from fare bidding depends on the travel time between the OD pairs, so, for absolute comparison, the distance between origins and destinations was classified and examined. The classification by travel time consists of short distances (within 15 min), medium distances (within 30 min), and long distances (over 30 min), based on the average travel time of 15 min when the origin and destination pairs are adjacent.

5.2.1. Analysis of the Profit and Matching Rate in District with Low Demand and Higher Supply

Gangnam-Gu indicates districts where demand is low compared to supply because the average OI index is small during the off-peak hours. This means that the demand for use is less than the arrival of vehicles, and it mainly has the characteristics of business areas and central commercial areas. Regardless of the distance from the destination, in the morning time, when demand is low compared to supply, driver A, who suggests a relatively low fare discount level, had a higher rate of profit and matching rate than driver B, who suggests an active discount level (Figure 6). However, it was found that driver B's profit and matching rate improved as demand-to-supply levels increased over the afternoon. In the morning time zone, driver A's matching rate was up to 57% on a short distance (a), and the profit rate was up to 89% on a long distance (c). Next, in the afternoon, driver B had a matching rate of up to 64% and a profit rate of up to 86%.
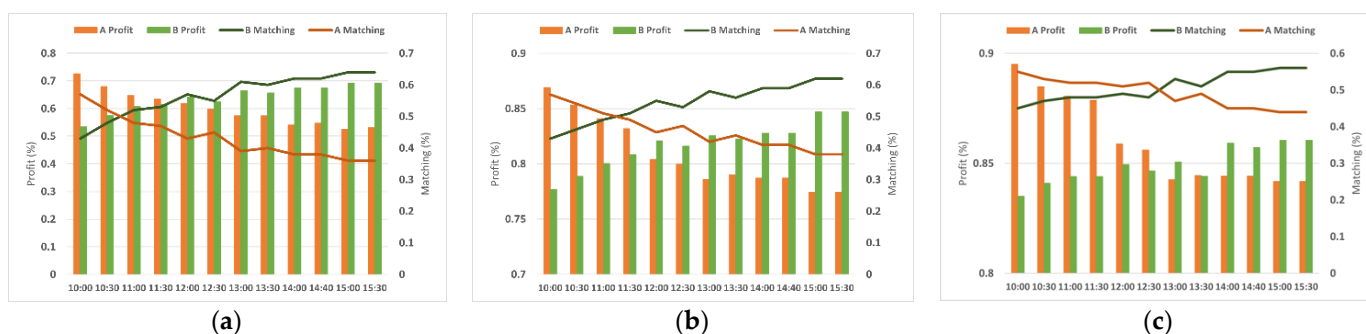
Figure 6. Profit and matching rate by drivers in Gangnam-Gu: (**a**) Short Distance; (**b**) Medium Distance; (**c**) Long Distance.

Incidentally, as the distance from the destination increases and the travel time increases, the difference between the two drivers' profits and matching rates decreases. In the case of profit rate, there was a difference of up to 20% in short distance (a) driving as of 10 a.m., but it decreased by 9% in intermediate distance (b) and 6% in long distance (c). The matching rate also differed by up to 28% in short distance driving as of 3 p.m., but decreased to 24% in intermediate distance (b) and 12% in long distance (c). This is because the longer the distance between OD pairs, the greater the value that has a positive impact on the driver's

profits, and the greater the impact compared to the negative factors. The increasing level of the basic fare for distance proportions is a positive factor for the driver's profit, and the waiting time value is a negative factor.

5.2.2. Analysis of the Profit and Matching Rate in District with Average Demand and Supply

Dongjak-Gu represents an area with an average OI Index value of 25 district nodes, meaning that the level of vehicle arrival and demand for use is similar. In all time zones, Dongjak-Gu has a high matching rate of Driver B, which has actively competed in the fare bidding simulation. In addition, as shown in Figure 7, driver B also accounted for maximum profit rate of 68% over short distance (a), 75% over intermediate distance (b) and 89% over long distance (c). The matching rate was up to 65% and 63% for both short and intermediate distance as of 11:30 and 61% for long distance, respectively. In particular, the difference from driver A was the largest between 10 a.m. and 1 p.m., and as of 11:30 a.m., there was a deviation of up to 30% from short distance driving (graph (a)). The figure declined to 4% over time, gradually decreasing in the afternoon hours. This is because the demand-to-supply level is lowered in the afternoon off-peak hours, so both driver agents suggest a high discount level.

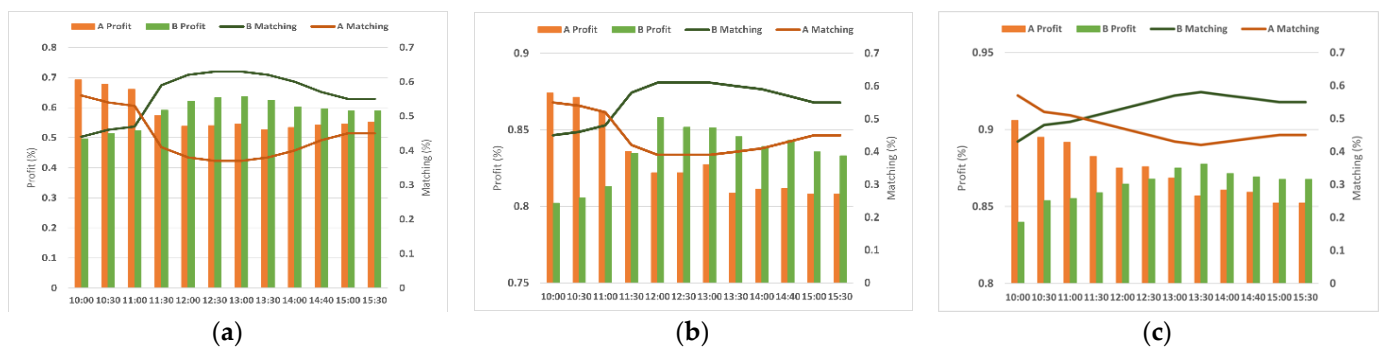**Figure 7.** Profit and matching rate by drivers in Dongjak-Gu: (**a**) Short Distance; (**b**) Medium Distance; (**c**) Long Distance.

5.2.3. Analysis of the Profit and Matching Rate in District with High Demand and Lower Supply

Eunpyeong-Gu is representative of regions with large OI Index values and high demand compared to supply. This means that the demand for use is relatively high compared to the arrival of vehicles, which is typically found in residential areas. Therefore, those areas are affected by the residual demand resulting from their commute to work at the time of the morning-peak hours. In other words, in the morning hours when demand for supply is high, both drivers offer relatively low fare discounts, so driver A's profits are high due to the relatively low waiting time values. As shown in Figure 8, the profit was up to 69% over short distances (a) and 87% over intermediate distances (b), and up to 90% over long distances (c). On the other hand, driver B's profits and matching rate were both high in the afternoon off-peak hours after the remaining demand was resolved, with the biggest difference between drivers, especially between 12:00 and 2:00 p.m. There was a difference of 26% for short distances (a) and 22% for intermediate distances (b), and up to 16% for long distances (c). This is believed to be due to the large difference in discount levels selected by drivers depending on the value of the individual's waiting time at a time when the OI Index values that affect the discount level are not extremely skewed.

Figure 8. Profit and matching rate by drivers in Eunpyeong-Gu: (**a**) Short Distance; (**b**) Medium Distance; (**c**) Long Distance.

Therefore, on average, in areas where demand is high compared to supply, there is already a demand for the use of services in the morning time zones, so it is unnecessary to present an assertive discount level. However, in some time zones, including off-peak hours in the afternoon, active discount levels mean that driver profits and matching rates can be maximized.

6. Conclusions

In the field of mobility, the problem of structural imbalance in terms of supply and demand has been continuously raised due to changes in traffic patterns by time zone. During off-peak hours in the Seoul metropolitan area, demand declines sharply while supply stays put, resulting in problems such as vacant vehicle problems and traffic congestion. As a solution to this problem, this study aimed to derive real-time fares by establishing a bidding system between drivers based on deep reinforcement learning to promote taxi uses and improve the driver's profits.

The process was carried out as follows. First, based on the data of taxi operation in Seoul, the amount of departure and arrival trips in the Gu unit and the basic fares between origins and destinations were calculated. In addition, the OI Index, a regional fare discount indicator, was calculated through an In-degree and Out-degree centrality analysis based on the value of departure and arrival. Areas with low OI Index included business and commercial areas such as Jongno-Gu, Jung-Gu, and Gangnam-Gu. On the other hand, the high OI value was derived mainly from residential areas such as Jungnang-Gu, Eunpyeong-Gu, and Gwanak-Gu.

Next, the simulation of fare bidding between drivers was conducted through a multi-agent based deep reinforcement learning. First, the traffic environment was constructed for 625 OD pairs consisting of 25 nodes. Driver agents, classified by the value of waiting time at each origin, were required to choose a level from the five discount ranges for the basic fare between OD-pairs. The driver who presented the highest discount level was given a positive reward value at that fare, and other agents were given a negative reward value considering the waiting time and hourly value.

Finally, the characteristics of supply and demand by regional time zone were identified and the optimal fare discount level was derived accordingly. As a result, high-demand areas, including Eunpyeong-Gu, Jungnang-Gu, and Gwanak-Gu, had lower discounts, and lower-demand areas with lower OI indexes, including Gangnam-Gu, had higher discounts. In particular, Gangnam-Gu had the highest discount level with an average of 36%, while Dongjak-Gu had 33% and Eunpyeong-Gu had an average of 17%.

In addition, we analyze the driver's profit and matching rate according to the distance between the OD pairs in each node. As a result, in areas with high demand for use, the higher the demand, the higher the matching rate of drivers who offered relatively passive discount levels. On the other hand, in areas with high supply, the higher the demand, the higher the matching rate of drivers who actively offered an active fare discount level. In addition, in areas where demand and supply are relatively balanced, drivers who actively offered high fare discounts at all times were found to have a high rate of matching.

Meanwhile, it has been shown that the pattern of drivers' profit and matching rates is similar at medium and long distances, but the difference decreases as the distance traveled increases. This means that discounts on fares have a significant impact on the fare and matching rates of drivers operating in short distances.

According to the results of the review of the literature, the dynamic pricing system was applied primarily to induce the taxi supply with surcharges, but this study proposed a method to encourage demand through a fare discount. In particular, we made a contribution by finding optimal solutions by applying network-based reinforcement learning methodologies. This study is also significant in that it has attempted to secure additional taxi demand in off-peak hours through competition for fare discounts. In particular, it is believed that securing demand for taxi use in terms of sustainable transportation planning will contribute to the establishment of eco-friendly transportation measures, such as reducing air pollutant emissions by reducing waiting time for empty cars. In addition, considering the current trend of spreading platform-based mobility services, it is possible to implement the technology for real-time price competition between drivers.

The methodological limitations and future challenges of this study are as follows. First, the simulation presented the discount level of the fare for one operation between the origins and the destinations. Therefore, it is necessary to reflect the expected demand for the next operation in consideration of the driver's series of trip chains and seek ways to derive a more realistic discount level of fares. Secondly, this simulation sets the basic unit of time to 30 min. Therefore, it is assumed that time units will need to be subdivided to take into account the level of change in demand and supply by more granular time units.

Author Contributions: Conceptualization, K.H.; Data curation, J.S.; Formal analysis, Y.C.; Investigation, Y.C.; Methodology, Y.C.; Project administration, K.H.; Resources, Y.C.; Supervision, J.S.; Validation, J.S.; Writing—Original draft, Y.C.; Writing—Review & editing, Y.C., M.K. and K.H. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Firnkorn, J.; Müller, M. Selling mobility instead of cars: New business strategies of automakers and the impact on private vehicle holding. *Bus. Strategy Environ.* **2012**, *21*, 264–280. [CrossRef]
2. Ministry of Land, Infrastructure and Transport. Transformation of Taxis Combined with the Platform Continues. Available online: http://www.molit.go.kr/USR/NEWS/m_71/dtl.jsp?id=95085112 (accessed on 1 September 2020).
3. Jittrapirom, P.; Marchau, V.; van der Heijden, R.; Meurs, H. Future implementation of Mobility as a Service (MaaS): Results of an international Delphi study. *Travel Behav. Soc.* **2018**, *21*, 281–294. [CrossRef]
4. Park, J.; Park, S.W.; Kim, Y. *Public Transportation Policies on Commuter Balance and Efficiency*; Korea Transport Institute: Sejong, Korea, 2014.
5. An, G. *Seoul's Taxi Usage and Operation Status and Improvement Plan*; Policy Report 186; Seoul Institute: Seoul, Korea, 2015.
6. Kakaomobility Report. 2019. Available online: https://brunch.co.kr/@kakaomobility/36 (accessed on 10 September 2020).
7. Uber Technologies, Inc. How Surge Pricing Works. Available online: https://www.uber.com/enGB/drive/partner-app/how-surge-works/ (accessed on 15 October 2019).
8. Lyft, Inc. How to Estimate a Lyft Ride's Cost. Available online: https://help.lyft.com/hc/en-us/articles/115013080308-How-to-estimate-a-Lyft-ride-s-cost (accessed on 15 October 2019).
9. *Kakaomobility Report, Data Driven Mobility Innovation*; Kakaomobility: Seongnam, Korea, 2020.
10. Moran, D.; Kanemoto, K.; Jiborn, M.; Wood, R.; Többen, J.; Seto, K. Carbon footprints of 13,000 cities. *Environ. Res. Lett.* **2018**, *13*, 064041. [CrossRef]

11. Sim, S.; Jeonga, S.; Park, H.; Park, C.; Kwak, K.-H.; Lee, S.-B.; Kim, C.H.; Lee, S.; Chang, J.S.; Kang, H.; et al. Co-benefit potential of urban CO_2 and air quality monitoring: A study on the first mobile campaign and building monitoring experiments in Seoul during the Winter. *Atmos. Pollut. Res.* **2020**, *11*, 1963–1970. [CrossRef]

12. *National Transportation Statistics 2019*; Korea Transport Institute: Sejong, Korea, 2019.

13. Gallego, G.; Van Ryzin, G. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Manag. Sci.* **1994**, *40*, 999–1020. [CrossRef]

14. Bitran, G.R.; Mondschein, S.V. Periodic pricing of seasonal products in retailing. *Manag. Sci.* **1997**, *43*, 64–79. [CrossRef]

15. Elmaghraby, W.; Keskinocak, P. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Manag. Sci.* **2003**, *49*, 1287–1309. [CrossRef]

16. Diao, J.; Zhu, K.; Gao, Y. Agent-based simulation of durables dynamic pricing. *Syst. Eng. Procedia* **2011**, *2*, 205–212. [CrossRef]

17. Yoon, M.G.; Lee, H.Y.; Song, Y.S. Discrete Choice Dynamic Pricing and Seat Control Problem in Airlines. *Bus. Sci.* **2012**, *29*, 91–103.

18. Cho, J.H. A Dynamic Pricing Negotiation Model in the Online Ticket Resale Market. *J. Electron. Trade Assoc. Korea* **2009**, *14*, 133–148.

19. Borenstein, S.; Jaske, M.; Rosenfeld, A. *Dynamic Pricing, Advanced Metering, and Demand Response in Electricity Markets*; UCEI: Berkeley, CA, USA, 2002.

20. Roozbehani, M.; Dahleh, M.; Mitter, S. Dynamic pricing and stabilization of supply and demand in modern electric power grids. In Proceedings of the 2010 First IEEE International Conference on Smart Grid Communications, Gaithersburg, MD, USA, 4–6 October 2010; pp. 543–548.

21. Jia, L.; Tong, L. Dynamic pricing and distributed energy management for demand response. *IEEE Trans. Smart Grid* **2016**, *7*, 1128–1136. [CrossRef]

22. Yoo, I.H. Study on Implementing Dynamic Pricing to Demand Response System. *J. Korean Electr. Soc.* **2009**, *40*, 541–542.

23. Kwon, J.H.; Han, S.M.; Park, S.W.; Choi, D.C.; Kim, B.H. Analysis of The Impact of Demand Response by Dynamic Pricing for Residential Customers. *J. Korean Electr. Soc.* **2010**, *2010*, 548–549.

24. Saharan, S.; Bawa, S.; Kumar, N. Dynamic pricing techniques for Intelligent Transportation System in smart cities: A systematic review. *Comput. Commun.* **2020**, *150*, 603–625. [CrossRef]

25. Pueboobpaphan, S.; Indra-Payoong, N.; Opasanon, S. Experimental analysis of variable surcharge policy of taxi service auction. *Transp. Policy* **2019**, *76*, 134–148. [CrossRef]

26. Qiu, H.; Li, R.; Zhao, J. Dynamic pricing in shared mobility on demand service. *arXiv* **2018**, arXiv:1802.03559.

27. Guan, Y.; Annaswamy, A.M.; Tseng, H.E. Towards Dynamic Pricing for Shared Mobility on Demand using Markov Decision Processes and Dynamic Programming. *arXiv* **2019**, arXiv:1910.01993.

28. Banerjee, S.; Johari, R.; Riquelme, C. Pricing in ride-sharing platforms: A queueing-theoretic approach. In Proceedings of the Sixteenth ACM Conference on Economics and Computation, Portland, OR, USA, 15–19 June 2015; p. 639.

29. Sayarshad, H.R.; Gao, H.O. A scalable non-myopic dynamic dial-a-ride and pricing problem for competitive on-demand mobility systems. *Transp. Res. Part C Emerg. Technol.* **2018**, *91*, 192–208. [CrossRef]

30. Yan, C.; Zhu, H.; Korolko, N.; Woodard, D. Dynamic pricing and matching in ride-hailing platforms. *Nav. Res. Logist.* **2019**, *67*, 705–724. [CrossRef]

31. Adnan, M.; Nahmias Biran, B.-H.; Baburajan, V.; Basak, K.; Ben-Akiva, M. Examining impacts of time-based pricing strategies in public transportation: A study of Singapore. *Transp. Res. Part A* **2020**, *140*, 127–141. [CrossRef]

32. e-Country Indicators. Public Transportation Sharing Rate. Available online: http://www.index.go.kr/potal/stts/idxMain/selectPoSttsIdxSearch.do?idx_cd=4081&stts_cd=408101&freq=Y (accessed on 29 July 2021).

33. Ha, H.G.; Park, M.Y.; Yang, T.W.; Kim, K.S. Directions of Taxi Fare System Improvements in the 4th Industrial Revolution Era. *Transp. Technol. Policy* **2019**, *16*, 44–55.

34. Choi, J.; Cho, Y.; Jeong, I. Multiple-Intersection Traffic Signal Control based on Traffic Pattern Learning. *J. Soc. Inf. Sci. Pract. Lett. Comput.* **2014**, *20*, 171–179.

35. Kim, D.H.; Jung, O. A Study on Cooperative Traffic Signal Control at multi-intersection. *J. Electr. Electron. Soc.* **2019**, *23*, 266–271.

36. Joo, H.; Lim, Y. Distributed Traffic Signal Control at Multiple Intersections Based on Reinforcement Learning. *J. Korea Commun. Assoc.* **2020**, *45*, 303–310.

37. Wei, C.; Wang, Y.; Yan, X.; Shao, C. Look-ahead insertion policy for a shared-taxi system based on reinforcement learning. *IEEE Access* **2017**, *6*, 5716–5726. [CrossRef]

38. Xu, Z.; Li, Z.; Guan, Q.; Zhang, D.; Li, Q.; Nan, J.; Ye, J. Large-scale order dispatch in on-demand ride-hailing platforms: A learning and planning approach. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19–23 July 2018; pp. 905–913.

39. Jindal, I.; Qin, Z.T.; Chen, X.; Nokleby, M.; Ye, J. Optimizing taxi carpool policies via reinforcement learning and spatio-temporal mining. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 1417–1426.

40. Lin, K.; Zhao, R.; Xu, Z.; Zhou, J. Efficient large-scale fleet management via multi-agent deep reinforcement learning. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19–23 July 2018; pp. 1774–1783.

41. Ke, J.; Xiao, F.; Yang, H.; Ye, J. Optimizing online matching for ride-sourcing services with multi-agent deep reinforcement learning. *arXiv* **2019**, arXiv:1902.06228.

42. He, S.; Shin, K.G. Spatio-temporal capsule-based reinforcement learning for mobility-on-demand network coordination. In Proceedings of the World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019; pp. 2806–2813.

43. Lei, C.; Jiang, Z.; Ouyang, Y. Path-based dynamic pricing for vehicle allocation in ridesharing systems with fully compliant drivers. *Transp. Res. Part B Methodol.* **2019**, *38*, 77–97. [CrossRef]

44. Rambha, T.; Boyles, S.D. Dynamic pricing in discrete time stochastic day-to-day route choice models. *Transp. Res. Part B Methodol.* **2016**, *92*, 104–118. [CrossRef]

45. Chen, M.; Shen, W.; Tang, P.; Zuo, S. Optimal vehicle dispatching schemes via dynamic pricing. *arXiv* **2017**, arXiv:1707.01625.

46. Wu, T.; Joseph, A.D.; Russell, S.J. *Automated Pricing Agents in the On-Demand Economy*; Berkeley EECS: Berkeley, CA, USA, 2016.

47. Guan, Y.; Annaswamy, A.M.; Tseng, H.E. Cumulative prospect theory based dynamic pricing for shared mobility on demand services. In Proceedings of the 2019 IEEE 58th Conference on Decision and Control (CDC), Nice, France, 11–13 December 2019; pp. 2239–2244.

48. Kamatani, T.; Nakata, Y.; Arai, S. Dynamic pricing method to maximize utilization of one-way car sharing service. In Proceedings of the 2019 IEEE International Conference on Agents (ICA), Jinan, China, 18–21 October 2019; pp. 65–68.

49. Song, J.; Cho, Y.J.; Kang, M.H.; Hwang, K.Y. An Application of Reinforced Learning-Based Dynamic Pricing for Improvement of Ridesharing Platform Service in Seoul. *Electronics* **2020**, *9*, 1818. [CrossRef]

50. Ortner, S.B. *Anthropology and Social Theory*; Duke University Press: Durham, NC, USA, 2006.

51. Jang, S.E.; Lee, S.H. A Study of Themes and Trends in Research of Global Maritime Economics through Keyword Network Analysis. *J. Korea Port Econ. Assoc.* **2016**, *32*, 79–95.

52. Seo, G.C. An Analysis on the Centrality of Network Structure in Land Market: Focused on Land Sales Markets in the Seoul Region. *Apprais. Stud.* **2019**, *18*, 203–232. [CrossRef]

53. Kwak, K.Y. *Social Network Analysis*; Chungram: Seoul, Korea, 2017.

54. Kim, Y. *Social Network Analysis*; Parkyoungsa: Seoul, Korea, 2016.

55. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Graves, A.; Antonoglou, I.; Wierstra, D.; Riedmiller, M. Playing atari with deep reinforcement learning. *arXiv* **2013**, arXiv:1312.5602.

56. Mnih, V.; Badia, A.P.; Mirza, M.; Graves, A.; Lillicrap, T.; Harley, T.; Kavukcuoglu, K. Asynchronous methods for deep reinforcement learning. In Proceedings of the International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016; pp. 1928–1937.

57. Korean Statistical Information Service. Available online: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B040A3&checkFlag=N (accessed on 25 July 2021).

58. National Association of Private Taxi Transport Business Association. Available online: https://www.xn--h49al54beid46l.org (accessed on 11 October 2020).

59. Seoul's Open Data Square. Available online: http://data.seoul.go.kr/rss/rssView.do?searchType=0000 (accessed on 1 April 2020).

60. NGII, Administrative Map of Seoul. Available online: http://nationalatlas.ngii.go.kr/pages/page_1952.php#prettyPhoto (accessed on 25 July 2021).

61. Uber Blog. California Drivers: Set Your Own Fares When You Drive with Uber. Available online: https://www.uber.com/blog/california/set-your-fares/ (accessed on 25 September 2020).

62. An, G.; Kang, T.; Kim, B. *Taxi Service Evaluation Service Report 2014*; Policy Report 186; Seoul Institute: Seoul, Korea, 2015.

63. Oroojlooyjadid, A.; Nazari, M.; Snyder, L.; Takác, M. A deep q-network for the beer game: A reinforcement learning algorithm to solve inventory optimization problems. *arXiv* **2017**, arXiv:1708.05924.

64. OpenAI Gym. Available online: https://gym.openai.com/ (accessed on 25 September 2020).