Trip Pricing in User-Based Relocation for Station-Based Carsharing Systems

Ruiyou Zhang, *Member, IEEE*, Haiyu Kan, and Ilkyeong Moon

***Abstract*— The imbalanced distribution of vehicles due to asymmetric demands has become a main operational challenge in increasingly popular carsharing services. This study investigates a user-based relocation strategy to alleviate the imbalances in widely adopted non-reserved station-based carsharing systems, in which users are induced to relocate vehicles via monetary incentives. An incentive-based trip pricing problem considering the strategic choices of users who aim to maximize the traveling utility is formulated to optimize the incentives. A bi-level mixed-integer programming model is developed and reformulated into a single-level one, with a valid inequality introduced to improve computational efficiency. An improved particle swarm optimization algorithm is designed because the problem is proved to be NP-hard. Extensive numerical results validate the mathematical formulations and the solution method. Results also indicate that the relocation strategy improves both the operating profit and the number of served users in the configuration of the experiments. Several managerial insights are provided for service operators as well.**

***Index Terms*— Bi-level programming model, carsharing, particle swarm optimization, trip pricing, user-based relocation, valid inequality.**

I. Introduction

CARSHARING services have increasingly spread worldwide with the trend of mobility on demand. Specifically, the renting and returning sites of a car are unnecessarily the same in one-way mode of carsharing, which provides users with high flexibility [1]. However, such a mode usually leads to an imbalanced distribution of vehicles due to uncertain and asymmetric demands over time and space. The imbalances result in low reliability of the system and decrease the satisfaction degree of users [2], [3]. As a result, alleviating the imbalanced distribution of vehicles is a main operational challenge for service operators [4], [5], [6].

Received 3 August 2023; revised 9 March 2024 and 7 July 2024; accepted 7 October 2024. Date of publication 13 November 2024; date of current version 9 January 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 71971050 and Grant 72471052 and in part by the National Research Foundation of Korea (NRF) Grant funded by Korea Government (Ministry of Science and ICT) under Grant RS-2023-00218923. The Associate Editor for this article was H. M. A. Aziz. *(Ruiyou Zhang and Haiyu Kan are co-first authors.) (Corresponding author: Ilkyeong Moon.)*

Ruiyou Zhang is with the College of Information Science and Engineering, Northeastern University, Shenyang 110819, China (e-mail: zhangruiyou@ise.neu.edu.cn).

Haiyu Kan is with the College of Information Science and Engineering, Northeastern University, Shenyang 110819, China, and also with the School of Management, Shenyang University of Technology, Shenyang 110870, China (e-mail: kanhaiyu@sut.edu.cn).

Ilkyeong Moon is with the Department of Industrial Engineering and the Institute of Engineering Research, Seoul National University, Seoul 08826, Republic of Korea (e-mail: ikmoon@snu.ac.kr).

Digital Object Identifier 10.1109/TITS.2024.3488894

Operator-based and user-based relocation strategies can be applied to address the imbalances. In operator-based relocation strategies, designated staff are employed to relocate vehicles from areas with surplus vehicles to areas with deficient vehicles [7], [8], [9]. As a comparison, users are encouraged to relocate vehicles by adjusting their itineraries in user-based relocation strategies. User-based relocations are more sustainable than operator-based relocations by saving employment costs of staff and avoiding vehicle occupancies [10], [11]. However, they are insufficiently researched. The relocations carried out by users are more difficult to control than those implemented by staff which can be optimized by solving vehicle routing problems [12], [13], [14].

This research focuses on user-based relocations for non-reserved carsharing systems. An incentive-based trip pricing (IBTP) problem is investigated, in which incentives are determined to induce users to relocate vehicles. In a non-reserved system, users choose renting and returning stations freely without submitting any information on the trip in advance. Users relocate vehicles under incentives automatically and strategically to maximize their traveling utility, which brings a great challenge for operators to control the relocations. This research incorporates such behaviors of users into a bi-level programming model.

The motivations for this research are summarized as follows. Non-reserved systems, such as EVCARD in China (https://www.evcard.com/outlets/), are widely applied in practice. However, user-based relocations for non-reserved carsharing are seldom academically investigated due to the high complexity of controlling the relocations. User-based relocations are more easily applied in non-reserved carsharing than in reservation-based modes owing to the unnecessity of interacting with users. Methodologically, simulation- and heuristic-based methods are commonly employed in the literature of user-based relocations. However, exact approaches are scarcely reported, let alone those based on bi-level programming models (see Section II for a detailed review).

Dealing with the IBTP problem in non-reserved carsharing involves great challenges. First, the demands of car renting and car returning are both time-varying and uncertain. This feature requires incentives to be updated dynamically such that the supply can better match the forthcoming demand. Furthermore, incentives should be optimized in almost real time due to the requirement of dynamic updates. Second, formulating users' choices under incentives requires excessive variables and constraints. Third, it is rather difficult to obtain the optimal solutions for real-world instances due to the NP-hardness of the IBTP problem.

This research makes threefold contributions. (i) An IBTP problem is proposed formally for non-reserved carsharing, where incentives are determined to encourage users to relocate vehicles with a fine-grained consideration of users' choices. (ii) The NP-hardness of the IBTP problem is proved. A bi-level mixed-integer programming model is established and reformulated into a single-level one, with a valid inequality introduced. An improved particle swarm optimization (IPSO) algorithm is also developed. (iii) Several managerial insights, including the critical factors and appropriate circumstances for applying user-based relocations, are provided.

The remainder of this paper is organized as follows. Section II reviews the literature on user-based relocation strategies. Sections III and IV describe the IBTP problem formally and formulate it into a bi-level mixed-integer programming model, respectively. Section V reformulates the bi-level model and introduces the valid inequality. The IPSO algorithm is presented in Section VI. Section VII validates the proposed methods, followed by conclusions and possible extensions in Section VIII.

II. Literature Review

This section reviews the literature related to the topic of this research. Section II-A presents a brief review of user-based relocation strategies. Section II-B reviews the trip pricing problems extensively in the field of the user-based relocation strategy. Some user-based relocation strategies proposed for bike-sharing can also be applied to carsharing, and hence, are also included for a comprehensive review.

A. User-Based Relocation Strategies

Users are encouraged to adjust their renting and/or returning sites as well as the corresponding times in user-based relocation strategies. Some studies focus on identifying the most effective form of applying user-based relocations. For instance, [15] compares the benefits of adjusting stations and adjusting times in decreasing the fleet size. Reference [16] concludes that users prefer adjusting returning sites to adjusting renting sites based on stated preference investigations. However, [17] points out that adjusting renting sites is most easily accepted by users, according to an online survey in a city different from the city in [16]. One reason for these diverging conclusions may be the differences in extra distances (caused by site adjustment) and the incentives, besides the differences in city characteristics and cultural customs.

Some other researchers mainly focus on revealing the potential of user-based relocations, but ignore the effect of extra distances and incentives on users' choices of transportation modes or renting/returning sites. It is commonly assumed that the share of users participating in relocations is constant, whereas the incentives given to users are not considered necessarily [18], [19], [20]. In practice, users may not participate in relocations if the incentive cannot offset the extra cost caused by relocations. Hence, optimizing the amount of incentives is a critical issue in user-based relocations. The incentive forms include coupons [21] and direct monetary incentives [22], [23]. Surcharges may also be applied to the user whose trip aggravates the imbalances of vehicle distribution [24]. The applications of incentives and surcharges result in differentiated price rates, which prompts the trip pricing problem.

B. Trip Pricing Problems

In trip pricing problems, area- and period-dependent trip prices are optimized to affect the pattern of demands [1], [25]. Associated problems can be grouped into two classes roughly. The first class of problems focuses on the choices between vehicle sharing and the other transportation modes [26], [27], [28]. The second class of problems mainly concerns the effects on users' paths with shared vehicles [29], [30], [31]. This research falls into the second class and this subsection focuses also on the second class.

Most studies in the second class focus on providing solutions for reservation-based systems [32], [33], [34]. Users in a reservation-based system need to submit information on their desired trips in advance. The system assigns vehicles (or parking slots) or recommends alternative trips based on the submissions. The choice of users, i.e., whether to accept alternative trips, is commonly formulated using binomial logistic models [22], [35], [36]. Price discrimination occurs usually in reservation-based systems, because incentives depend on both the initially desired trip and the actual trip. Several users taking the same trips in the same period may be charged different fees because their initially desired trips are different [23], [37]. Such price discrimination may cause unfairness among users and loss of potential demands at the long-term level.

Non-reserved systems are rarely reported in the second class of problems despite their wide applications (as shown in Table I). Users in a non-reserved system are free to choose renting and returning stations without being recommended any alternative trips. Users have many choices on trips, and hence, binomial logistic models are unable to formulate their choices. Reference [21] employs a multi-logit model to formulate users' choices of trips under coupons. A rule-based strategy is proposed to update the coupons in an evolution model. Reference [29] applies the Monte Carlo simulation to formulate the choice of users under incentives for non-reserved bike-sharing.

As can be seen from Table I, trip pricing problems for user-based relocation strategies have attracted increasing attentions over recent years. However, the investigation into flexible non-reserved systems is insufficient. In addition, users' strategic behaviors have not been well considered. Some studies incorporating users' strategic choices employ either simulation- or heuristic-based approaches, whereas exact methods are seldom employed. This research enriches the literature by investigating an IBTP problem for non-reserved systems with the consideration of users' strategic choices. Both exact (based on a bi-level programming model) and meta-heuristic approaches are developed.

TABLE I
MAIN ARTICLES ON THE SECOND CLASS OF TRIP PRICING PROBLEMS

Article	Non-served mode ?	Approach	Strategic behavior [2] ?
Ref. [10]	N	Polynomial-time algorithm	×
Ref. [11]	N	Lyapunov-based model	√
Ref. [15]	N	Mixed-integer programming (MIP) model	×
Ref. [16]	N/A [1]	Stated preference	√
Ref. [17]	N/A	Stated preference	√
Ref. [18]	N	Two-stage heuristic	×
Ref. [19]	N	MIP model	×
Ref. [20]	N	Simulation-based optimization	×
Ref. [21]	Y	Evolution model	√
Ref. [22]	N	Simulation and Markovian chains	√
Ref. [23]	N	Online learning algorithm	√
Ref. [24]	N	Rolling-horizon and simulated annealing	√
Ref. [25]	N	Approximate dynamic programming-based approach	√
Ref. [29]	Y	Simulation	√
Ref. [30]	N	Simulation	×
Ref. [31]	N	MIP model and approximation algorithm	×
Ref. [32]	N	MIP model	√
Ref. [33]	N	Approximation algorithm	√
Ref. [34]	N	MIP model	×
Ref. [35]	N	Simulation and two-stage heuristic	√
Ref. [36]	N	MIP model and approximate dynamic programming	√
Ref. [37]	N	Online learning algorithm	√
This study	Y	Bi-level MIP model and IPSO algorithm	√

1. "N/A": not mentioned.
2. "Strategic behavior": users' strategically choosing renting/ returning sites under differentiated price rates for maximizing the traveling utility.

III. INCENTIVE-BASED TRIP PRICING PROBLEM

This section formally describes the IBTP problem from the following three viewpoints: the carsharing system, the incentive policy, and the strategic choices of users. Sets and parameters describing the problem are presented in Table II.

TABLE II
SETS AND PARAMETERS IN THE IBTP PROBLEM

Notation	Meaning
Sets:	
I	Set of stations in the system, which is period-independent
I_D	Set of car-deficient stations in the considered period
Parameters:	
M_1	Number of cars in the system
C_i	Number of specific parking slots at station i, which is period-independent
q_i	Estimated number of cars in demand at station i in the considered period
q_i'	Estimated number of cars in demand at station i in the next period
r_i^B	Number of cars available at station i at the beginning of the considered period
r_i^E	Estimated number of cars available at station i at the end of the period under Condition NR
s_i	Estimated number of rented cars at station i in the considered period
a_i	Estimated number of returned cars at station i in the considered period under Condition NR
$[L_i, U_i]$	Target inventory range at station i at the end of the considered period
d_{ij}	Distance from station i to station j
c^F	Cost of per unit additional distance exhibited by users
Δ	Cost of participating in relocations exhibited by users
f_1, f_2	Weightings of the DET part and the TIE part, respectively
P	Maximum incentive offered to a user

A. The Carsharing System

A one-way station-based carsharing system offers services within a given area. The system is composed of a set of stations (say I) and a fixed number M_1 of homogeneous cars. A given number of parking slots specific to the system and some extra parking slots exist at each station. If the number of shared cars parked at station $i \in I$ exceeds the number of specific parking slots C_i, additional cars can still be parked at extra slots of station i, but the system needs to pay for occupying extra slots.

The system provides users with high flexibility in renting and returning cars. Users are allowed to rent cars at any time over a planning horizon. Users' demands for cars can always be accommodated if cars are available at the renting stations. Users are also free to return the cars to any station in the service area without submitting any information in advance. The service is fully self-serviced via a service application, on which some important information such as incentives is updated in almost real time.

Imbalances in car distribution occur regularly due to spatially asymmetric demands over time. The system applies a user-based relocation (UBR) strategy to alleviate the imbalances, which induces users to adjust their returning stations via monetary incentives. The planning horizon (usually one day) is divided into multiple periods to address the fluctuation and uncertainty in demands. The IBTP problem for each period is solved to determine the incentives dynamically with the consideration of users' strategic choices. Sections III-B and III-C detail the incentive policy and strategic choice of users, respectively. Note that the initial distribution of cars in the first period is given as an outcome of the previous planning horizon.

B. The Incentive Policy

The relocations are proactive in the sense that they are implemented to better match supply with forthcoming demand. At the beginning of each period (excluding the last one), the system first estimates the number of cars available at each station at the end of the period without any relocations. Based on the estimation, the stations with insufficient cars (called car-deficient stations) at the end of the considered period can be easily found. Station-dependent incentives are offered at the car-deficient stations during the considered period to attract more users to return cars, such that the resulting car distribution at the end of the period can be better. For the convenience of description, the condition without any relocations is referred to as Condition NR. We assume that the number of cars in demand at each station in each period can be estimated based on historical data. Readers interested in demand estimation are directed to [28] and [38].

The number of cars available at station $i \in I$ at the end of the period under Condition NR is estimated as

$$r_i^E = r_i^B - s_i + a_i, \tag{1}$$

where r_i^B should be observable in practice and is obtained by simulation in the experiments, s_i is calculated as $\min\left\{r_i^B, q_i\right\}$, and a_i is estimated prior by simulating the operations under Condition NR. Station i is a *car-deficient station* if the estimated number of cars r_i^E is less than the estimated number of cars in demand q_i'. The set of car-deficient stations under Condition NR is given as

$$I_D = \left\{i \in I | r_i^E < q_i'\right\}. \tag{2}$$

The IBTP problem is formulated at the beginning of each period to determine the incentives at car-deficient stations. The objective is to minimize the weighted sum of the deviation from the estimated number of cars available at the end of the period after applying the incentives to the target inventory range at all stations (called the DET part), and the total incentive expense (called the TIE part). For a specific station at its target range, the car demand can be well satisfied, and meanwhile, parking slots over there are still available [29], [39]. Therefore, the *lower* (*upper*) *limit of the target inventory range* at station i at the end of the period is specified as the estimated number of cars in demand in the next period (the number of specific slots) at the station, i.e., $L_i = q_i'(U_i = C_i)$. We assume that $q_i' \le C_i$ for any station i at any period. For any station i, the *deviation* from the estimated number of cars available at the end of the period after applying the incentives, denoted as n_i, to the target inventory range $[L_i, U_i]$ is defined as: the distance between n_i to the closer endpoint of the range $[L_i, U_i]$ if n_i is out of the range; or zero otherwise. It can be expressed as $\max\{n_i - U_i, 0\} - \min\{n_i - L_i, 0\}$.

The system charges users based on the traveling distance. The service fee of a user returning a car to a station $i \in I$ is

$$c = c^{RE} w - p_i, \tag{3}$$

where c^{RE} is the fee of unit traveling distance, w is the traveling distance from the renting station to the returning station i, and p_i is the incentive of returning a car to station i in the returning period.

C. The Strategic Choices of Users

The incentives are assumed to affect users' choices on returning stations without influencing their willingness to enjoy the service [22], [23]. Users choose returning stations strategically to maximize their perceived utility. The *perceived utility* is defined as the received incentive minus the sum of the extra cost due to station adjusting and the inertia cost. We assume that station adjustments do not change the returning period, and that users consider only the incentives in the returning period. If a user who initially desires station $i \in I$ finally returns a car to the station $j^* \in I \backslash \{i\}$, we have

$$j^* = \underset{j \in I\backslash\{i\} p_j - c^F d_{ij} - \Delta \ge p_i}{\operatorname{argmax}} \left(p_j - c^F d_{ij} - \Delta\right), \tag{4}$$

where d_{ij} is the distance from station i to station j, c^F is the cost of per unit additional distance exhibited by users (referred to as the *flexibility cost*), and Δ is the cost of participating in relocations (referred to as the *inertia cost*). That is, a user chooses the station with the highest perceived utility among all the choices with nonnegative perceived utilities.

Assumption 1 [40]: Both the flexibility cost and the inertia cost are homogeneous among the users.

The returning station of a user may not be his/her final destination in practice. For a user adjusting the returning station, the time from the initially desired returning station to the final destination and the time from the actual returning station to the final destination may be different slightly. The service fees for arriving at the initially desired returning station and for arriving at the actual returning station may also be different slightly. However, the above differences are difficult to formulate precisely. We simplify the extra cost due to station adjusting as the product of the flexibility cost and the distance from the initially desired returning station to the actual returning station, as in [34] and [35].

IV. Bi-Level Mixed-Integer Programming Model

The IBTP problem can be formulated as a leader-follower game. The system first determines the incentives, followed by the users choosing their returning stations. This section formulates this game into a bi-level mixed-integer programming model (called *Model B-MIP*).

A. Upper-Level Model

The upper level formulates the decision of the system, with the decisions of users treated as given parameters. The following variables were introduced:

Decision variable:

$p_j\ (\ge 0)$	Incentive offered to a user returning a car to station $j \in I_D$ in the considered period.

Auxiliary variables:

n_i	Estimated number of cars available at station $i \in I$ at the end of the considered period after applying the incentives;
e_i	1 if $n_i \le L_i$; or 0 otherwise;
o_i	1 if $n_i \ge U_i$; or 0 otherwise.

The decision of the system can be formulated as the following mixed-integer programming model (called *Model B-MIPU*).

Objective function (5), shown at the bottom of the next page, minimizes the weighted sum of the DET part and the TIE part. $(n_i - U_i)\, o_i + (L_i - n_i)\, e_i$ is equivalently a tractable formulation of $\max\{n_i - U_i, 0\} - \min\{n_i - L_i, 0\}$. $p_j\left(a_j + \sum_{i \in I} a_i m_{ij} - a_j \sum_{i \in I_D} m_{ji}\right)$ is the incentive expense at station j, where m_{ij} depends on the decision of users to be formulated in Section IV-B. If a user who initially desires station $i \in I$ finally returns a car to station $j \in I_D$, $m_{ij} = 1$; otherwise, $m_{ij} = 0$. $a_j + \sum_{i \in I} a_i m_{ij} - a_j \sum_{i \in I_D} m_{ji}$ represents the estimated number of returned cars at station j under the incentives, where $\sum_{i \in I} a_i m_{ij}$ is the sum of estimated returned cars switching to station j, and $a_j \sum_{i \in I_D} m_{ji}$ is the sum of estimated returning cars switching from station j to stations belonging to set I_D.

Constraints (6) and (7), shown at the bottom of the page, are the flow conservation constraints under the incentives. Constraint (6) estimates the number of cars available at the car-deficient stations at the end of the period. Constraint (7) calculates those at stations $I \backslash I_D$, where $a_i - a_i \sum_{j \in I_D} m_{ij}$ is the estimated number of returned cars at station i under the incentives. Constraint (8), shown at the bottom of the page, determines the relationship between n_i and L_i. If $e_i = 1$, Constraint (8) becomes $-M_1 + L_i \le n_i \le L_i$, with its left-hand side relaxed. Otherwise, i.e., if $e_i = 0$, Constraint (8) becomes $L_i \le n_i \le M_1 + L_i$, with its right-hand side relaxed. Constraint (9), shown at the bottom of the page, determines the relationship between n_i and U_i. Constraints (10) and (11), shown at the bottom of the page, restrict the ranges of the variables.

Model B-MIPU can be extended by incorporating operator-based relocations. For instance, the relocations implemented by staff and the scheduling of staff can be integrated with the user-based relocations. However, the IBTP problem itself is quite difficult. An integrated optimization will be extremely complicated. Collaborative optimization of user- and operator-based relocations is one of our research directions in the future.

B. Lower-Level Model

The lower level formulates users' choices on returning stations to maximize their perceived utilities under given incentives. It treats m_{ij} $(i \in I, j \in I_D)$ as decision variables and p_j $(j \in I_D)$ as given parameters. The decision of users is formulated as the following linear programming model (called *Model B-MIPL*):

$$\max \sum_{j \in I_D} \sum_{i \in I} \left(p_j - c^F d'_{ij} \right) m_{ij} \tag{12}$$

$$\text{s.t.} \sum_{j \in I_D} m_{ij} \le 1, \quad \forall i \in I \tag{13}$$

$$m_{ij} \ge 0, \quad \forall i \in I, \forall j \in I_D \tag{14}$$

Objective function (12) maximizes the perceived utility of users, where

$$d'_{ij} = \begin{cases} d_{ij}, & \text{if } i = j, \\ d_{ij} + \Delta / c^F, & \text{otherwise} \end{cases}$$

Constraint (13) guarantees that a user who initially desires station i finally returns a car to at most one car-deficient station. Constraint (14) relaxes the binary variable m_{ij} as a positive continuous one. The relaxation does not change the optimality of a solution to Model B-MIPL because the right-hand side of Constraint (13) are integers and the coefficient matrix of the left-hand side forms a totally unimodular matrix [40].

Model B-MIPL does not involve any specific user, because the users initially desiring station $i \in I$ behave consistently in selecting the final returning station under Assumption 1.

Theorem 1: The IBTP problem is NP-hard (see Appendix for the proof).

V. The Single-Level Mixed-Integer Model and Valid Inequality

Model B-MIP cannot be handled directly by state-of-the-art optimizers due to its bi-level nature. Section V-A reformulates it into a single-level model, with the nonlinear constraints linearized in Section V-B. Section V-C introduces a valid inequality to improve computational efficiency.

A. Single-Level Mixed-Integer Programming Model

Model B-MIPL is convex with a linear objective function and linear constraints. Therefore, the Karush-Kuhn-Tucker (KKT) conditions are both sufficient and necessary to judge the optimality of its solutions [41]. We replace it with its KKT conditions and integrate them with Model B-MIPU. A single-level model (called *Model S-MIP*) is obtained, with the objective function (5), constraints (6)-(11), (13)-(14) and

$$c^F d'_{ij} - p_j + \mu_i - \lambda_{ij} = 0, \quad \forall i \in I, \forall j \in I_D \tag{15}$$

$$\mu_i \left(\sum_{j \in I_D} m_{ij} - 1 \right) = 0, \quad \forall i \in I \tag{16}$$

$$-\lambda_{ij} m_{ij} = 0, \quad \forall i \in I, \forall j \in I_D \tag{17}$$

B-MIPU:

$$\min F = f_1 \sum_{i \in I} ((n_i - U_i) o_i + (L_i - n_i) e_i) + f_2 \sum_{j \in I_D} p_j \left(a_j + \sum_{i \in I} a_i m_{ij} - a_j \sum_{i \in I_D} m_{ji} \right) \tag{5}$$

$$\text{s.t. } n_j = r_j^B - s_j + a_j + \sum_{i \in I} a_i m_{ij} - a_j \sum_{i \in I_D} m_{ji}, \quad \forall j \in I_D \tag{6}$$

$$n_i = r_i^B - s_i + a_i - a_i \sum_{j \in I_D} m_{ij}, \quad \forall i \in I \backslash I_D \tag{7}$$

$$-M_1 e_i + L_i \le n_i \le M_1 (1 - e_i) + L_i, \quad \forall i \in I \tag{8}$$

$$-M_1 o_i + n_i \le U_i \le M_1 (1 - o_i) + n_i, \quad \forall i \in I \tag{9}$$

$$0 \le p_j \le P, \quad \forall j \in I_D \tag{10}$$

$$n_i \ge 0, e_i \in \{0, 1\}, o_i \in \{0, 1\}, \quad \forall i \in I \tag{11}$$

$$0 \le \mu_i \le P, \quad \forall i \in I \tag{18}$$

$$0 \le \lambda_{ij} \le \max_{i\in I, j\in I_D} c^F d'_{ij} + P, \quad \forall i \in I, \forall j \in I_D \tag{19}$$

$$m_{ij} \in \{0,1\}, \quad \forall i \in I, \forall j \in I_D \tag{20}$$

Constraints (15)-(19), with the primal feasibility constraints (13) and (14), are the KKT conditions of Model B-MIPL. Constraint (15) is the stationary equation, where μ_i and λ_{ij} are the dual variables. Constraints (16) and (17) are the complementary slackness conditions of Constraints (13) and (14), respectively. Constraints (18) and (19) restrict the ranges of the dual variables. For a feasible solution with $m_{ij} = 1$, λ_{ij} should be zero to satisfy Constraint (17). According to Constraint (15), μ_i should be $p_j - c^F d'_{ij}$, and bounded up to P. Otherwise, for a feasible solution with $m_{ij} = 0$, μ_i may be zero or larger than zero, depending on the value of $\sum_{j\in I_D} m_{ij}$ in Constraint (16). According to Constraint (15), λ_{ij} should be $c^F d'_{ij} - p_j + \mu_i$, and bounded up to $\max_{i\in I, j\in I_D}\left(c^F d'_{ij} + \mu_i\right)$. Therefore, a feasible upper-bound combination of the dual variables are $\mu_i \le P$ and $\lambda_{ij} \le \max_{i\in I, j\in I_D} c^F d'_{ij} + P$ for any $i \in I$ and $j \in I_D$. Constraint (20) strengths the restriction of the variable m_{ij} to maintain the optimality of a solution to Model B-MIP.

B. Linearization of Constraints (16) and (17)

It is still difficult to solve Model S-MIP directly due to the nonlinearity in Constraints (16) and (17). The nonlinear constraints are linearized by introducing auxiliary binary variables. Specifically, Constraint (16) can be replaced by

$$\mu_i \le M_2 \alpha_i, \quad \forall i \in I \tag{21}$$

$$1 - \sum_{j\in I_D} m_{ij} \le M_2 (1 - \alpha_i), \quad \forall i \in I \tag{22}$$

$$\alpha_i \in \{0,1\}, \quad \forall\, i \in I \tag{23}$$

where $M_2 = \max\{P, 1\}$. If $\alpha_i = 1$, Constraint (21) is relaxed, and Constraint (22) becomes $1-\sum_{j\in I_D} m_{ij} \le 0$. $\sum_{j\in I_D} m_{ij} - 1$ is forced to be zero to satisfy Constraints (13) and (22), and hence Constraint (16) holds. Otherwise, Constraint (21) becomes $\mu_i \le 0$, and Constraint (22) is relaxed. In such a case, μ_i is compelled to be zero to satisfy Constraints (18) and (21), and Constraint (16) holds. Similarly, Constraint (17) can be replaced by

$$\lambda_{ij} \le M_3 \beta_{ij}, \quad \forall i \in I, \forall j \in I_D \tag{24}$$

$$m_{ij} \le M_3 \left(1 - \beta_{ij}\right), \quad \forall i \in I, \forall j \in I_D, \tag{25}$$

$$\beta_{ij} \in \{0,1\}, \quad \forall i \in I, \forall j \in I_D \tag{26}$$

where $M_3 = \max\left\{\max_{i\in I, j\in I_D} c^F d'_{ij} + P, 1\right\}$.

In summary, Model B-MIP is reformulated equivalently into a new model with Objective function (5) and Constraints (6)-(11), (13)-(15), (18)-(26) (*Model S-MIP-N*).

C. The Valid Inequality

Model S-MIP-N can be solved by commercial solvers such as CPLEX if the problem scale is not too large. An inequality is introduced to improve the computational efficiency as follows:

$$c^F d'_{ij} m_{ij} \le p_j, \quad \forall i \in I, \ \forall j \in I_D. \tag{27}$$

Inequality (27) excludes the solution space where users' perceived utilities are smaller than zero. Specifically, if $c^F d'_{ij} > p_j$, adjusting the returning station from station $i \in I$ to station $j \in I_D$ results in a negative perceived utility. In such a case, m_{ij} is compelled to be zero. We refer to Model S-MIP-N with Inequality (27) as *Model S-MIP-I*.

VI. IPSO Algorithm

Model S-MIP-I is still computationally intractable for large-scale instances due to the NP-hardness of the problem. As a result, we develop an IPSO algorithm, motivated by the outstanding ability of PSO algorithms to address continuous optimization problems that include various planning and operational problems in carsharing [20], [42]. Similarly as in [43], starting from an initial population, the IPSO algorithm iterates until either the given maximum number of iterations (say N^{ITE}) or the given computation time (say T^{ITE}) is reached.

A. Encoding and Generation of the Initial Population

A solution is a vector with $|I_D|$ real elements within the range $[0, P]$, where $|\cdot|$ is the size of a given set. Each element denotes an incentive of returning a car to a car-deficient station. Given a population size N^{POP}, the solution corresponding to particle $l(\le N^{POP})$ is represented by $\boldsymbol{p_l}$, where the uth dimension p_{lu} represents the incentive of returning a car to the uth car-deficient station (denoted as $\pi(u)$).

For any station $i \in I$, we can sort the stations $I\setminus\{i\}$ in ascending order of distances from them to station i. The sorting result is named as the vector of neighborhood stations of station i and denoted as $\boldsymbol{E}_i$, where E_{ih} $(h \ge 1)$ is the hth element in $\boldsymbol{E}_i$.

The initial solutions are generated using the vectors of neighborhood stations of the car-deficient stations. They can be classified into *refined* ones and *coarse* ones. For a refined one $\boldsymbol{p_l}$, the incentive p_{lu} is applied to induce the users who initially desire station $E_{\pi(u)l}$ to return cars to the car-deficient station $\pi(u)$, if doing so does benefit. The refined initial solutions ensure good starting points. For a coarse one $\boldsymbol{p_l}$, the incentive p_{lu} is adopted to induce the users who initially desire station $E_{\pi(u)h}$ to return cars to station $\pi(u)$, where the station $E_{\pi(u)h}$ is selected randomly. Coarse initial solutions ensure the population diversity. Given the number N^{I_REF} of the refined initial solutions, the procedure of generating the initial population is detailed as follows.

Step 1: Initialization. Let $\boldsymbol{p}_0 = [0, 0, \cdots, 0]_{|I_D|}$, and $l = u = h = 1$. Let $\boldsymbol{p}'_{\boldsymbol{l}} = \boldsymbol{p}_l = \boldsymbol{p}_0$, and $C_l = f(\boldsymbol{p}_0)$, where $f(\cdot)$ is the objective value of the given solution.

Step 2: If station $i = E_{\pi(u)h}$ is not a car-deficient station, let

$$p'_{lu} = \begin{cases} c^F\ d_{i\pi(u)} + \Delta, & \text{if } c^F\ d_{i\pi(u)} + \Delta \le P, \\ p_{lu}, & \text{otherwise}, \end{cases} \tag{28}$$

otherwise, let $p'_{lu} = p_{lu}$.

Step 3: If $p'_{lu} > p_{lu}$ and $h > 1$, for any station i in the set $I_D \cap \left(\bigcup_{g=1}^{h-1} E_{\pi(u)g}\right)$, let

$$p'_{l\pi^{-1}(i)} = \max\left\{p'_{lu} - c^F d_{i\pi(u)} - \Delta, p_{l\pi^{-1}(i)}\right\}, \quad (29)$$

where $\pi^{-1}(i)$ is the index of car-deficient station i in the solution. If $\boldsymbol{p}_l \neq \boldsymbol{p}'_l$ and $f\left(\boldsymbol{p}'_l\right) \le C_l$, let $C_l = f\left(\boldsymbol{p}'_l\right)$ and $\boldsymbol{p}_l = \boldsymbol{p}'_l$; if $\boldsymbol{p}_l \neq \boldsymbol{p}'_l$ and $f\left(\boldsymbol{p}'_l\right) > C_l$, let $\boldsymbol{p}'_l = \boldsymbol{p}_l$.

Step 4: If $h < l$ and $u \le |I_D|$, let $h = h + 1$ and return to Step 2; if $h = l$ and $u < |I_D|$, let $h = 1$, $u = u + 1$, and return to Step 2; otherwise, let $h = 1, u = 1$, and $l = l + 1$.

Step 5: If $l \le N^{I\_REF}$, let $\boldsymbol{p}'_l = \boldsymbol{p}_l = \boldsymbol{p}_0$, $C_l = f\left(\boldsymbol{p}_0\right)$, and return to Step 2.

Step 6: Generate a random positive integer h bounded up to a given parameter N^{NEI}, and let $i = E_{\pi(u)h}$. Let

$$p_{lu} = \begin{cases} c^F d_{i\pi(u)} + \Delta, & \text{if } c^F d_{i\pi(u)} + \Delta \le P, \\ P, & \text{otherwise}, \end{cases} \quad (30)$$

and let $u = u + 1$.

Step 7: If $u \le |I_D|$, return to Step 6; otherwise, let $u = 1$ and $l = l + 1$. If $l \le N^{POP}$, return to Step 6; otherwise, output the initial solutions and the procedure stops.

The initial velocity for each dimension of each solution is randomly generated in the range $[-V, V]$, where V is the maximum velocity.

B. Update of the Solutions

The entire solutions are first updated with linearly varying inertial weighting and asynchronous time-varying learning factors. With the increment in the number of iterations, the values of both the inertial weighting and the self-learning factor decrease linearly; the value of the social-learning factor increases linearly. The IPSO algorithm can possess a better global exploration ability in the earlier phase and a better convergence ability in the later phase.

Each dimension of each solution requires an update. Given the best-so-far solution $\boldsymbol{P}^G$ for the population and the best-so-far solution $\boldsymbol{P}_l^{IND}$ for particle l, the velocity $v_{lu}(\kappa+1)$ of the uth dimension of particle l in the $(\kappa+1)$ th iteration is updated as

$$\begin{aligned} &v_{lu}(\kappa+1) \\ &= \left(\varpi_1^{INT} - \omega\left(\varpi_1^{INT} - \varpi_2^{INT}\right)\right) v_{lu}(\kappa) \\ &\quad + \theta\left(\varpi_1^{SEL} - \omega\left(\varpi_1^{SEL} - \varpi_2^{SEL}\right)\right)\left(P_{lu}^{IND} - p_{lu}(\kappa)\right) \\ &\quad + \theta\left(\varpi_1^{SOC} + \omega\left(\varpi_2^{SOC} - \varpi_1^{SOC}\right)\right)\left(P_u^G - p_{lu}(\kappa)\right), \end{aligned} \quad (31)$$

where $p_{lu}(\kappa)$ represents the uth dimension of solution l in the κth iteration; θ is a random number in the range [0, 1]; $\omega = (\kappa+1)/N^{ITE}$ is varying over the iterations; ϖ_1^{INT} and $\varpi_2^{INT} (< \varpi_1^{INT})$ are the initial and final values of the inertial weighting, respectively; ϖ_1^{SEL} and $\varpi_2^{SEL}\left(< \varpi_1^{SEL}\right)$ are the initial and final values of the self-learning factor, respectively; ϖ_1^{SOC} and $\varpi_2^{SOC}\left(> \varpi_1^{SOC}\right)$ are the initial and final values of the social-learning factor, respectively. The updated velocity is truncated if it exceeds the range $[-V, V]$. Furthermore, we have

$$\begin{aligned} &p_{lu}(\kappa+1) \\ &= \begin{cases} 0, & \text{if } p_{lu}(\kappa) + v_{lu}(\kappa+1) \le 0, \\ P, & \text{if } p_{lu}(\kappa) + v_{lu}(\kappa+1) \ge P, \\ p_{lu}(\kappa) + v_{lu}(\kappa+1), & \text{otherwise}. \end{cases} \end{aligned} \quad (32)$$

Some incentives in a solution can be slightly reduced whereas the adjustment of returning stations remains. That is, some solutions can be further improved by avoiding unnecessary incentive expenses. A so-called elitist modification strategy is applied to potentially improve the solutions. Specifically, when the entire particles are arranged in ascending order of the objective values of their best-so-far solutions, only the first $\lfloor N^{IMP} N^{POP} \rfloor$ particles are selected to avoid premature convergence. Here, $N^{IMP}\left(0 \le N^{IMP} \le 1\right)$ is the probability of the modification, and $\lfloor\cdot\rfloor$ represents the largest integer not larger than the given value. Each element of each selected solution needs to be modified. For the incentive $p_{lu}(\kappa+1)$, there exists a set of stations $Q_{lu}(\kappa+1)$. Users may adjust the returning station from stations in the set $Q_{lu}(\kappa+1)$ to station $\pi(u)$, i.e.,

$$\begin{aligned} &Q_{lu}(\kappa+1) \\ &= \left\{i \mid p_{lu}(\kappa+1) - c^F d_{i\pi(u)} - \Delta \ge 0, i \in I \backslash \{\pi(u)\}\right\}. \end{aligned} \quad (33)$$

If $Q_{lu}(\kappa+1)$ is empty, let $p_{lu}(\kappa+1) = 0$; otherwise, let

$$p_{lu}(\kappa+1) = \max_{i \in Q_{lu}(\kappa+1)} \left(c^F d_{i\pi(u)} + \Delta\right). \quad (34)$$

C. Decoding

Model B-MIP is applied to calculate the objective value. Given a solution, the adjustment of returning stations is first calculated based on Formula (4). Subsequently, the numbers of cars available at stations are calculated based on Constraints (6) and (7). Finally, the objective value can be obtained.

VII. Validation and Evaluation

This section validates the proposed methods based on numerical experiments. Section VII-A presents the settings of experiments and instances. Section VII-B validates the mathematical formulations and the IPSO algorithm. The proposed UBR strategy is evaluated in Section VII-C.

A. Setting of Experiments and Instances

All the experiments were run on an Intel I5-7300HQ quad-core CPU at 3.60 GHz and 8.0 GB RAM in a Windows 10 operating system. The mathematical models were solved using IBM ILOG CPLEX 12.6, where the *depth-first search* strategy was applied in a 2048 MB working memory. The C++ language in Visual Studio 2013 was used to call CPLEX and code the IPSO algorithm.

1) Generation of the Instances: Instances of three scales (i.e., small, medium, and large) were investigated. Some data such as data on demand involve business privacy and usually are not publicly available. Synthetic instances are commonly applied in various carsharing problems [2], [7], [44], [45]. We generated the instances according to [7], [44], and [46]. Eight instances were generated for each scale, with the instances named in the scale and number (such as S1, M1, and L1).

The service areas were square regions. The side lengths were selected in the sets $\{7, 9\}$ km, $\{10, 12, 13\}$ km and $\{12, 13\}$ km in the small-, medium- and large-scale instances, respectively. Stations were scattered randomly in the service area. The numbers of stations were in the ranges [40, 100], [130, 200] and [210, 250] in the three scales of instances sequentially. In each instance, the numbers of specific parking slots among stations followed normal distribution $N\left(20, 5^2\right)$. The initial numbers of cars followed $N\left(15, 5^2\right)$. For station pair i and j $(\neq i)$, the traveling distance d_{ij} was the Euclidean distance multiplied by a random number in the range [1, 2] to incorporate the impact of road layouts, as in [2] and [46]. Traveling times were calculated based on an average traveling speed of 25 km/h. For the trips with returning stations coincided with renting stations, the traveling times were set to one hour.

The planning horizon ranged from 6:00 a.m. to 10:00 p.m., where 6:00 a.m.-9:00 a.m. and 5:00 p.m.-8:00 p.m. were peak traffic time frames. The length of a period was set as 15 minutes. The demand at each station in each period was assumed to follow the Poisson distribution [4], [27]. Reference [7] found that the demand rate for cars during off-peak hours in a working day was approximately 0.7 times the demand rate during peak hours. The mean of cars in demand at each station was a random integer in the range [0, 25] in each period within peak traffic time frames, and was a random integer in the range [0, 20] in each of the other periods. The mean of cars in demand at each station was regarded as the estimated number of cars in demand when optimizing the incentives. The operations of 90 planning horizons under Condition NR were simulated. The obtained average number of returned cars at station i was estimated as the number of returned cars a_i.

The flexibility cost c^F was set, unless stated explicitly, as 10¥/km, referring to [44] which set parameters according to a report from EVCARD in China. The inertia cost Δ was set as ¥1. For a considered period, f_1 was set as $\left(\sum_{i\in I} D_i^0\right)^{-1}$, where $D_i^0 = \max\left\{r_i^E - U_i, 0\right\} - \min\left\{r_i^E - L_i, 0\right\}$. f_2 was set as f_1/f^{RA}, where $f^{RA} = 10$¥/vehicle, unless stated explicitly. The maximum incentive P offered to a user was ¥30.

2) Setting of Parameters in the IPSO Algorithm: The parameters in the IPSO algorithm include ϖ_1^{INT}, ϖ_2^{INT}, ϖ_1^{SEL}, ϖ_2^{SEL}, ϖ_1^{SOC}, ϖ_2^{SOC}, V, T^{ITE}, N^{ITE}, N^{POP}, N^{I_REF}, N^{NEI}, and N^{IMP}. We adopted the recommended values in [43] for the inertial weighting, learning factors, and the maximum velocity, i.e., $\varpi_1^{INT} = 0.9$, $\varpi_2^{INT} = 0.4$, $\varpi_1^{SEL} = 2.5$, $\varpi_2^{SEL} = 0.5$, $\varpi_1^{SOC} = 0.5$, $\varpi_2^{SOC} = 2.5$, $V = P - ¥0 = ¥30$. The computation time T^{ITE} was set to one minute. The parameters N^{ITE}, N^{POP}, N^{I_REF}, N^{NEI} and N^{IMP} were tuned sequentially using a popular method as in [47] based on the eight large-scale instances. The tuning results were $N^{ITE} = 500$, $N^{POP} = 150$, $N^{I\_REF} = 10$, $N^{NEI} = 10$, and $N^{IMP} = 0.1$.

TABLE III
RESULTS OF MODEL S-MIP-N AND MODEL S-MIP-I

Scale	Model	Avg. optimality gap (%)	Avg. computation time (s)
Small	S-MIP-N	0.00	32.34
	S-MIP-I	0.00	7.94
Medium	S-MIP-N	3.14	971.24
	S-MIP-I	0.00	32.20
Large	S-MIP-N	19.94	3,184.74
	S-MIP-I	1.33	2,094.12

B. Validation Based on a Single Period

This section validates the mathematical formulations and the IPSO algorithm. The third period was considered for all the instances.

1) Performance of Model S-MIP-N and Inequality (27): Each instance was solved using CPLEX based on Model S-MIP-N and Model S-MIP-I separately, with the computation time limited to one hour. Model S-MIP-N obtains the optimal solutions for all the small-scale and five medium-scale instances. Ten of the above 13 instances were solved to optimality even within one minute. The above results indicate the effectiveness of Model S-MIP-N.

The comparative results in Table III validate Inequality (27). The averaged computation time of Model S-MIP-I decreases significantly for each scale compared to that of Model S-MIP-N. Nineteen instances (including all the small- and medium-scale instances and three large-scale instances) obtain the optimal solutions by Model S-MIP-I. Fifteen of the above 19 instances obtain the optimal solutions even within one minute.

2) Performance of the IPSO Algorithm: The effectiveness and scalability of the IPSO algorithm are explored by comparing them to the performance of Model S-MIP-I. The algorithm was run three times independently for each instance, with the best solution presented. The solution time of CPLEX was limited to one minute for actual applications. Fig. 1 shows the computation times of the IPSO algorithm and of CPLEX for each instance, which indicates the high efficiency of the algorithm.

The IPSO algorithm provides the optimal solutions for instances S1-S8, M1-M6 and M8. For instance M7, the solution provided by the algorithm is worse than the optimal solution by 0.11%. For the large-scale instances, the improvements in the solutions provided by the IPSO algorithm compared to those obtained by CPLEX within one minute and one hour, i.e., Obj. Imp1 and Obj. Imp2, and the optimality gap

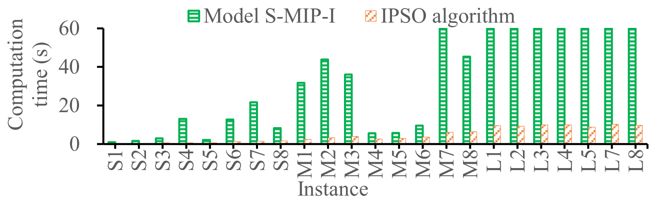
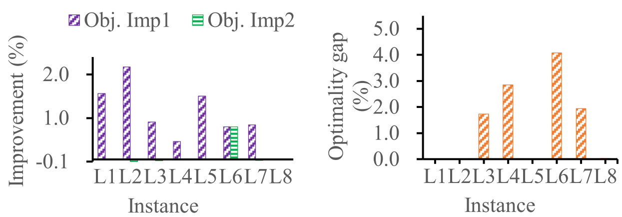


Fig. 1. Computation times of Model S-MIP-I and the IPSO algorithm for the three-scale instances.



(a) Improvement in the objective value (b) Optimality gap

Fig. 2. Comparison of solutions obtained by Model S-MIP-I and the IPSO algorithm for the large-scale instances.

given by CPLEX within one hour are shown in Fig. 2. The Obj. Imp1 for instances L1-L8 is 0.99% on average. For instances L2-L4 and L7, the solutions provided by the IPSO algorithm are slightly worse than those obtained by CPLEX within one hour. However, the solutions provided by the algorithm are still acceptable owing to the small optimality gaps and the small values of Obj. Imp2.

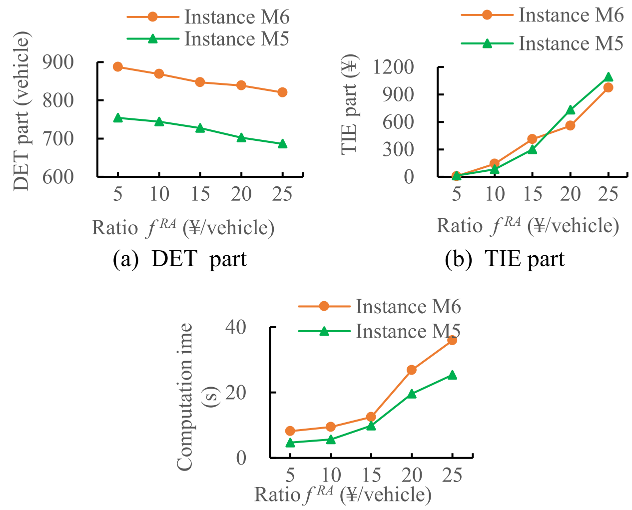
The robustness of the IPSO algorithm is investigated as well based on the large-scale instances. Results indicate that the IPSO algorithm is fairly stable. The objective values are invariant in the three repeats of six instances. The maximum relative difference of the three objective values is 0.03% in the remaining two instances.

3) Sensitivity Analysis of the Ratio f^{RA}: The ratio $f^{RA} = f_1/f_2$ represents the degree of the trade-off between the two parts in the objective function (5). We changed the ratio f^{RA} from 10 ¥/vehicle to 5 ¥/vehicle, 15 ¥/vehicle, 20 ¥/vehicle and 25 ¥/vehicle to conduct a sensitivity analysis. Instances M5 and M6 were selected for analysis, with Model S-MIP-I applied to obtain solutions.

As can be seen from Fig. 3, the system puts more emphasis on decreasing the DET part with the increment of the ratio f^{RA}. The decrement of the DET part is accompanied by the increment in the TIE part. A larger ratio f^{RA} also leads to a longer computation time because of a more complicated structure of incentives.

C. Validation of the User-Based Relocation Strategy

This section validates the UBR strategy based on the Monte Carlo simulation in the rolling-horizon framework. The incentives were updated dynamically from the first to the second-to-last period in a planning horizon, where the number of cars available at station $i \in I$ at the beginning of a period is an outcome of the previous period. Five planning horizons were simulated for each instance, with the average value of a performance metric approximated as the expectation. The following performance metrics are considered:



(a) DET part (b) TIE part

(c) Computation time for obtaining the optimal solution

Fig. 3. The effect of the ratio f^{RA}.

(1) The operating profit is the total revenue (calculated in Eq.(3)) minus the sum of the fuel consumption cost and the cost of occupying extra slots. The distance-based fee was obtained using a setting of ¥2.4 per kilometer. The fuel consumption cost was ¥0.3 per minute when cars were driven on the road. The above settings were referred from [44], as was the setting of the flexibility cost. The cost of occupying extra slots was calculated based on a cost of ¥1.0 per period for an extra slot.

(2) The number of served users is the sum of served users at all stations over the planning horizon.

1) Improvements in the System Performance: Results in Section VII-B.3 indicate that an appropriate ratio f^{RA} may improve both the operating profit and the number of served users, owing to a balance between the incentive expense and the increment in the number of served users. However, appropriate ratios for different systems may be disparate due to the differences in operational characteristics. We changed the ratio f^{RA} from 10 ¥/vehicle to 8¥/vehicle, 15¥/vehicle, 20¥/vehicle and 25 ¥/vehicle for each instance, and simulated the operations under each value by the Monte Carlo simulation with the other settings remained. The highest operating profit under the five values and the corresponding number of served users were adopted as the performance of the UBR strategy. The incentives for instances S1-S8 and M4-M6 were optimized using Model S-MIP-I with the computation time limited to one minute. The incentives for the other instances were obtained using the IPSO algorithm.

The UBR strategy has great potential for enhancing system performance and is scalable (as shown in Fig. 4). The increment in the operating profit varies from 2.67% to 30.70%, with the average increments reaching 18.62%, 5.65%, and 9.85% for the small-, medium- and large-scale instances, respectively. The numbers of served users increase by 20.37%, 8.12%, and 9.87% on average for the three-scale instances. More users are served owing to the incentive-based relocations, and the profit obtained from the additionally served users compensates well for the incentive expense. The ratios f^{RA*} obtaining the highest operating profit are disparate among the instances,

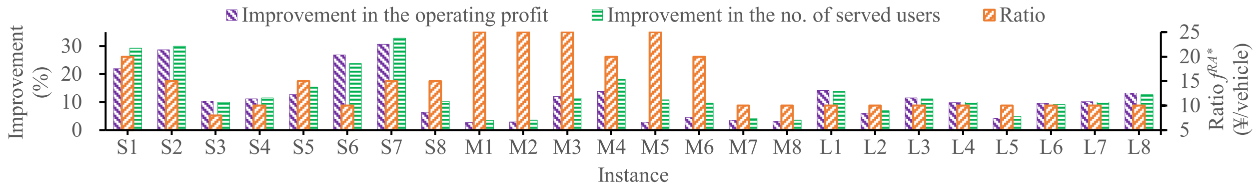


Fig. 4. Performance of the UBR strategy ("Improvement" is the improvement in the system performance under the UBR strategy compared to that under Condition NR).

which indicates the importance of properly balancing the incentive expense and the increased served users for each system.

2) Comparison to Operator-Based Relocations: The UBR strategy is compared with a widely considered operator-based relocation strategy. The incentive-based relocations in the UBR strategy are assumed to be implemented by staff after users return cars to their initially desired stations. The operator-based relocations are assumed to be finished by the end of each period. According to [35] and [40], the cost of relocating a car by staff is

$$c_v = c_0 + c_f d / v_s, \quad (35)$$

where c_0 is the employment cost of one staff for a period, c_f is the fuel consumption cost per unit time, d is the traveling distance of the relocation, and v_s is the traveling speed. The settings of d and v_s can be found in Section VII-A.1, and the setting of c_f can be found in this section. We set c_0 as ¥ 16.25 per period, referring to [44].

Results indicate that in 20 out of the 24 instances, the UBR strategy outperforms the operator-based relocation strategy, with the incentive expense lower than the cost of relocating the equivalent cars by staff. The incentive expense is approximately 0.72 times the cost of operator-based relocations on average for the 20 instances. For the other four instances, the former is approximately 1.25 times the latter on average.

3) Effect of the Flexibility Cost: The flexibility cost influences users' choice on returning stations and further affects the performance of the UBR strategy. Therefore, we changed the flexibility cost from 10 ¥/km to 8 ¥/km, 15¥/km, 20 ¥/km and 30 ¥/km, and ran the Monte Carlo simulation under each value of the flexibility cost, with the details described in Section VII-C.1. Instances S5, M4, and L5 were selected for the analysis. Fig. 5 presents the improvements in the operating profit obtained by the UBR strategy (compared to the operating profit under Condition NR) and the ratio f^{RA*}.

Fig. 5 (a) indicates the robustness and scalability of the UBR strategy. An improvement in the operating profit can always be guaranteed for each instance within a large range of the flexibility cost. It can also be observed that the improvement in the operating profit decreases with the increment in the flexibility cost in most cases. The above result is expected because the system needs to pay more incentives to improve the distribution of cars when users exhibit higher flexibility costs. However, the improvements in the operating profit slightly increase in a few cases, for example, when the flexibility cost increases from 20 ¥/km to 30 ¥/km at Instance M4. Such an outcome might result from the estimation error

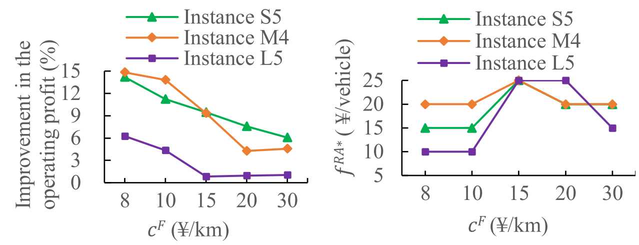


Fig. 5. Effect of the flexibility cost.

in the number of rented/returned cars or a coarse-grained setting of the ratio f^{RA}. Fig. 5 (b) indicates that the ratio f^{RA*} is affected by the flexibility cost even for a specific system.

4) Effect of the Number of Cars in Demand: The demand for cars possibly increases in the near future with the development of services. Therefore, we increased the number of cars in demand to investigate its effect on the UBR strategy. The following scenario, besides the baseline scenario described in Section VII-A, is considered:

Scenario 1: The mean of cars in demand at each station in each period within the two peak traffic time frames was a random integer in the range [0, 30]. In each of the other periods, the mean of cars in demand at each station was a random integer in the range [0, 20].

Instances S5, M4, and L5 were selected for the analysis. For each instance and scenario, the system operations were simulated. The number of cars in demand, number of served users, car utilization ratio, operating profit, revenue, incentive expense, and the cost of occupying extra slots are presented in Table IV. The car utilization ratio is the duration in which cars are driven on the road divided by the total car duration over the planning horizon.

As can be seen from Table IV, the incentive expense rises as the demand increases, due to an insufficient supply of cars and the resulting larger need for relocations. However, the number of served users, car utilization ratio, revenue, and operating profit may not increase with the increment of the demand, as in instances M4 and L5. The reason may be that the higher demand aggregates the degree of asymmetry in demand, which leads to less match between supply and demand. The above results indicate that the UBR strategy may not improve the operating profit with the increment in demands when the car utilization ratio reaches a certain level. In such a case, the system can additionally apply operator-based relocations or increase the fleet size.

TABLE IV
EFFECT OF THE NUMBER OF CARS IN DEMAND

Instance	Scenario	No. of cars in demand	No. of served users	Car utilization ratio (%)	Operating profit (¥)	Revenue (¥)	Incentive expense (¥)	Cost of occupying extra slots (¥)
S5	Baseline scenario	41,025.8	24,895.8	58.35	405,203	611,142	18,752	3,845
	Scenario 1	42,197.2	26,081.0	60.82	405,614	637,051	37,954	2,368
M4	Baseline scenario	76,518.6	43,152.0	69.04	869,719	1,318,860	50,994	2,489
	Scenario 1	76,655.4	42,905.8	68.50	858,501	1,308,551	54,685	2,800
L5	Baseline scenario	147,980.0	77,430.0	61.16	1,568,350	2,310,240	22,651	26,166
	Scenario 1	149,489.0	75,802.6	59.81	1,524,630	2,259,200	24,327	32,481

VIII. Conclusion and Discussions

This research investigates an IBTP problem for applying a UBR strategy in non-reserved carsharing. A bi-level mixed-integer programming model is developed and reformulated into a single-level one, with a valid inequality introduced. An IPSO algorithm is developed. Extensive experiments are conducted to validate the methods.

Several managerial insights are provided for service operators. (1) The UBR strategy is applicable in real-world non-reserved systems because of its great potential to improve the operating profit and the number of served users. (2) The degree of trade-off between the incentive expense and the increased served users in the UBR strategy may vary with some critical parameters, which should be determined by simulation. (3) The benefits of the UBR strategy are affected by the flexibility cost significantly. Careful calibrations on the flexibility cost are required before actual applications. (4) The UBR strategy has restrictions in improving the operating profit with the increment in the potential demand. The system needs to additionally apply operator-based relocations or increase the fleet size if necessary.

This study can be followed in several aspects. First, the consideration of heterogeneous or uncertain flexibility costs in UBR strategies may be more valuable for applications. Second, (autonomous) electric cars are adopted increasingly in carsharing. Developing UBR strategies considering the charging of electric cars and measuring the environmental benefits (i.e., reduction in carbon emission caused by the application of electric cars) are both promising research angles. The integration of self-relocations by autonomous cars is also an expected angle of research. Third, user-based relocations alone might not improve the operating efficiency to its maximum. Hence, collaborative optimizations of user- and operator-based relocations are anticipated, especially in non-reserved systems.

Appendix: Proof of Theorem 1

We first introduce the *simplified incentive-based trip pricing*(SIBTP) problem as follows. Consider the set of stations $V_j = \left\{ i \in I | c^F d'_{ij} \leq P \right\}$, from any of which users may adjust the returning station to station $j \in I_D$, if station j provides the maximum incentive P. A subset V'_j of the set V_j consists of the stations from which users finally adjust their returning stations to station j under the assignment. The incentive, $p_j\left(V'_j\right)$, for returning a car to station j is at least $\max_{i \in V'_j} c^F d'_{ij}$ to guarantee nonnegative perceived utilities for users.

The utility of the returning adjustments regarding the subset V'_j is defined as

$$F_j\left(V'_j\right) = f_1 \sum_{i \in V'_j} \left(D_i^0 - D'_i\left(V'_j\right)\right) - f_2 p_j\left(V'_j\right) a'_j\left(V'_j\right), \tag{A1}$$

where $D'_i(V'_j)$ is the counterpart of the deviation D_i^0 under the adjustments regarding the set V'_j, and $a'_j\left(V'_j\right)$ represents the number of returned cars at station j under the adjustments regarding the set V'_j. The utility of the returning adjustments for the whole system is defined as

$$\psi = f_1 \sum_{j \in I_D} \sum_{i \in V'_j} \left(D_i^0 - D'_i(V'_j)\right) - f_2 \sum_{j \in I_D} p_j\left(V'_j\right) a'_j\left(V'_j\right). \tag{A2}$$

The goal of the SIBTP problem is to maximize the utility ψ by assigning users to return cars to car-deficient stations, while $V'_j \cap V'_k = \emptyset$ holds for any two selected subsets V'_j and V'_k $(j \neq k)$.

Lemma 1: The SIBTP problem is NP-hard.

Proof: The NP-hardness of the SIBTP problem can be proved by a reduction from the weighted set packing problem (which is NP-hard) considered in [48].

Given a weighted set packing problem, the construction of the SIBTP problem is as follows: a) The set of stations I corresponds to the finite set. b) The set of stations V_j for station $j \in I_D$ corresponds to a given subset of the finite set. c) The utility $F_j\left(V_j\right)$ for any station $j \in I_D$ corresponds to the weighting of the corresponding subset. The utility $F_j\left(V'_j\right)$ equals to zero constantly for any subset $V'_j \in V_j$.

The above links can be constructed in a linear time, and thus the SIBTP problem is NP-hard. □

Proof of Theorem 1: We denote $I^{IN} = \bigcup_{j \in I_D} V'_j$ as the set of stations, from any of which users adjust their returning stations to a car-deficient station under the assignment. The utility ψ can be expressed as

$$\psi = f_1 \sum_{i \in I^{IN}} \left(D_i^0 - D'_i\left(I^{IN}\right)\right) - f_2 \sum_{j \in I_D} p_j\left(V'_j\right) a'_j\left(V'_j\right). \tag{A3}$$

For any station $i \in I \setminus I^{IN}$, $D_i^0 = D_i'\left(I^{IN}\right)$, we have

$$\psi - f_1 \sum_{i \in I} D_i^0 = -f_1 \left(\sum_{i \in I \setminus I^{IN}} D_i^0 + \sum_{i \in I^{IN}} D_i'\left(I^{IN}\right) \right) - f_2 \sum_{j \in I_D} p_j\left(V_j'\right) a_j'\left(V_j'\right) = -F. \quad \text{(A4)}$$

The maximization of the utility ψ is equivalent to the objective function (5), because $f_1 \sum_{i \in I} D_i^0$ is a constant.

The above analysis indicates that the SIBTP problem relaxes Formula (4) as $m_{ij^*}\left(p_{j^*} - c^F d_{ij^*} - \Delta\right) \geq 0$ for $j^* \in I \setminus \{i\}$, compared to the IBTP problem. Therefore, the IBTP problem is NP-hard. □

Acknowledgment

The authors thank the Editor-in-Chief, the Associate Editor, and the anonymous reviewers for their valuable and constructive comments which have improved this article significantly.

References

[1] R. Zhang, H. Kan, Z. Wang, and Z. Liu, "Relocation-related problems in vehicle sharing systems: A literature review," *Comput. Ind. Eng.*, vol. 183, Sep. 2023, Art. no. 109504.

[2] H. Kan, Z. Liu, and R. Zhang, "Operating policies with temporally and spatially varied parameters in station-based carsharing services," *IEEE Trans. Eng. Manag.*, vol. 71, pp. 5101–5116, 2024.

[3] Y. Zhou, J. Wang, and H. Yang, "Resilience of transportation systems: Concepts and comprehensive review," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 12, pp. 4262–4276, Dec. 2019.

[4] T. Benarbia, K. W. Axhausen, and B. Farooq, "Modeling, relocation, and real-time inventory control of one-way electric cars sharing systems in a stochastic Petri nets framework," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 5, pp. 2846–2861, May 2021.

[5] D. Kypriadis, G. Pantziou, C. Konstantopoulos, and D. Gavalas, "Optimizing relocation cost in free-floating car-sharing systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 9, pp. 4017–4030, Sep. 2020.

[6] D. Gavalas, C. Konstantopoulos, and G. Pantziou, "Design and management of vehicle sharing systems: A survey of algorithmic approaches," in *Smart Cities and Homes: Key Enabling Technologies*, M. S. Obaidat and P. Nicopolitidis, Eds., Boston, MA, USA: Morgan Kaufmann, 2016, pp. 261–289.

[7] R. Iacobucci, R. Bruno, and C. Boldrini, "A multi-stage optimisation approach to design relocation strategies in one-way car-sharing systems with stackable cars," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 17048–17061, Oct. 2022.

[8] K. Huang, K. An, J. Rich, and W. Ma, "Vehicle relocation in one-way station-based electric carsharing systems: A comparative study of operator-based and user-based methods," *Transp. Res. E, Logistics Transp. Rev.*, vol. 142, Oct. 2020, Art. no. 102081.

[9] U. Eilertsen, O. M. Falck-Pedersen, J. V. Henriksen, K. Fagerholt, and G. Pantuso, "Joint relocation and pricing in electric car-sharing systems," *Eur. J. Oper. Res.*, vol. 315, no. 2, pp. 553–566, Jun. 2024.

[10] M. Schiffer, G. Hiermann, F. Rüdel, and G. Walther, "A polynomial-time algorithm for user-based relocation in free-floating car sharing systems," *Transp. Res. B, Methodol.*, vol. 143, pp. 65–85, Jan. 2021.

[11] J. Wang, F. Li, S. Yang, Y. Li, and Y. Wang, "A real-time bike trip planning policy with self-organizing bike redistribution," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 10646–10661, Aug. 2022.

[12] D. N. Das, R. Sewani, J. Wang, and M. K. Tiwari, "Synchronized truck and drone routing in package delivery logistics," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 9, pp. 5772–5782, Sep. 2021.

[13] Q. Liu, Q. Ye, X. Mei, and Q. Zhang, "Inter-cell scheduling based on a transportation vehicle sharing strategy," *Eur. J. Ind. Eng.*, vol. 17, no. 6, pp. 939–973, Dec. 2023.

[14] I. K. Singgih and B. I. Kim, "Multi-type electric vehicle relocation problem considering required battery-charging time," *Eur. J. Ind. Eng.*, vol. 14, no. 3, pp. 335–368, Aug. 2020.

[15] P. Ströhle, C. M. Flath, and J. Gärttner, "Leveraging customer flexibility for car-sharing fleet optimization," *Transp. Sci.*, vol. 53, no. 1, pp. 42–61, Feb. 2019.

[16] C. Wu, S. L. Vine, S. Philips, W. Tang, and J. Polak, "Free-floating carsharing users' willingness-to-pay/accept for logistics management mechanisms," *Travel Behaviour Soc.*, vol. 21, pp. 154–166, Oct. 2020.

[17] R. Curtale, F. Liao, and P. van der Waerden, "Understanding travel preferences for user-based relocation strategies of one-way electric car-sharing services," *Transp. Res. C, Emerg. Technol.*, vol. 127, Jun. 2021, Art. no. 103135.

[18] A. Angelopoulos, D. Gavalas, C. Konstantopoulos, D. Kypriadis, and G. Pantziou, "Incentivized vehicle relocation in vehicle sharing systems," *Transp. Res. C, Emerg. Technol.*, vol. 97, pp. 175–193, Dec. 2018.

[19] M. P. Fanti, A. M. Mangini, M. Roccotelli, and B. Silvestri, "Innovative approaches for electric vehicles relocation in sharing systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 1, pp. 21–36, Jan. 2022.

[20] M. Clemente, M. P. Fanti, G. Iacobellis, M. Nolich, and W. Ukovich, "A decision support system for user-based vehicle relocation in car sharing systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 8, pp. 1283–1296, Aug. 2018.

[21] S. Zhang, H. Sun, Y. Lv, and J. Wu, "Day-to-day dynamics of traveler learning behavior and the incentivization scheme of the operator for one-way carsharing services," *Comput. Ind. Eng.*, vol. 155, May 2021, Art. no. 107170.

[22] P. Stokkink and N. Geroliminis, "Predictive user-based relocation through incentives in one-way car-sharing systems," *Transp. Res. B, Methodol.*, vol. 149, pp. 230–249, Jul. 2021.

[23] J. Wang and Y. Wang, "A two-stage incentive mechanism for rebalancing free-floating bike sharing systems: Considering user preference," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 82, pp. 54–69, Oct. 2021.

[24] L. Wang, W. Ma, M. Wang, and X. Qu, "Demand control model with combinatorial incentives and surcharges for one-way carsharing operation," *Transp. Res. C, Emerg. Technol.*, vol. 125, Apr. 2021, Art. no. 102999.

[25] C. Müller, J. Gönsch, M. Soppert, and C. Steinhardt, "Customer-centric dynamic pricing for free-floating vehicle sharing systems," *Transp. Sci.*, vol. 57, no. 6, pp. 1406–1432, 2023.

[26] S. Banerjee, D. Freund, and T. Lykouris, "Pricing and optimization in shared vehicle systems: An approximation framework," *Oper. Res.*, vol. 70, no. 3, pp. 1783–1805, May 2022.

[27] G. Guo and Y. Hou, "Rebalancing of one-way car-sharing systems considering elastic demand and waiting time," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 12, pp. 23295–23310, Dec. 2022.

[28] S. Wang, Y. Yang, Y. Chen, and X. Zhao, "Trip pricing scheme for electric vehicle sharing network with demand prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 20243–20254, Nov. 2022.

[29] J. Pfrommer, J. Warrington, G. Schildbach, and M. Morari, "Dynamic vehicle redistribution and online price incentives in shared mobility systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 4, pp. 1567–1578, Aug. 2014.

[30] A. B. Brendel, S. Lichtenberg, B. Brauer, I. Nastjuk, and L. M. Kolbe, "Improving electric vehicle utilization in carsharing: A framework and simulation of an e-carsharing vehicle utilization management system," *Transp. Res. D, Transp. Environ.*, vol. 64, pp. 230–245, Oct. 2018.

[31] Y. Duan and J. Wu, "Spatial-temporal inventory rebalancing for bike sharing systems with worker recruitment," *IEEE Trans. Mobile Comput.*, vol. 21, no. 3, pp. 1081–1095, Mar. 2022.

[32] S. Zhang, H. Sun, X. Wang, Y. Lv, and J. Wu, "Optimization of personalized price discounting scheme for one-way station-based carsharing systems," *Eur. J. Oper. Res.*, vol. 303, no. 1, pp. 220–238, Nov. 2022.

[33] B. Shi, Y. Deng, and H. Yuan, "An auction based task dispatching and pricing mechanism in bike-sharing," *Knowl.-Based Syst.*, vol. 235, Jan. 2022, Art. no. 107631.

[34] B. Boyacı and K. G. Zografos, "Investigating the effect of temporal and spatial flexibility on the performance of one-way electric carsharing systems," *Transp. Res. B, Methodol.*, vol. 129, pp. 244–272, Nov. 2019.

[35] A. Di Febbraro, N. Sacco, and M. Saeednia, "One-way car-sharing profit maximization by means of user-based vehicle relocation," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 628–641, Feb. 2019.

[36] Y. Liu, J. Xie, and N. Chen, "Stochastic one-way carsharing systems with dynamic relocation incentives through preference learning," *Transp. Res. E, Logistics Transp. Rev.*, vol. 166, Oct. 2022, Art. no. 102884.

[37] Y. Cheng, J. Wang, and Y. Wang, "A user-based bike rebalancing strategy for free-floating bike sharing systems: A bidding model," *Transp. Res. E, Logistics Transp. Rev.*, vol. 154, Oct. 2021, Art. no. 102438.

[38] S. Datner, T. Raviv, M. Tzur, and D. Chemla, "Setting inventory levels in a bike sharing network," *Transp. Sci.*, vol. 53, no. 1, pp. 62–76, Feb. 2019.

[39] Z. Wu, J. Wu, Y. Chen, K. Liu, and L. Feng, "Network rebalance and operational efficiency of sharing transportation system: Multi-objective optimization and model predictive control approaches," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 17119–17129, Oct. 2022.

[40] Z. Haidera, A. Nikolaevb, J. E. Kangb, and C. Kwona, "Inventory rebalancing through pricing in public bike sharing systems," *Eur. J. Oper. Res.*, vol. 270, pp. 103–117, Oct. 2018.

[41] R. Nair and E. Miller-Hooks, "Equilibrium network design of shared-vehicle systems," *Eur. J. Oper. Res.*, vol. 235, no. 1, pp. 47–61, May 2014.

[42] Y. Deng and M.-A. Cardin, "Integrating operational decisions into the planning of one-way vehicle-sharing systems under uncertainty," *Transp. Res. C, Emerg. Technol.*, vol. 86, pp. 407–424, Jan. 2018.

[43] A. Ratnaweera, S. K. Halgamuge, and H. C. Watson, "Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients," *IEEE Trans. Evol. Comput.*, vol. 8, no. 3, pp. 240–255, Jun. 2004.

[44] K. Huanga, G. H. de Almeida Correiab, and K. An, "Solving the station-based one-way carsharing network planning problem with relocations and non-linear demand," *Transp. Res. C, Emerg. Technol.*, vol. 90, pp. 1–17, May 2018.

[45] M. Xu, Q. Meng, and Z. Liu, "Electric vehicle fleet size and trip pricing for one-way carsharing services considering vehicle relocation and personnel assignment," *Transp. Res. B, Methodol.*, vol. 111, pp. 60–82, May 2018.

[46] G. Molnar and G. H. D. A. Correia, "Long-term vehicle reservations in one-way free-floating carsharing systems: A variable quality of service model," *Transp. Res. C, Emerg. Technol.*, vol. 98, pp. 298–322, Jan. 2019.

[47] D. Wang, I. Moon, and R. Zhang, "Multi-trip multi-trailer drop-and-pull container drayage problem," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 19088–19104, Oct. 2022.

[48] G. Gao, L. Ning, H.-F. Ting, Y. Xu, Y. Zhang, and Y. Zou, "Approximation algorithms for the partial assignment problem," *Theor. Comput. Sci.*, vol. 838, pp. 231–237, Oct. 2020.

Ruiyou Zhang (Member, IEEE) received the B.Sc. degree in automation and the Ph.D. degree in systems engineering from Northeastern University, Shenyang, China, in 2002 and 2007, respectively.

From 2007 to 2008, he was a Post-Doctoral Fellow with the Department of Industrial Engineering, Pusan National University, Busan, South Korea. From 2010 to 2011, he was a Visiting Scholar with the School of Industrial and Systems Engineering, Georgia Institute of Technology, Georgia, USA. He is currently a long-termed Professor with the College of Information Science and Engineering, Northeastern University. He has published more than 60 refereed technical articles on international/domestic journals. His major research interests include modeling and optimization, intelligent manufacturing, logistics and transportation, and service science and engineering.

Haiyu Kan received the B.S. degree in automation from the North China University of Science and Technology, Tangshan, China, in 2017, and the Ph.D. degree in systems engineering from Northeastern University, Shenyang, China, in 2024. Her research interests include modeling and optimization, and logistics and transportation.

Ilkyeong Moon received the B.S. and M.S. degrees in industrial engineering from Seoul National University, South Korea, and the Ph.D. degree in operations research from Columbia University. He is currently a Professor of industrial engineering with Seoul National University. He published over 180 articles in international journals. His research interests include supply chain management, logistics, and inventory management. He is a fellow of Asia Pacific Industrial Engineering and Management Society and a Board Member of an International Foundation for Production Research and WG 5.7 of IFIP. He was the former Editor-in-Chief of *Journal of the Korean Institute of Industrial Engineers* which is a flagship journal of Korean Institute of Industrial Engineers (KIIE). He was the President of KIIE in which he had served from 2019 to 2020. He is the Editor-in-Chief of *European Journal of Industrial Engineering*.