https://doi.org/10.1016/j.tre.2020.102031

A Static Bike Repositioning Model in a Hub-and-Spoke Network Framework

Abstract

This paper addresses a static bike repositioning problem by embedding a short-term demand forecasting process, the Random Forest (RF) model, to account for the demand dynamics in the daytime. To tackle the heterogeneous repositioning fleets, a novel repositioning operation strategy constructed on the hub-and-spoke network framework is proposed. The repositioning optimization model is formulated using mixed-integer programming. An artificial bee colony algorithm, integrated with a commercial solver, is applied to address computational complexity. Experimental results show that the RF can achieve a high forecasting accuracy, and the proposed repositioning strategy can efficiently decrease the users' dissatisfaction.

Keywords: bike repositioning, demand forecasting, random forests, hub-and-spoke network framework, hub-first-route-second.

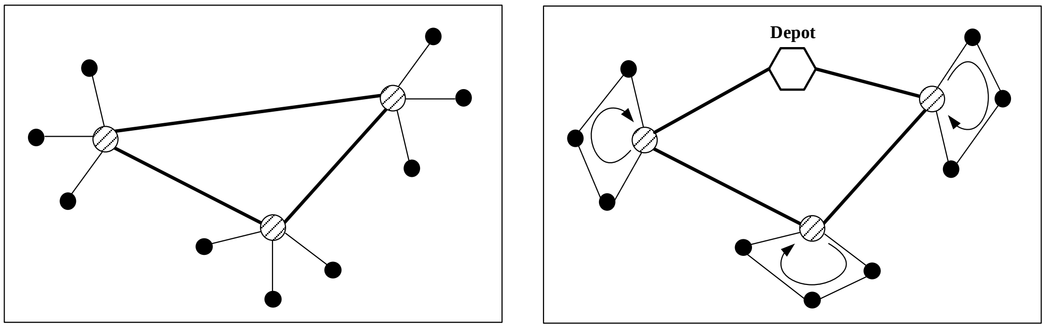
1. Introduction

The bike-sharing is becoming increasingly popular worldwide as a convenient, efficient, and green travel mode. It is designed to complement the existing multimodal transit system, and encourage the use of public transportation by addressing the first-/last-mile problem. In practice, one of the major issues faced by the bike-sharing system is the imbalance between demand and supply (i.e., the tidal effect), especially during rush hours (Fishman, 2016; Bai et al., 2017; Ji et al., 2017; Liu et al., 2018; Ji et al., 2020). For instance, during morning peak hours, users generally have difficulties in renting bikes from docking stations located in the residential areas as they are usually deficient in bikes. Whereas, stations located in the central business district area may have surplus bikes. The bike repositioning problem (BRP) is defined as the operational issue of rebalancing bike inventory at stations to meet potential user demand (Raviv et al., 2013; Dell'Amico et al., 2014; Forma et al., 2015; Szeto et al., 2016; Si et al., 2018).

The bike repositioning operation, from the perspective of a bike operator, can be categorized into two types: user-based and vehicle-based. The user-based repositioning is realized by a reward system that encourages users to return bikes to underused stations (Singla et al., 2015; Ghosh et al., 2017; Haider et al., 2018). In the vehicle-based case, a fleet of trucks is deployed by the operator to rebalance the bike inventory among stations. There is a large body of studies on the vehicle-based BRP, which has been widely implemented in the real-world operation as well (Chemla et al., 2013; Raviv et al., 2013; Dell'Amico et al., 2014; Erdoğan et al., 2014; Szeto et al., 2016; Ho & Szeto, 2017; among many others). Considering the time-dependent demand fluctuation, the vehicle-based BRP can be further classified into static and dynamic cases. In the static case, the repositioning operation is commonly performed during night-time or before the morning peak when the customer demand is low, and the demand fluctuation is negligible. In the dynamic case, the repositioning system operates in the daytime while users are continuously renting/returning bikes.

The state of docking stations (i.e., deficit or surplus) depends on the existing number of bikes at each station as well as the potential demand in the near future. In the static case, users' activities of rental and return are not considered during the repositioning process, while the quantity of bikes at each station is known in advance and unchanged. The dynamic BRP considers real-time station demand variations which is more intricate to manage because of user's behaviors and activities (Contardo et al., 2012; Forma et al., 2015; Ghosh et al., 2017; Ghosh & Varakantham, 2017; Shui & Szeto, 2018; Legros, 2019). Data mining techniques can identify and estimate the underlying demand patterns from historical data with respect to (w.r.t) demand variations (Albiński et al., 2018; Mellou, 2019; Liu, Y. et al., 2019; Liu, Z. et al., 2019). In this regard, the short-term demand forecasting, including both rental and return demands at each station for a given period in the daytime, can help the operator to optimize its myopic repositioning decisions. The forecasted result of a station represents the final inventory state of this station as the forecasting procedure implicitly considers demand fluctuation during any given period. A static BRP with forecasted demand can also cope with the fluctuating demand by determining each station's final inventory state during each period.

When dealing with real-world operational issues, the substantial size of bike-sharing system imposes a significant burden on the repositioning efficiency to serve all stations (Szeto et al., 2016). Some studies select a subset of stations w.r.t station characteristics (e.g., location, demand features) and solve the BRP efficiently by narrowing the solution search space (Ho & Szeto, 2014; Regue & Recker, 2014).



(a) **(b)**

Figure 1. (a) The hub-and-spoke network framework; (b) Illustration of the BRP in a hub-and-spoke network framework.

In this paper, a hub-and-spoke framework is applied to select crucial stations and to construct a hub network connecting all other nodes (i.e., non-hub nodes, also known as spokes). The hub-and-spoke framework is employed widely in transportation, logistics, telecommunications, and computer networks as it efficiently routes flows between multiple origins and destinations (Gelareh & Nickel, 2011; Lin & Yang, 2011; Huang et al., 2018). As shown in Fig. 1, at the operational level, selected bike stations play roles of hubs that distribute and collect bikes. The determination of the optimal number and location of hubs is usually described as a hub location problem (HLP).

A demand forecasting process using a machine learning technique is embedded in the BRP to capture demand fluctuations. The objective function aims to minimize the unmet demand and the total routing cost. Given a short-term planning horizon, the routing decisions of vehicles that perform the redistribution and loading/unloading quantities at each station are optimized based on the forecasted user demand. The proposed methodology is divided into two steps: demand forecasting and vehicle routing. The random forests (RF) model is used to estimate both rental and return demands at each station, which is an ensemble learning method that combines a multitude of decision trees (conducting classification, regression, or other tasks). The implication of “random” is twofold: i) random sample with replacement of the training set; and ii) random selection of features. Compared with decision trees, the RF overcomes the weakness of overfitting by averaging multiple deep decision trees with different subsets of the same training set to reduce the variance and give an unbiased estimation (Friedman et al., 2001; Lahouar & Slama, 2017).

1.1 Literature review

1.1.1 Bike repositioning problem

The goal of BRP is to ensure the cost-efficient allocation of bikes to stations while attaining an optimal system service level considering the spatial and temporal distributions of bike demands (Chemla et al., 2013; Raviv et al., 2013; Ho & Szeto, 2014; Szeto et al., 2016). Most of the studies view the BRP as an extension of the classical traveling salesman problem (TSP) and vehicle routing problem (VRP). However, the BRP is more complicated than those problems as it should make routing and bike allocation decisions simultaneously. The BRP can also be classified as a variation of the VRP with pickup and delivery (Forma et al., 2015).

The objective of BRP is designed according to the bike-sharing operator’s concerns, which can be roughly divided into operator- and user-oriented. In the first case, the operating cost incurred during the repositioning process is minimized, i.e., the total routing cost (including time-, labor- and cost-related attributes) (Raviv et al., 2013; Erdoğan et al., 2014; Kadri et al., 2016). The user-oriented objectives concern the issues related to user satisfaction or the system’s level of service, e.g., the total unmet demand (Contardo et al., 2012), the total deviation between final and expected inventory (Rainer-Harbach et al., 2013), and the total penalty cost incurred at each station (Raviv et al., 2013; Ho & Szeto, 2014).

Another concern of BRP is to rebalance the bike inventory caused by asymmetric demand and to keep each station at a desired inventory level. There have been some works in developing the optimal inventory level for each station using inventory models. Raviv & Kolka (2013) describe the BRP as a closed-loop inventory problem and introduce a convex user dissatisfaction function. Rainer-Harbach et al. (2015) consider a combined problem of inventory balancing and vehicle routing, and propose several construction heuristics to obtain high-quality solutions efficiently. Schuijbroek et al. (2017) also consider the combined problem and present a constraint programming formulation to obtain the exact solution for small-scale problems and the benchmark for heuristics.

Table 1. A summary of the existing solution methods and for BRP.

Solution method		Publication
Exact algorithm	Branch-and-bound	Kadri et al. (2016)
	Branch-and-cut	Erdoğan et al. (2014, 2015); Dell'Amico et al. (2013, 2016); Bulhões et al. (2018)
	Benders decomposition	Contardo et al. (2012); Erdoğan et al. (2014)
Heuristics or metaheuristics	Cluster-first-route-second	Forma et al. (2015); Schuijbroek et al. (2017)
	Iterated tabu search	Ho & Szeto (2014)
	Chemical reaction optimization	Szeto et al. (2016)
	Generic algorithm	Li et al. (2016)
	Ant colony optimization	Di Gaspero et al. (2013)
	ABC algorithm	Shui & Szeto (2018); Szeto & Shui (2018)
	Construction heuristics	Rainer-Harbach et al. (2015)
Approximation method	Large neighborhood search	Di Gaspero et al. (2016)
Hybrid algorithm	Branch-and-cut with tabu search	Chemla et al. (2013)
	3-step math heuristic	Forma et al. (2015)

As aforementioned, the BRP, as an extension of TSP and VRP, has been proved to be NP-hard because it contains the NP-hard problem as a special case (Chemla et al., 2013). Table 1 presents a summary of the prevalent solution approaches adopted in the existing literature. Exact methods, such as branch-and-bound algorithm (Kadri et al., 2016), branch-and-cut algorithm (Dell'Amico et al., 2014; Erdoğan et al., 2014) and Benders decomposition (Contardo et al., 2012) are only designed to solve small-scale repositioning experiments and are, therefore, intractable in realistic or large-scale instances. Thus, heuristics or metaheuristics methods are widely adopted to reduce the problem scale and to address large-scale or realistic repositioning operations. Ho & Szeto (2017) also point out that, in practice, not all stations need to be visited. The reasons are i) the rental and return demands at a station are roughly equal in a given period, termed as a “balanced” station; ii) the marginal repositioning costs incurred by some stations is larger than the penalty cost; and iii) the total supply is insufficient from all pickup stations to delivery stations. Hence, how to select crucial stations to serve is of great importance to narrow the solution space and expedite the solution algorithm.

During the daytime, especially in rush hours, the station demand varies dynamically. Monitoring real-time system status or forecasting stations' state is of high value to operators as they are required to rebalance the system in a dynamic manner (Lathia et al., 2012). Contardo et al. (2012) first formulate the dynamic BRP on a space-time network by discretizing the time horizon into periods to capture time-dependent demand. Shu et al. (2013) develop a bike network flow model on a time-expanded graph in both deterministic and stochastic systems. The results show that the proposed deterministic model could approximate the actual system performance closely. Ghosh et al. (2017) describe the dynamic BRP as a sequential decision-making model in the Markov Decision Process (MDP). Decomposition and aggregation techniques are employed to obtain a near-optimal solution. Legros (2019) also uses the MDP to develop a decision-support tool to decide which station should be visited first and the find optimal inventory at each station. Shui & Szeto (2017) adopt a rolling horizon approach to address time-varying demand and decomposed the whole service time horizon into smaller static sub-problems that can be solved efficiently.

Stochastic optimization has also been applied to solve the BRP under uncertainty. For example, the robust optimization techniques are widely introduced to handle uncertain demand, where the dynamics of user demand is modeled by different demand scenarios extracted from historical trip data. Lu (2016) formulates a time-space network to capture the time-dependent bike flows. The uncertain bike demand is prescribed by uncertainty sets, based on which the worst-case scenario of the bike system is analyzed. Ghosh et al. (2016) propose an online and robust repositioning model to minimize the unmet demand in the bike system. A scenario generation approach is developed based on an iterative game. The operator makes routing and repositioning decisions according to the worse case lost demand in each iteration. Jin et al. (2019) develop a two-stage stochastic programming model to maximize user demand. The user's demand scenario is determined by time periods, travel intensity, and distribution. Ghosh et al. (2019) propose a dynamic bike repositioning approach aiming to maximize the probability of satisfying the user's demand. The demand uncertainty is obtained from the historical demand data.

1.1.2 Demand analysis and forecasting

The uncertain and fluctuating bike demand during the daytime affects the efficiency of bike repositioning operation significantly. It is, therefore, necessary to conduct a demand analysis to capture the spatial and temporal characteristics of the pattern of bike usage. The near-future demands can be then forecasted based on the understanding of influencing factors in bike user's decision making process. Many efforts are devoted to applying machine learning techniques in the demand forecasting problem of the bike-sharing system, including regression (Regue & Recker, 2014; Hulot et al., 2018), classification (Ruffieux et al., 2018), clustering (Vogel et al., 2011; Guido et al., 2019), time series analysis (Kaltenbrunner et al., 2010), and neural networks (Xu et al., 2018). Hulot et al. (2018) conduct a comprehensive comparison between four demand forecasting methods, i.e., linear regression, neural network, gradient boosted tree, and RF. The author uses temporal and weather features to predict hourly demand for rentals and returns. The prediction performance results show that the Singular Value Decomposition method embedded in the gradient boosted tree predictor outperformed other models. A satisfactory score was achieved by RF as well. The RF illustrates better accuracy for short-term predictions in Ruffieux et al. (2017) and Ruffieux et al. (2018).

One of the challenges in demand forecasting problem is the lost demand. It occurs when there are no bikes available at a station or all docks are occupied. However, it cannot be observed from the trip data collected in the bike-sharing system. Few researchers have made attempts to analyze and forecast the true (or latent) demand of the bike-sharing system. O'Mahony & Shmoys (2015) propose a pure data-driven approach by using the average number of trips for each time window as the lower bound of the true demand. A censoring process is applied to omit outage data (i.e., the zero demands at the same station and at the same time almost every day) to ensure the accuracy of demand prediction. It is a fact that valid observations of underlying demand are insufficient and cannot be extracted from the trip data of the bike-sharing system. In this regard, several researchers intend to estimate true demand theoretically. The random arrival of bike users is usually considered a Poisson process, with arrival at each time step at each station following a Poisson distribution. Mellou & Jaillet (2019) estimate the lost demand while considering both average station behavior and daily demand trends. Goh et al. (2019) estimate the primary (first choice) demand in a rank-based demand model while considering the user's choice substitution. Users are allowed to switch to nearby stations when their first-choice is not available. Datner et al. (2017) also introduce a user behavior model, which minimized users' journey dissatisfaction w.r.t the existing state of the system. Schuijbroek et al. (2017) define the net demand process as a non-stationary stochastic process and determined the number of bikes to meet the service level requirement. They also argue that the "lost sales" bias (i.e., the unmet demand is not recorded) cannot be omitted because it is stochastically complex to capture. Negahban (2019) first propose a simulation-based approach to estimate the true demand in the bike-sharing system. The proposed novel methodology combined simulation, bootstrapping, and subset selection to reveal the underlying demand hidden in the usage data.

1.2 Objectives and contributions

In the existing literature, the static and dynamic bike repositioning are differentiated clearly for overnight and daytime operations, respectively (Regue & Recker, 2014; Forma et al., 2015; Schuijbroek et al., 2017; Shui & Szeto, 2018). However, in practice, the distinction between static and dynamic repositioning operations is blurred. In the daytime, the operator can hardly capture the real-time demand fluctuation and frequently change the vehicle's repositioning route. The repositioning operation always lags behind the user demand. Moreover, to improve the efficiency of the operator's rebalancing program, different types of vehicles (w.r.t the vehicle capacity) are widely deployed in practice, e.g., the Citi Bike in New York City, which employs fleets of 3-bike trailers and larger capacity trucks (Urbica, 2016). Though very few, some efforts have been devoted to deploying heterogeneous vehicles in the bike repositioning operation (Dondo et al., 2007; Contardo et al., 2012; Raviv et al., 2013). However, it is still an open question when it comes to constructing a coordinated and efficient network framework to tackle heterogeneous fleets. This paper attempts to remedy this gap by applying a hub-and-spoke network framework as the backbone of the bike repositioning operation, where the inter-hub reposition is served by a fleet of vehicles with a larger capacity to increase operational efficiency. The hub-and-spoke network is applied to address the strategic design problem of the bike-sharing system, such as the number of locations of stations, the inventory level (Lin et al., 2013). Despite this, no previous work has applied the hub-and-spoke network framework in the bike repositioning operation.

This paper closes the gap in the literature and makes the following contributions. First, a novel bike repositioning operation strategy is proposed based on the hub-and-spoke network framework. Compared with existing hub-first-route-second (HFRS) approaches, the operator could benefit from the economies of scale by consolidating user demand from and to spoke nodes, which improves the overall operational efficiency. Second, to address the problem of demand fluctuation during the bike repositioning operation in the daytime, a demand forecasting system is constructed based on the RF model. The essential features are identified and selected using a practical feature engineering process. Hence, unlike the robust BRP, which considers the demand uncertainty only based on the historical trip scenarios, the proposed demand forecasting approach considers essential factors that affect bike usage and analyze the dynamics of user demand comprehensively.

The rest of the paper is organized as follows. The next section presents some of the related works. Section 3 introduces the principle and procedures of the RF model. Section 4 presents and models the HFRS problem. Section 5 describes the solution algorithm for the proposed HLP. Section 6 includes a set of experiments based on randomly generated instances. A case study of a real-world bike-sharing system in Nanjing, China, is conducted in Section 7. Finally, we conclude the paper with some remarks and future work perspectives.

2. Demand forecasting system

One of the key characteristics that differentiate the bike-sharing system with other transit modes is that its usage is highly affected by weather conditions (e.g., temperature, humidity, rain, wind) (Lathia et al., 2012). Additionally, other factors like the type of day (weekday or weekends) and spatial dependency (residential, commercial, or uphill stations) have an evident impact on the bike usage as well (Ruffieux et al., 2017). Hence, the demand forecasting model for the bike-sharing system is required to accommodate high feature dimensionality. The RF model, which is a non-parametric and ensemble machine learning approach, is proved to have a high generalization ability to identify the importance of selected features. The basic principle of RF is to extract a set of samples from the training set and to fit them into decision trees, which is known as the bootstrap aggregating, or bagging (Breiman, 2001; Jiang et al., 2009; Lahouar & Slama, 2017). The following subsections present the basic component and forecasting procedure of the RF, followed by the measuring criterion of forecasting performance.

2.1 Decision tree

The decision tree is represented by a statistical model that classifies the samples based on the outputs in terms of a set of input features (Breiman et al., 1984). Based on the ability to handle discrete and continuous variables, decision trees can be divided into two categories: classification and regression. In the bike-sharing system, the quantities of rental and return demands are usually considered as continuous variables (Ruffieux et al., 2017; Ruffieux et al., 2018). Hence, a regression model is applied to build the decision trees.

Generally, a tree is a set of nodes and branches which are organized hierarchically without any loops. Each node stores a test function for the incoming data. If each node has two outgoing branches, the tree is considered a binary tree. Let X denote an input vector containing m characteristic variables (or features), Y an output scalar (i.e., the objective value), and S_k a training set containing k observations denoted by $\left(X_k,\ Y_k\right)$. During the training process, the input vector branches at each node so that the variables of split functions are optimized to fit with the training set S_k. A recursive split is applied to search optimal sub-partitions from X. Specifically, each step in the training of a decision tree intends to branch at the best node into two different sub-partitions. Each branch should minimize the variance of the child node to select the best split (Lahouar & Slama, 2017). The variance of node p is defined as:

$$Var(p) = \sum_{i:X_i \in p} (Y_i - \bar{Y}_p)^2 , \tag{1}$$

where $\bar{Y}_p$ is the mean of the objective value Y_i in node p.

The child nodes continue to repeat the above branching process until a termination criterion is satisfied. In general, the branching process terminates when the sample number of a node is greater than or equal to a predefined value, $R_{\min}$. A prediction function $f(X, S_k)$ is constructed from the training data once the training process stops. For any new input vector X', an estimation $\hat{Y}$ can be obtained by the prediction function $\hat{f}(X, S_k)$, given as follows,

$$\hat{Y} = \hat{f}(X', S_k). \tag{2}$$

2.2 Demand forecasting using the RF model

As aforementioned, the RF model is an ensemble method that contains several decision trees generated through a bootstrap sampling process (Breiman, 2001; Lahouar & Slama, 2017). The bootstrap sampling process randomly extracts L sample subsets from the training set. The selected L sample subsets are trained in L decision trees by $f(X, S_k)$, $k = 1, 2, ..., L$. Each decision tree outputs a prediction value, while the final estimation $\hat{Y}$ of the RF is obtained by averaging the predicted values of all the decision trees. Fig. 2 shows the structure of the RF model.

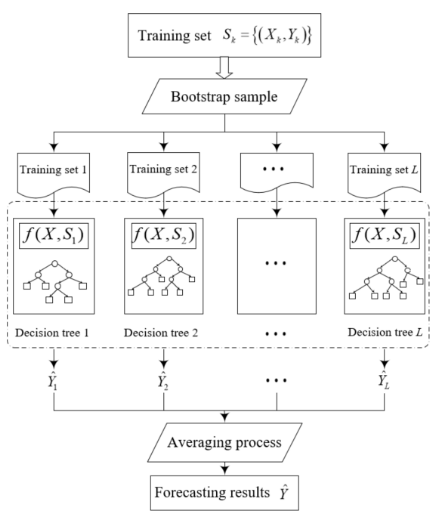


Figure 2. The structure of the RF model.

Note that the result of each decision tree is independent and identically distributed. The average output of the RF is obtained by the expected value of each tree, given by,

$$E(\frac{1}{L}\sum_{k=1}^{L} T_k) = E(T_{k'}),\ k' = 1, ..., L. \tag{3}$$

The generalization performance of the RF model can be improved by reducing the variance. Assume that the variance of each decision tree is $D(T_k) = \sigma^2$, and the correlation coefficient of each arbitrary decision tree is ρ, and $\rho > 0$. The variance of RF is:

$$D(\frac{1}{L}\sum_{k=1}^{L} T_k) = \rho\sigma^2 + \frac{1-\rho}{L}\sigma^2 . \qquad (4)$$

According to Eq. (4), it is clear that increasing the number of decision trees would result in a smaller model variance, which guarantees the quality of prediction. Additionally, while generating the decision tree, each node needs to select m ($m \le n$) characteristic variables from n input variables before branching. The correlation coefficient would decrease as m decreases, which would, in turn, result in a small model variance. However, it increases the deviation of RF as well. According to Breiman (2001), m is taken as $\lceil n/3 \rceil$, where $\lceil \cdot \rceil$ denotes the ceiling function.

In sum, the steps of the RF method are presented as follows:

Step 1: Set the number of decision trees, L;

Step 2: Extract a training subset S_k (where $k = 1,...,L$) from the full training set S by using the bootstrap method, where the size of S_k is N;

Step 3: Repeat the following steps in S_k until the number of samples for the node does not exceed $R_{\min}$. Thus, obtaining a decision tree T_k:

(1) m characteristic variables are selected randomly among n characteristic variables;

(2) $\theta_k(j,s)$ is obtained by selecting the best variable j and the segmentation point s from m feature variables;

(3) According to $\theta_k(j,s)$, the node is divided into two child nodes;

Step 4: Consider the output of all the decision tree sets $\{T_k\}_1^L$ to form random forests. The regression output of the model is as follows:

$$\hat{Y} = \frac{1}{L}\sum_{k=1}^{L} T(x, S_k, \theta_k) . \qquad (5)$$

2.3 Performance measure

Let $y(i)$ and $\hat{y}(i)$ denote the actual and the predicted values, respectively. Three criteria widely used in existing studies are adopted to evaluate the performance of the prediction model (see Table 2). The first two criteria use raw error values to evaluate prediction results. It may result in bias due to the different magnitude of the predicted values. To address this problem, the mean absolute percentage error is also adopted to evaluate the prediction result, which is the ratio of error over the real value. Note that when the actual value is close to zero, the mean absolute percentage error tends to approach infinity (Lahouar & Slama, 2017). Hence, to ensure a comprehensive evaluation, these three criteria are used simultaneously.

Table 2. The calculation functions of evaluation criteria.

Criterion	Calculation function
Mean absolute error	$M_{AE} = \frac{1}{n}\sum_{t=1}^{n}\lvert y(i) - \hat{y}(i)\rvert$
Root mean square error	$R_{MSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y(i) - \hat{y}(i))^2}$
Mean absolute percentage error	$M_{APE} = \frac{1}{n}\sum_{t=1}^{n}\frac{\lvert y(i) - \hat{y}(i)\rvert}{y(i)}$

3. The hub-first-route-second problem in BRP

The RF model can be used to predict bike demand at each station for any given period by embedding the model with the exogenous factors (the time of day, weather, temperature, etc.) that influence bike usage. For an operator, it is relatively straight forward to optimize the repositioning strategy by considering only the near-future demand to ignore the fluctuation of user demand for any given period. However, it is far less cost-efficient to serve all stations with limited labor force and resources.

This section introduces an HFRS repositioning strategy implemented on a hub-and-spoke network framework. The strategy is composed of two basic decisions: i) which stations should be selected as hubs to minimize the total repositioning cost, and ii) how to construct a route via hubs while minimizing the customer's dissatisfaction. The proposed approach addresses these two issues by selecting a set of stations to visit, sequencing them, and determining the loading/unloading quantities at each station. The iteration lies in the fluctuation between a pure HLP and a VRP. This study approximates the delivery costs w.r.t the number of bikes transferring between hub stations while assessing the possible hub locations. The subsequent routing problem is modeled as a single-vehicle VRP with pick-up and delivery. The following section will introduce an optimization model for the two-stage bike repositioning problem. The notations used are listed in Table 3.

Table 3. List of notations.

Notation	Description
Set	
A	set of links
V	set of nodes
V_0	set of stations
$\bar{V}$	set of hub stations
V_p	set of non-hub stations allocated to hub station p, $p \in \bar{V}$
Parameter	
c_{ij}	transportation cost between arc $(i, j) \in A$
D_i	demand dissatisfaction at station i
f_{ij}	number of bikes carried by the vehicle when it travels between arc $(i, j) \in A$
P	number of hubs
Q_1, Q_2	capacity of the vehicle serving hub stations and vehicles serving non-hub stations, and $Q_1 > Q_2$
q_i	predicted user demand at station i
s_i^0	initial inventory at station
s_i	final inventory at station i
v	user's walking speed
α	parameter of inter-hub travel cost
μ_1, μ_2	parameters associated with the penalty function
Variable	
Hubbing stage	
x_{ij}	binary variable for hub-spoke assignment, i.e., $x_{ij} = 1$ if node i is allocated to a hub located at station j, and $x_{ij} = 0$, otherwise
y_{ij}	binary variable for inter-hub routing, i.e., $y_{ij} = 1$ vehicles travels from hub i to hub j, and $y_{ij} = 0$, otherwise
Routing stage	
g_j	auxiliary variable applied for the sub-tour elimination constraint
n_i^L	quantity of bikes loading at station i

n_i^U	quantity of bikes unloading at station i
z_{ij}	binary variable for routing, i.e., $z_{ij} = 1$ vehicle travels directly from node i to node j, and $z_{ij} = 0$, otherwise

In this study, the bike-sharing system is represented by a complete direct graph $G = (V, A)$, where V and A are sets of nodes and arcs, respectively. The nodes consist of a set of stations denoted by $V_0 = \{1, 2, ..., N\}$ and a depot (denoted by $\{0\}$), where N is the total number of stations. The number of hubs is defined exogenously and is denoted by P. The target number of bikes at a station $i \in V_0$ is q_i, where a positive value represents the station has a surplus of bikes, whereas a negative value indicates that bike at the station is deficient. Furthermore, there is no additional cost for constructing hubs, and the docking stations are uncapacitated. The binary decision variable x_{ij} is used to define the hub location, where x_{ij} equals 1 if node i is allocated to a hub located at node j; and 0, otherwise. x_{jj} equals 1 if node j is the hub node. Let c_{ij} be the transportation cost between arc $(i, j) \in A$. The binary decision variable y_{ij} is used to indicate the inter-hub routing, where y_{ij} equals 1 if vehicles travel from hub i to hub j; and 0, otherwise. Besides, the number of bikes loading or unloading at each station is denoted by n_i^L and n_i^U, respectively.

In this study, two types of repositioning vehicles are considered. The hub stations are assumed to be served by a truck with large capacity and high operation costs. Whereas, a fleet of pickup trucks with low capacity serves the spokes. Hence, a parameter α ($\alpha > 1$) is introduced and multiplied with the inter-hub travel cost incurred by the truck.

3.1 Hubbing stage

HLP is a popular research area in location theory. It has been widely applied to solve various transportation problems, e.g., public transit network design, logistics and shipping, by constructing hubs to connect all other nodes (i.e., non-hub nodes or spokes). Compared to a fully connected network, the hub-and-spoke network framework can considerably decrease the number of transportation links and provide cost-efficient services. Assume that the spatial distribution of all stations and the bike demand obtained from the demand forecasting system are known. The determination of hub locations can be described as a discrete HLP, embedded with a hub routing problem. In the hub-and-spoke framework, the non-hub stations or spokes can be allocated to one hub station, indicating that the unmet users gathered in spokes need to walk to the hubs to fulfill their rental or return demands. Hence, the total cost of the HLP is composed of two components: the user's walking cost and vehicle routing cost. The mathematical formulation of the single allocation p-HLP of the bike repositioning problem is as follows:

[P1]

$$\min z_1 = \sum_{i \in V_0} \sum_{j \in V_0, i \neq j} \frac{|q_i| c_{ij}}{v} x_{ij} + \alpha \sum_{i \in V_0} \sum_{j \in V_0, i \neq j} c_{ij} y_{ij} \tag{6}$$

subject to

$$\sum_{j \in V_0} x_{ij} = 1, \ \forall i \in V_0, \tag{7}$$

$$\sum_{j \in V_0} x_{jj} = P, \ \forall j \in V_0, \tag{8}$$

$$x_{ij} - x_{jj} \leq 0, \ \forall i, j \in V_0, \tag{9}$$

$$\sum_{i \in V_0, i \neq j} y_{ij} = 1, \ \forall j \in V_0, \tag{10}$$

$$\sum_{i \in V_0, i \neq j} y_{ji} = 1, \ \forall j \in V_0, \tag{11}$$

$$x_{ij} \cdot y_{ij} = 0, \ \forall i, j \in V_0, \tag{12}$$

$$x_{ij}, y_{ij} \in \{0,1\}, \ \forall i, j \in V_0. \tag{13}$$

The objective function (6) minimizes the total transportation cost of the bike-sharing system. The first term is the connection cost or the user's walking cost of all transportation from non-hub node i to hub node k; if i is allocated to k. The second term is the vehicle's inter-hub routing cost. Constraint (7) ensures that non-hub node i is assigned to one hub node. Constraint (8) defines the number of hub nodes to be selected. Constraint (9) ensures that node i is assigned to a hub node j only if a hub is located at node j. Constraints (10)-(12) define the inter-hub routing problem. Constraint (13) defines the binary decision variables.

Several existing studies have proved that the *p*-HLP is NP-hard even if the locations of hubs are known and fixed (Alumur & Kara, 2008). Sohn & Park (2000) first proved that the single allocation problem with *p* fixed hubs is NP-hard when the number of hubs is larger than three. Hence, in this study, a heuristic algorithm, namely, the ABC algorithm, is applied to address the large-scale hub location problem.

3.2 Routing stage

Once the HLP is solved, the hub locations and the allocation of non-hub stations to hubs are determined. The routing problem is then divided into $P+1$ subproblems, taking the hub stations and the assignment of non-hub stations obtained from the hubbing stage as the inputs. In the case of inter-hub routing, the vehicle starts from the depot and visits hub stations. While for spoke routing, the vehicle starts from the hub station and visits the selected non-hub stations sequentially to load/ unload bikes, and finally returns to the hub station.

Let $\bar{V}$ denote the set of hubs obtained in the hubbing stage, and $\bar{V}_0 = \bar{V} \cup \{0\}$. Let V_p denote the set of non-hub stations allocated to hub p, $p \in \bar{V}$, and $\bar{V}_p = V_p \cup \{p\}$. Similar to the formulation proposed by Szeto et al. (2016), each hub station i, $i \in \bar{V}$ is characterized by its initial inventory s_i^0, final inventory s_i, and demand q_i. Note that the capacity constraint of the docking station is relaxed by jointly considering the number of lockers at each station and inventory bikes. The capacities of the redistributing vehicles for hub stations and non-hub stations are assumed as Q_1 and Q_2 (in terms of the number of bikes).

The decision variables are defined as binary, in which z_{ij} equals 1 if the vehicle travels directly from node i to j, and $i, j \in \bar{V}$. The number of bikes transported by the vehicle while traveling from node i to node j is denoted by f_{ij}. Moreover, the number of bikes loading or unloading at each station is denoted by n_i^L and n_i^U, respectively.

The objective function of the routing stage is to minimize the weighted sum of unmet customer demand and operational time on the vehicle route. The inter-hub routing problem can be formulated as follows. Note that for the non-hub routing problem, we only need to replace $\bar{V}$ with V_p, and $\bar{V}_0$ with $\bar{V}_p$.

[P2]

$$\min\ z_2 = \mu_1 \sum_{i\in\bar{V}} D_i + \mu_2 \sum_{i,j\in\bar{V}_0, i\neq j} c_{ij} z_{ij} \tag{14}$$

subject to

$$D_i \geq q_i - s_i,\ \forall i \in \bar{V}, \tag{15}$$

$$\sum_{j\in\bar{V}} z_{0j} = 1, \tag{16}$$

$$\sum_{j\in\bar{V}_0, j\neq i} z_{ij} \leq 1,\ \forall i \in \bar{V}, \tag{17}$$

$$\sum_{j\in\bar{V}_0, j\neq i} z_{ij} = \sum_{j\in\bar{V}_0, j\neq i} z_{ji},\ \forall i \in \bar{V}_0, \tag{18}$$

$$g_j \geq g_i + 1 - M\left(1 - z_{ij}\right),\ \forall i \in \bar{V}_0, j \in \bar{V}, i \neq j, \tag{19}$$

$$\sum_{j\in\bar{V}} f_{0j} = 0, \tag{20}$$

$$f_{ij} \leq Q_k \cdot z_{ij},\ \forall i \in \bar{V}_0, j \in \bar{V}, i \neq j, k = \{1,2\}, \tag{21}$$

$$s_i = s_i^0 - n_i^L + n_i^U,\ \forall i \in \bar{V}, \tag{22}$$

$$n_i^L - n_i^U = \sum_{j\in\bar{V}_0, i\neq j} f_{ij} - \sum_{j\in\bar{V}_0, i\neq j} f_{ji},\ \forall i \in \bar{V}, \tag{23}$$

$$n_i^U = \min\left\{\max\left\{q_i - s_i^0, 0\right\} \cdot \sum_{j\in\bar{V}_0, i\neq j} z_{ij}, \sum_{j\in\bar{V}_0, i\neq j} f_{ij}\right\},\ \forall i \in \bar{V}, \tag{24}$$

$$n_i^L \leq \min\left\{\max\left\{s_i^0 - q_i, 0\right\} \cdot \sum_{j\in\bar{V}_0, i\neq j} z_{ij}, Q - \sum_{j\in\bar{V}_0, i\neq j} f_{ij}\right\},\ \forall i \in \bar{V}, \tag{25}$$

$$z_{ij} = \left\{0,1\right\},\ \forall i \in \bar{V}_0, j \in \bar{V}, i \neq j, \tag{26}$$

$$z_{ii} = 0,\ \forall i \in \bar{V}_0, \tag{27}$$

$$n_i^L, n_i^U \geq 0,\ \text{integer},\ \forall i \in \bar{V}, \tag{28}$$

$$f_{ij} \geq 0,\ \forall i \in \bar{V}_0, j \in \bar{V}, i \neq j, \tag{29}$$

$$g_i \geq 0,\ \forall i \in \bar{V}, \tag{30}$$

where μ_1 and μ_2 are weighting values for each term.

Constraint (16) indicates that the vehicle leaves the depot only once. Constraint (17) states that each station is visited by the vehicle at most once. Constraint (19) is the vehicle flow conservation equation. Constraint (19) is the sub-tour elimination constraint. Constraint (20) ensures that the vehicle is empty when it leaves the depot. Constraint (21) is the capacity constraint that ensures the number of bikes on the vehicle is no greater than the vehicle capacity Q on each link. Constraint (22) depicts the bike loading/unloading conservation equation at each station, which defines the final quantity of bikes. Constraint (23) depicts the conservation of bikes on each vehicle: the number of bikes loading or unloading at a station is equal to the difference between the number of bikes on the vehicle when it arrives and departs that station. Constraints (24) and (25) ensure the relationship between the number of loading or unloading bikes and the capacity of the repositioning vehicle.

Constraint (26) defines z_{ij} as a binary variable. Constraint (27) ensures that the repositioning vehicle cannot visit the same node in a single route. Constraint (28) defines that the number of loading and unloading bikes at a station are non-negative integers. Constraints (29) and (30) define the auxiliary variables as non-negative.

4. Solution algorithm

Both the HLP and VRP are widely acknowledged to be NP-hard (Dell'Amico et al., 2014). Applications of exact algorithms, such as branch-and-cut, branch-and-price. on real-size networks are limited due to the overwhelming computational burden when dealing with a large number of variables and constraints (Ho & Szeto, 2017). Hence, heuristics or metaheuristics are widely applied to obtain good solutions in relatively short computing time. Some of the examples include chemical reaction optimization (Szeto et al., 2016), the genetic algorithm (GA) (Szeto et al., 2016; Bie et al., 2020), and the particle swarm optimization (PSO) (Chen et al., 2019). This study employed the ABC algorithm (Karaboga & Akay, 2009; Szeto & Jiang, 2014; Huang et al., 2016; Liu et al., 2017). It is proven to outperform other existing evolutionary algorithms because of its inherent local search mechanism. In this algorithm, three colonies of bees are introduced: employed bees, onlookers, and scouts. The employed bees take charge of exploring food sources (i.e., solutions to the optimization model) until the food source is exhausted (i.e., a local search subroutine). The onlookers and scouts evaluate and search for new food sources, respectively.

In this algorithm, each food source indicates a feasible solution, the representation of which should be specifically designed to search all possible hub-and-spoke structures. Fig. 3 illustrates the structure of a feasible solution in the ABC algorithm. A solution consists of N elements, each of which represents a node in the network. The value of each element represents the number of hubs the node is allocated to. A node is considered as a hub if the value of the element is equal to the number of the node.

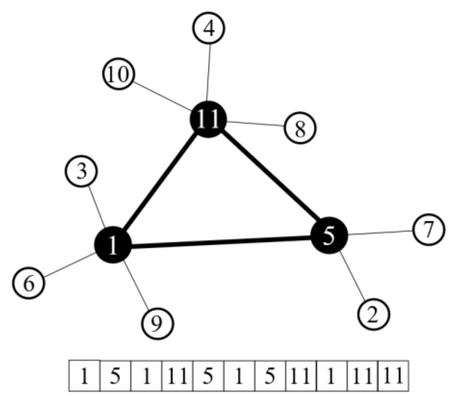


Figure 3. Illustration of a food source (a feasible solution).

Firstly, an initial population is randomly generated that contains N_{ini} food sources, where $N_{ini} = N_c/2$, and N_c is the total population size. Each food source represents a feasible solution, which is a N-dimensional vector. The other parameters involved in the ABC algorithm are introduced as follows: the number of employed bees N_e; the number of onlooker bees N_o; the number of scout bees N_s; the limit counter $L_{\max}$; the counter of iterations I; and the maximum number of iterations $I_{\max}$.

The pseudo-code for the ABC algorithm is provided as follows.

ABC Algorithm

Input: N_c, N_e, N_o, N_s, $L_{\max}$, I, $I_{\max}$

Output: x_{ij}, y_{ij}

1: **Initialization:** Initialize the food sources z_m, $m = 1,\ldots,N_c$;
Evaluate the fitness of the population, fit_m;
Send the employed bees to the current food sources;
Set $I = 1$;

while $I < I_{\max}$ **do**
  for each employed bee **do**
    Find a new food source in its neighborhood using
    $z'_{m,n} = z_{m,n} + \phi(z_{m,n} - z_{k,n})$;
    Caculate the fitness of the new food source;
    Apply the greedy selection process;
  end for
  Calculate the probability values P_m for each solution z_m
  using $P_m = fit_m \Big/ \sum_{k=1}^{N_c} fit_k$;
  for each onlooker bee **do**
    Select a food source z_m depending on P_m;
    Find a new food source in its neighborhood;
    Caculate the fitness of the new food source;
    Apply the greedy selection process;
  end for
  if any employed bee becomes scout bee **then**
    Randomly send the scout bee to a new food source;
  end if
  Memorize the best solution obtained so far;
  $I = I + 1$;
end while

The ABC algorithm is first applied to solve P1 of the hubbing stage. A small set of hub stations is obtained as the input of the routing stage. It decreases the computational burden while addressing the routing problem. In this regard, the routing stage could be solved by commercial solvers efficiently.

5. Numerical experiments

In this section, numerical experiments are conducted to evaluate and compare the proposed model and algorithm with existing studies. The algorithm is coded in python, while all computational experiments are conducted on an Intel Core i7-9750H CPU at 2.60 GHz with 16 GB RAM.

5.1 Experimental setup

The proposed model and algorithm are tested with random instances including, 25- 200 stations following the random generation method adopted in the literature (Toth & Vigo, 2002; Dell’Amico et al., 2014). The number of hubs is predetermined, which is given by 10%, 20%, and 30% of the total number of stations. The coordinates of the stations are generated randomly between 0 and 100. The depot is set in $(50,50)$. The cost matrix depends on the Euclidean distance.

5.2 Comparison between different HFRS approaches

The first set of experiments aims to analyze the performance of different clustering approaches in finding optimal hub locations. The optimal hub location-decision problem of hub location can be classified into two categories: discrete and continuous HLPs (Farahani et al., 2013). In the first case, the HLP is formulated in MIP on a strongly connected network where all nodes can be considered as candidate hubs. Whilst in the second case, the solution is a subset of points in the plane. In the following subsection, we first compare the computational performance of the proposed HFRS with the clustered routing approach based on the maximum spanning star (Schuijbroek et al., 2017). Besides, we also adopt the geographical clustering approach as a benchmark against the discrete HLP.

5.2.1 Clustering based on MIP

To compare the effectiveness of the optimization model, both HFRS and clustered routing problems are solved by the MIP solver GUROBI 9.0. Table 4 summarizes the computational results for random instances. Each instance runs 20 times. The computational time required to run the algorithm is measured in CPU seconds. The time limit is set to 3600 seconds.

Table 4 shows that the number of hubs has a significant influence on the routing cost and the total dissatisfaction (represented by the user cost). The increase in the number of hubs may improve routing efficiency as the repositioning vehicle does not need to visit remote stations to guarantee the coverage of services. The impact of hubs on user costs varies with the instance scale. For example, for $|V| = 25$ instance families, the user cost increases with more hubs.

In Table 4, computational complexity is reported in terms of the running time and optimality gap. It can be observed that the proposed HFRS finds the optimal solutions in 2 minutes for randomly generated instances. Whereas, the clustered routing approach could hardly obtain a feasible solution within the given time limit, which is in accordance with Schuijbroek et al. (2017). In the case of large-scale instances, the computational complexity lies in the routing problem with an approximation of routing distance. Therefore, the optimal solution of these instances is reported in the table as "-". As the hubbing and routing stages in the proposed HFRS are solved separately, the reduction in the problem size and the search space, in turn, speeds the computation time. The experiments on randomly generated instances also confirm the quality of the proposed formulations.

Table 4. Comparison results of the proposed HFRS and the clustered routing.

$\|V\|$	$\|\bar{V}\|$	HFRS				Clustered routing			
		Routing cost	User cost	CPU time	Gap (%)	Routing cost	User cost	CPU time	Gap (%)
25	3	481.35	5.00	0.09	0.00	426.74	74.80	0.48	0.63
	5	312.60	6.00	0.09	0.00	286.07	59.84	3.85	7.82
	8	200.41	31.80	0.10	0.00	114.28	176	3600	17.00
50	5	755.84	38.00	0.37	0.00	312.66	135.00	356.88	0.23
	10	447.19	61.00	0.36	0.00	152.79	110.4	3600	40.40
	15	306.59	60.20	0.37	0.00	-	-	3600	76.10
75	8	858.80	80.60	0.77	0.00	-	-	3600	63.80
	15	533.07	69.40	0.93	0.00	-	-	3600	80.30
	23	358.26	78.00	0.92	0.00	-	-	3600	64.30
100	10	1055.49	93.40	1.46	0.00	-	-	3600	71.40
	20	640.56	54.20	1.81	0.00	-	-	3600	79.00
	30	449.80	85.40	3.25	1.26	-	-	3600	68.00
125	13	1136.59	96.60	2.65	0.00	-	-	3600	86.20
	25	707.41	63.80	5.45	0.46	-	-	3600	82.50
	38	485.85	83.60	6.53	1.02	-	-	3600	66.20
150	15	1258.41	67.80	4.75	0.00	-	-	3600	88.50
	30	743.00	63.80	6.34	0.01	-	-	3600	81.60
	45	510.29	113.20	21.37	0.82	-	-	3600	70.40
200	20	1434.91	142.86	9.958	0.00	-	-	3600	98.50
	40	845.36	135.43	13.87	0.26	-	-	3600	81.90
	60	581.02	178.00	68.97	0.41	-	-	3600	66.80

To further compare the performance of the proposed HFRS and the clustered routing approach, we consider the instance with $|V| = 50$ and $|\bar{V}| = 5$. For a fair comparison, the unmet demands in both models are calculated based on the same service level, which guarantees station inventory after rebalancing operations. According to Schuijbroek et al. (2017), the service level of a station can be determined from the observation of user demand, which is modeled as a stochastic process. For simplicity, the service levels of all the stations are generated randomly in this paper.

Table 5. Comparison results with $|V| = 50$ and $|\bar{V}| = 5$.

	Unmet demand		Fleet size	
	HFRS	Clustered routing	HFRS	Clustered routing
Mean	64.85 (0.09[a])	210.10 (0.28)	5+1	5
Max	171.00 (0.23)	356.00 (0.43)	5+1	5
Std. dev.	39.17 (0.05)	76.07 (0.09)	5+1	5
	Cluster routing cost		**Total routing cost**	
	HFRS	Clustered routing	HFRS	Clustered routing
Mean	273.75	310.85	1368.80	1554.28
Max	315.21	389.42	1576.05	1947.12
Std. dev.	16.45	26.69	82.27	133.47

[a]Note: The value in the parenthesis represents the ratio of unmet demand over total demand.

Table 5 summarizes the statistics for different approaches. Due to the arbitrary setting of the number of hubs, the fleet size of HFRS is always one more than that of the clustered routing approach, where the inter-hub bike repositioning is served by a dedicated vehicle. In this regard, the total routing cost of HFRS could be obtained by the summation of within- and inter-hub routing costs, while the total routing cost of the clustered routing is composed of the within-cluster cost and the routing costs between the depot and the first/last stations of each cluster. Due to the inter-hub redistribution, the HFRS reduces the overall unmet demand on average by 20% when compared to the clustered routing approach. The worst-case unmet demand of the HFRS is lower than the average unmet demand of the clustered routing approach. The proposed HFRS also performs better in terms of reducing the routing cost within clusters, as well as the total routing cost.

5.2.2 Geographical clustering

In practice, the spatial distribution of bike docking stations depends highly on the land-use type, population density, and other demographic/environmental issues. Hence, in the areas with high station density (i.e., commercial and residential areas), the number of links between each station would increase exponentially and incur additional computational complexity. For this reason, two classical geographical clustering approaches, namely the agglomerative hierarchical clustering (AHC) and the density-based spatial clustering of applications with noise (DBSCAN), are adopted as the benchmark for our proposed model.

Table 6. Comparison results of the geographical clustering approaches.

$\lvert V \rvert$	AHC				DBSCAN			
	$\lvert \bar{V} \rvert$	Routing cost	User cost	CPU time	$\lvert \bar{V} \rvert$	Routing cost	User cost	CPU time
	3	364.91	74.84	0.01				
25	5	350.39	59.84	0.01	2.40	383.67	74.82	0.01
	8	367.40	44.88	0.01				
	5	696.35	246.2	0.01				
50	10	620.90	196.96	0.01	3.80	752.82	246.20	0.01
	15	628.57	147.72	0.01				
	8	889.42	350.2	0.01				
75	15	839.60	280.16	0.01	5.00	900.56	350.20	0.01
	23	908.66	210.12	0.01				
	10	1174.48	460.37	0.01				
100	20	1040.98	368.48	0.01	5.20	904.64	341.65	0.01
	30	1158.88	276.54	0.01				
	13	1273.13	565.27	0.01				
125	25	1253.93	452.16	0.01	7.60	1443.38	565.20	0.01
	38	1430.11	339.12	0.01				
	15	1424.95	630.24	0.01				
150	30	1421.54	504.16	0.01	9.20	1458.02	630.20	0.01
	45	1641.76	378.12	0.01				
	20	1848.53	919.80	0.01				
200	40	1842.45	735.84	0.01	12.60	2214.85	919.80	0.01
	60	2147.17	551.88	0.01				

Table 6 reports the computational results for the same instance used in Section 5.2.1. It shows that both AHC and DBSCAN can solve the clustering problem in a relatively short time. The number of clusters does not have a significant influence on the routing cost. The user cost (i.e., the amount of unsatisfied demand) is higher than that of the mathematical programming approach. The reason for this might be because the objective of the geographical clustering approach is to cluster/classify the stations with spatial similarity. However, in the bike-sharing system, the bike docking station is characterized by other features, such as inventory level and temporal distribution of bike demand. Hence, the geographical clustering approach can be considered as an efficient tool for the long-term location planning of the bike-sharing system when the dynamics of user demand is not crucial during the decision process.

5.3 Heuristics performance

In this section, we display an experiment of the comparison between ABC algorithm and popular metaheuristics i.e., GA and PSO. Fig. 4 shows a single run of the three metaheuristics on a random instance with 50 stations. It shows that the ABC converges faster than both GA and PSO and has the better solution quality than the other two heuristics.

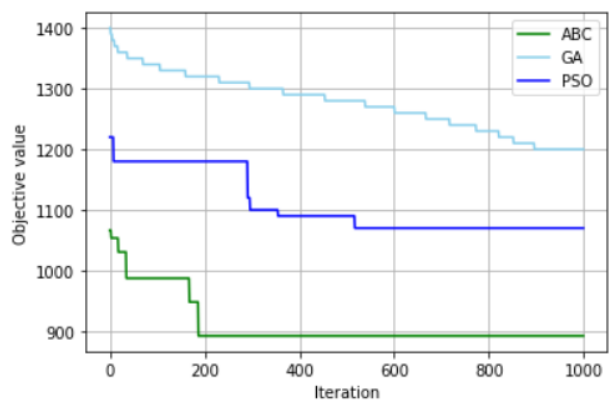


Figure 4. Convergence performance of metaheuristics.

6. Case study

In this section, a case study for real-world bike-sharging system in Nanjing, China, is presented.

6.1 Data description

The initial data set consists of the bike-ride information in the form of IC card serial numbers, rental and return stations, and corresponding timestamps (see Table 7). The data covers approximately 4,932,902 rides from October 1st to December 31st, 2016, in the downtown area of Nanjing, China (see Fig. 5).

Data preprocessing is carried out to remove invalid data records while following various criteria (Vogel et al., 2011), including: i) rides that last less than 60 seconds, which start and end at the same stations; ii) rides that have negative trip durations (caused by system error); iii) rides that have incomplete records; and iv) stations with only a few pickups or returns records. After removing the abnormal records, the number of valid trip data and the number of stations is reduced to 2,487,737 and 151, respectively.

Table 7. Data structure.

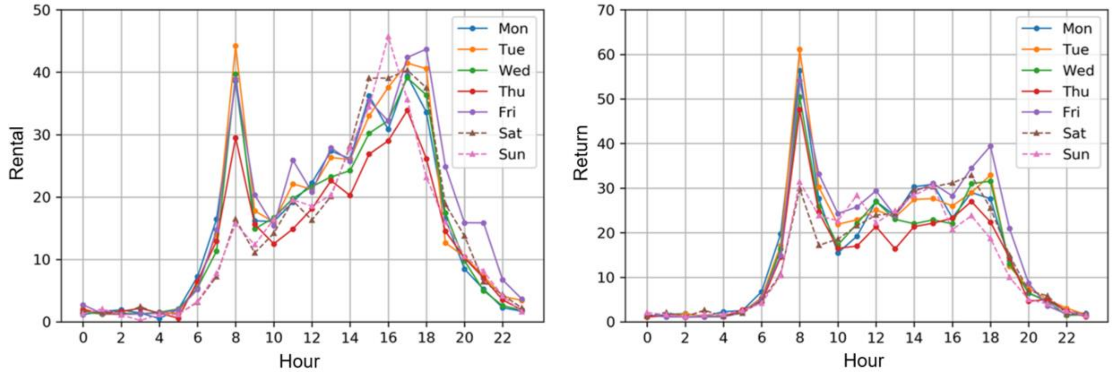
No. of IC card	No. of bike	Rental station	Rental timestamp	Return station	Return timestamp
10225940	2008129	13029	2016/12/1 8:00:11	13029	2016/12/1 8:15:54

Figure 5. Distribution of bike stations in the downtown area of Nanjing.

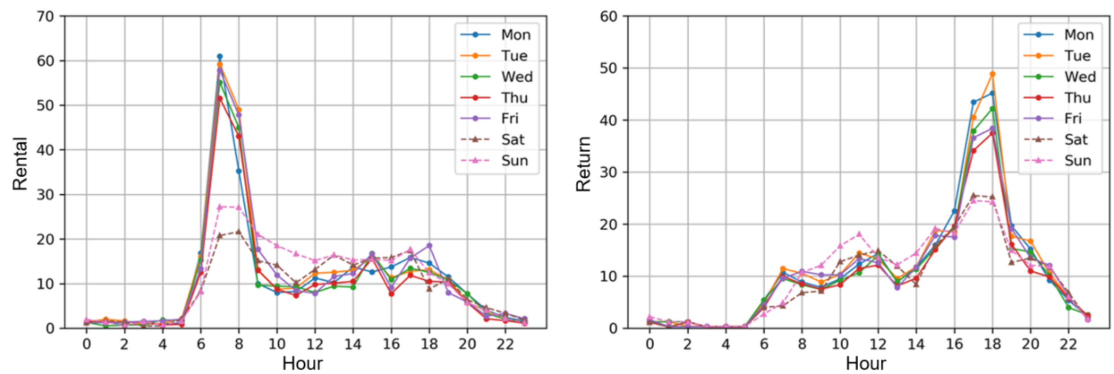
6.2 Feature engineering

Existing literature usually categorizes the input feature of demand forecasting for the bike-sharing system into two types: time-related and weather (Hulot et al., 2018). To comprehensively analyze the characteristics of the usage of bikes, two more types of features are further investigated, namely, usage and spatial features.

(1) Time-related feature



(a) Xinjiekou Rail Station

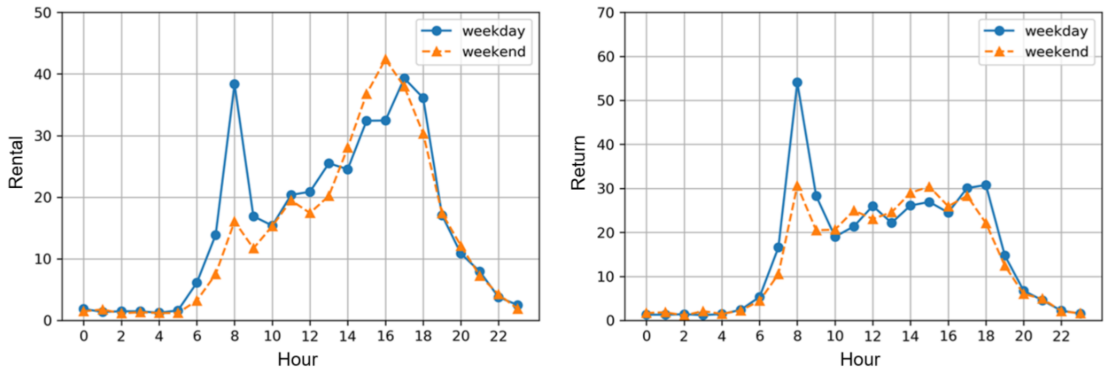


(b) Vanke Bright City Station

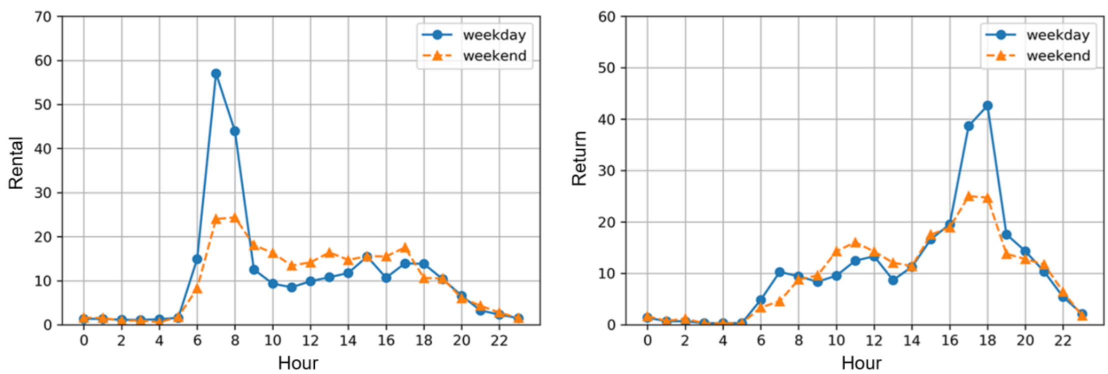
Figure 6. Rentals and returns over the course of the day for selected stations.

The time-related features include the month, the day, the weekday, and the hour of a day. In this study, hour and day type (i.e., weekday or weekend) are selected as influential features for the usage of bikes. To validate the rationale of this selection, we present the statistics results of two typical docking stations, i.e., Xinjiekou Rail Station and Vanke Bright City Station, located in commercial and residential areas, respectively.

As presented in Fig. 6, the bike usages in both stations show significant tidal characteristics. The usage of the Xinjiekou Rail Station shows a double-peak distribution w.r.t the morning and evening rush hours. Most of the rentals and returns in the residential area are aggregated in the morning rush hour and evening hour, respectively. We can also observe that the bike usage of the station in the commercial area stays at a higher level than that of the residential areas in the non-rush hours. Fig. 7 presents the bike usage during a weekday and weekend. The bike usages of both stations in the morning rush hours of the weekend decrease significantly. However, the distributions of demand in the non-rush and evening rush hours are quite similar during both weekdays and weekends.



(a) Xinjiekou Rail Station



(b) Vanke Bright City Station

Figure 7. Rentals and returns on a weekday and weekend for selected stations.

(2) Weather feature

It has been widely acknowledged that the traveler's willingness to ride a bike is more sensitive to weather conditions than any other travel modes (Lathia et al., 2012). In general, the weather-related features include temperature, humidity, visibility, wind speed, air pressure, and weather (cloudy, sunny, rainy, snow, etc.). Figs. 8 and 9 depict the influence of temperature and weather on the bike usage. It shows that users are more willing to ride bikes when it is warmer on sunny or cloudy days.

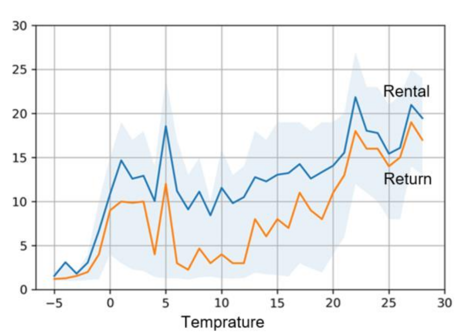


Figure 8. Influence of temperature on bike usage.

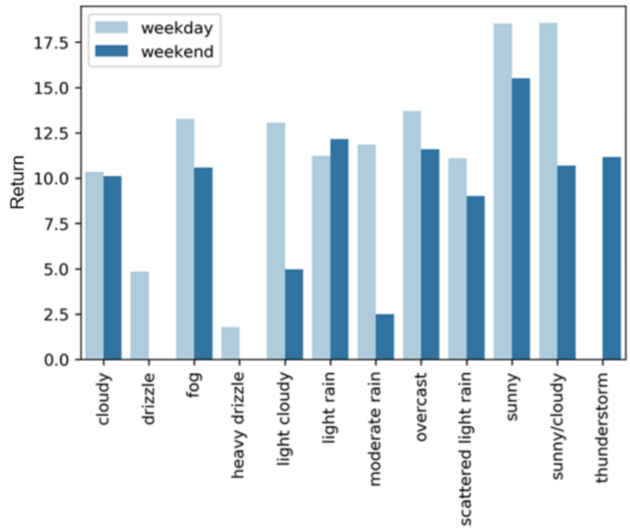


Figure 9. Influence of weather on bike usage.

(3) Usage feature

The usage feature is related to the passenger's historical bike usage trend for some stations in time series. Denote the target hour as k and day as n. To reflect the periodical changes in the bike usage of each station, we extract the bike usage information in adjacent hours, i.e., $k-2$, $k-1$, and $k+1$, as well as the same hour of past days, i.e., $n-1$, $n-2$, $n-3$ and $n-7$.

(4) Spatial feature

Fig. 10 presents the distribution of riding distance and time over the course of a day. It shows that the average distance is between 1.5~2.5 km, which is in accordance with the role of the bike-sharing system that aims to solve the last mile problem. The average riding time reaches the peak value in the morning rush hour.

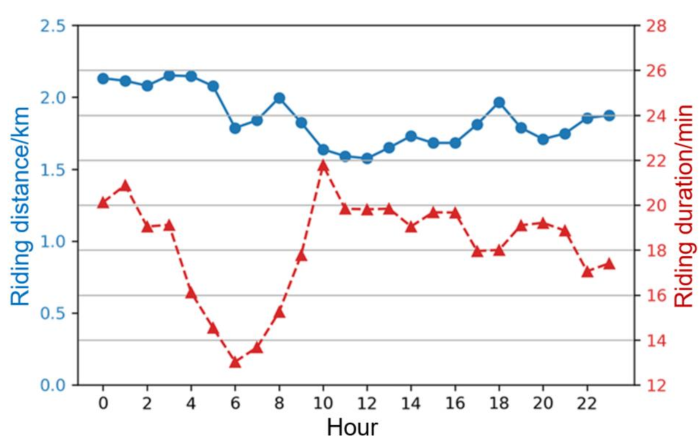


Figure 10. Average riding distance and duration.

Table 8 presents the list of variables used in the analysis of the demand forecasting model. To measure the importance of the selected features, the RF model provides a unique indicator called the variable importance, VI. As mentioned in Section 3.2, the bootstrap sampling randomly extracts samples from the training set. The out of bag error ($OOBE$) is used to calculate the average prediction error of the observations in the trees that do not contain these observations, thus, providing built-in cross-validation. Hence, the $OOBE$ can also be called the generalization error, which can be calculated as follows:

$$OOBE = \frac{1}{L}\sum_{k=1}^{L}\left(Y_k - \hat{Y}_k\right)^2 . \tag{31}$$

By definition, the difference between decision trees lies in both samples and features. Hence, VI can then be obtained by transposing a variable and averaging the difference of $OOBE$ before and after transposing over all trees. Hence the importance of the k-th variable can be obtained as follows

$$VI\left(X_k\right) = \frac{1}{L}\sum_{k=1}^{L}\left(OOBE_k - OOBE\right), \tag{32}$$

where $OOBE_k$ is the new $OOBE$ obtained after transposing. The degree-of-importance of the explanatory variables considered in the demand forecasting model is shown in Fig. 11.

Table 8. Influencing factors of the demand forecasting model.

Variable	Variable description
X_1	number of bikes taken in the $k-1$ hour of day $n-3$
X_2	number of bikes taken in the k hour of day $n-3$
X_3	number of bikes taken in the $k+1$ hour of day $n-3$
X_4	number of bikes taken in the $k-1$ hour of day $n-2$
X_5	number of bikes taken in the k hour of day $n-2$
X_6	number of bikes taken in the $k+1$ hour of day $n-2$
X_7	number of bikes taken in the $k-1$ hour of day $n-1$
X_8	number of bikes taken in the k hour of day $n-1$
X_9	number of bikes taken in the $k+1$ hour of day $n-1$
X_{10}	number of bikes taken in the $k-1$ hour of day n
X_{11}	number of bikes taken in the $k-2$ hours of day n

X_{12}	number of bikes taken in the $k-1$ hour of day $n-7$
X_{13}	number of bikes taken in the k hour of day $n-7$
X_{14}	number of bikes taken in the $k+1$ hour of day $n-7$
X_{15}	average time of bikes taken in the k hour of day $n-1$
X_{16}	average riding distance in the k hour of day $n-1$
X_{17}	temperature in the n-th day
X_{18}	weather in the n-th day
X_{19}	weekday or weekend
y	number of bikes taken in the k hours of day n

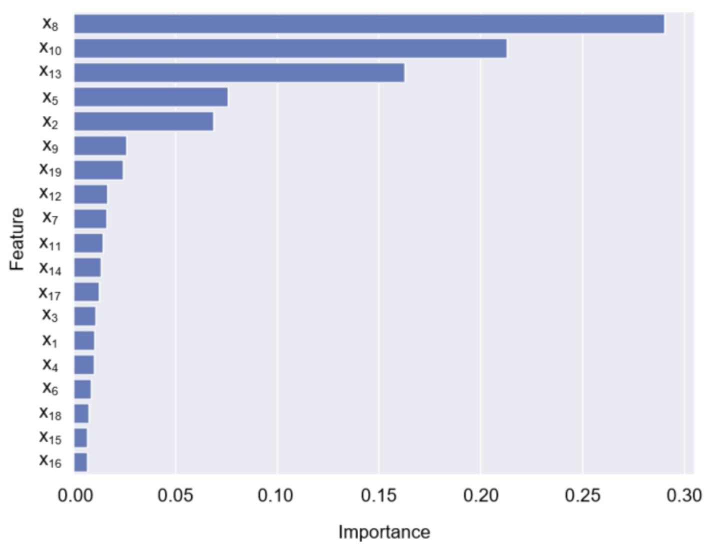


Figure 11. The importance degree of explanatory variables.

6.3 Lost demand estimation

As discussed in Section 2.2, the lost or unmet demand happens due to a lack of bikes or docks. It is difficult to capture the number of lost users due to the lack of valid observations from trip data of the bike-sharing system. However, to guarantee the system's level-of-service and decrease the number of unmet demands, it is essential to consider lost demand from historical data in repositioning operation. In this study, a pure data-driven estimation approach proposed by Mellou & Jailet (2019) is adopted, which is based on a basic assumption that historical trip data can reveal the user behavior of each station.

We can extract two kinds of behavior: i) average station behavior: the user behavior of the station at the time interval of the previous days; and ii) daily demand trend: the user behavior around that time interval of the same day when bikes/docks are available. By definition, the lost outgoing demand of station i in time interval $[t,t']$, $\text{AVE}\_q_{out,i}^{t,t'}$, the station behavior can be obtained by averaging the observed demand in previous days when station i is not empty. Note that we can also calculate the lost incoming demand in the same way. The demand trend of the target interval within a day is related to the user behavior on adjacent intervals in a time series. Let $r_{out,i}^{t_1,t_2}$ denote the outgoing demand rate of station i in time interval $[t_1,t_2]$, which measures the number of users departing from station i per minute:

$$r_{out,i}(t_1,t_2)=\frac{1}{t_2-t_1}\sum_{t=t_1}^{t_2}q_{out,i}^{t}\ , \tag{33}$$

where $q_{out,i}^{t}$ is the observed outgoing demand from station i at time t. The outgoing demand rate of station i in time interval $[t,t']$ can then be calculated by using the demand rate before and after this interval, that is,

$$r_{out,i}(t,t')=\frac{1}{2}\left[r_{out,i}(t-60,t)+r_{out,i}(t',t+60)\right]. \tag{34}$$

The estimation of lost outgoing demand for the daily demand trend, $\text{TREND}\_q_{out,i}^{t,t'}$, can be obtained as follows,

$$\text{TREND}\_q_{out,i}^{t,t'}=r_{out,i}(t,t')\cdot(t'-t). \tag{35}$$

The total lost outgoing demand can then be obtained by the convex combination of these two estimations:

$$\tilde{q}_{out,i}^{t,t'}=\lambda\cdot\text{AVE}\_q_{out,i}^{t,t'}+(1-\lambda)\cdot\text{TREND}\_q_{out,i}^{t,t'}\,. \tag{36}$$

where λ is a parameter, and $0<\lambda<1$.

6.4 Evaluation of forecasted results

The RF method is then implemented using the Scikit-learn framework in python (Pedregosa et al., 2011). The data set is separated into two subsets: the training set comprising of records from October 1st to November 29th and the test set with records in the peak hours of November 30th. Additionally, two essential parameters in the RF model should be specified: i) the number of decision trees, $L = 500$, and ii) the number of branching characteristic variables, $m = 7$. Fig. 12 shows the scatter diagram of actual values and forecasted values. The forecasted values are mainly distributed near the line $y = x$, which validates the accuracy of the proposed prediction method.

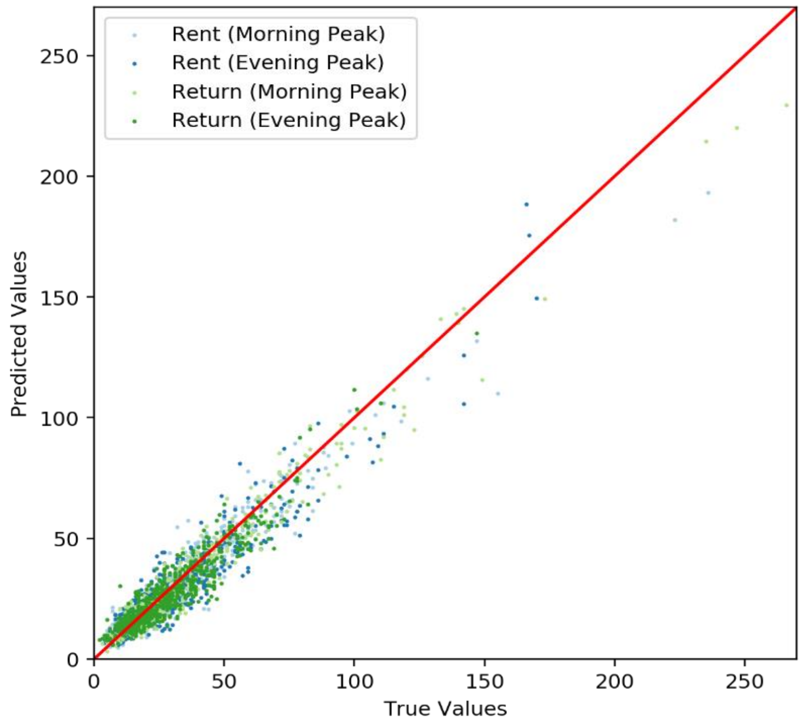


Figure 12. The distribution of forecasted values.

To compare the effectiveness of the forecasting model, three other prevalent models which have been widely used in the bike-sharing system are applied, namely, the Linear Regression (LR) model (Rudloff & Lackner, 2014), the Neural Network (NN) model (Ruffieux et al., 2017), and the Autoregressive Integrated Moving Average (ARIMA) model (Dias et al., 2015). The comparison of the evaluation values of the forecasted results is shown in Table 9. It shows that the RF model outperforms the other three forecasting models in both volatility (R_{MSE}) and accuracy (M_{AE} and M_{APE}). The performance of ARIMA is relatively closer to that of the RF model. The main reason for the improved performance of ARIMA is that it considers the lost demand inherently by averaging the user demand from previous days.

Table 9. Evaluation of the forecasted results.

Index	LR			NN		
	R_{MSE}	M_{AE}	M_{APE}	R_{MSE}	M_{AE}	M_{APE}
Rentals in morning peak hour	13.18	9.58	29.98%	11.56	8.07	24.12%
Returns in morning peak hour	15.32	10.32	28.44%	12.23	8.01	22.92%
Rentals in evening peak hour	10.51	7.32	28.24%	9.54	6.69	24.31%
Returns in evening peak hour	8.69	6.68	26.74%	7.48	5.70	22.55%
Index	**ARIMA**			**RF**		
	R_{MSE}	M_{AE}	M_{APE}	R_{MSE}	M_{AE}	M_{APE}
Rentals in morning peak hour	9.40	6.98	25.98%	7.45	5.31	19.22%
Returns in morning peak hour	8.46	5.99	21.37%	6.70	4.69	16.87%
Rentals in evening peak hour	8.14	6.06	26.14%	7.40	5.38	21.40%
Bike returned in evening peak hour	7.16	5.56	23.99%	6.27	4.86	20.43%

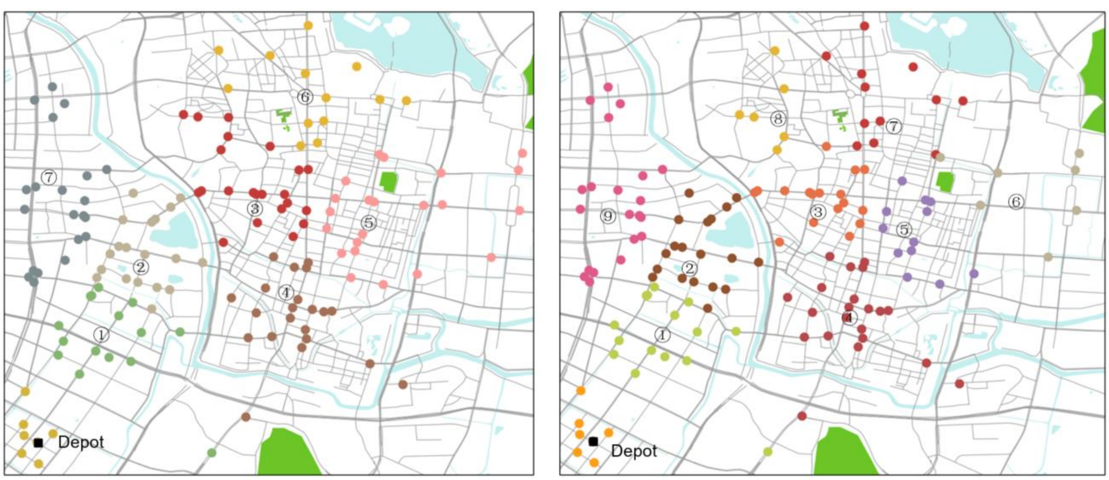
6.5 Bike repositioning based on the prediction results

Based on the predicted results, the proposed bike repositioning strategy is implemented in the morning and evening rush hours, respectively. The optimal results of the routing distance and the proportion of unmet demand to the total demand are summarized in Table 10. The ratio between the total predicted demand (the sum of rentals and returns) and total real demand is over 63%, which also illustrates that the forecasting model captures most of the passenger demand. After the repositioning process, the ratio of met demand reaches nearly 90% of total demand.

It is found that the influence of the number of hubs on the routing cost is not significant. In the morning peak hour, the influence of the number of hubs on the unmet demand is more significant, while fewer hubs are more efficient as it enables transporting of bikes to exhaust stations in a short time. In the evening peak hour, about 90% of the total demand can be satisfied after the repositioning strategy. The details of the repositioning routes are shown in Fig. 13. The number in each cluster represents the optimal visiting sequence.

Table 10. Optimal routing distance and unmet demand in different time-or-day.

Time-of-day	No. of hubs	Routing distance (km)	Total demand (predicted)	Total demand (real)	Ratio (%)	Unmet demand (%)
7:00 a.m.	8	22.5	6848	10792	0.63	0.09
	10	22.8				0.30
	12	23.1				0.14
8:00 a.m.	8	22.8	9488	13241	0.72	0.04
	10	22.8				0.16
	12	22.3				0.11
5:00 p.m.	8	22.1	9375	12686	0.74	0.10
	10	21.1				0.11
	12	20.9				0.12
6:00 p.m.	8	19.7	8767	10336	0.85	0.17
	10	20.0				0.10
	12	21.1				0.17



(a) 7:00 a.m. (8 hubs)

(b) 6:00 p.m. (10 hubs)

Figure 13. The optimal bike repositioning route in the morning and evening rush hours.

7. Conclusions

In this paper, a static bike repositioning strategy is proposed and investigated. The paper addressed three principal issues concerning the efficiency of the bike-sharing system: i) the future demand of each station, ii) the visiting sequence of the repositioning vehicle, and iii) the number of bikes that need to be loaded or unloaded at visiting stations. The customer demand for both rental and return can be forecasted accurately with the help of the machine learning approach. The RF model is applied to forecast demand in the bike-sharing system considering four influencing factors that have evident effects on customer demand. Based on the demand forecasting, an HFRS repositioning strategy is proposed. The hubbing stage is described as an HLP, which is modeled by an integer programming aiming to identify the hub stations. In the hub network, the demands of non-hub stations are allocated to hubs. It increases the operational efficiency of the bike-sharing system as the repositioning vehicle only needs to visit the hub stations. The visiting sequence of repositioning vehicles and the number of loading/unloading bikes are determined simultaneously in the routing stage. A comprehensive comparison is also conducted to verify the effectiveness of the proposed model. The results show that: i) the RF model outperforms other models in the short-term prediction of bike usage; ii) the clustering based on MIP could achieve more reasonable results than the geographical clustering algorithms by considering some unique features of the bike-sharing system, such as the inventory level and temporal distribution of bike demand.

Future studies can examine the impact of several potential enhancements. First, several influencing factors could be considered in the demand forecast model, including the characteristics of the station, such as land use, capacity. Second, the routing stage can further be extended to the multi-vehicle or the multi-depot cases. Third, the number of hubs can also be optimized along with the HLP, considering the tradeoffs between the vehicle routing cost and the customer's dissatisfaction.

References

Albiński, S., Fontaine, P., & Minner, S. (2018). Performance analysis of a hybrid bike sharing system: A service-level-based approach under censored demand observations. *Transportation Research Part E, 116*, 59-69.

Alumur, S., & Kara, B. Y. (2008). Network hub location problems: The state of the art. *European Journal of Operational Research, 190*(1), 1-21.

Bai, L., Liu, P., Chan, C. Y., & Li, Z. (2017). Estimating level of service of mid-block bicycle lanes considering mixed traffic flow. *Transportation Research Part A, 101*, 203-217.

Bie, Y., Xiong, X., Yan, Y., & Qu, X. (2020). Dynamic headway control for high-frequency bus line based on speed guidance and intersection signal adjustment. *Computer-Aided Civil and Infrastructure Engineering, 35*, 4-25.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and Regression Trees*: CRC Press, Boca Raton, Florida, USA.

Bulhões, T., Subramanian, A., Erdogan, G., & Laporte, G. (2018). The static bike relocation problem with multiple vehicles and visits. *European Journal of Operational Research, 264*(2), 508-523.

Chemla, D., Meunier, F., & Calvo, R. W. (2013). Bike sharing systems: Solving the static rebalancing problem. *Discrete Optimization, 10*(2), 120-146.

Chen, Z., Hu, Y., Li, J., & Wu, X. (2019). Optimal deployment of electric bicycle sharing stations: Model formulation and solution technique. *Networks and Spatial Economics*, 1-38.

Dondo, R., & Cerdá, J. (2007). A cluster-based optimization approach for the multi-depot heterogeneous fleet vehicle routing problem with time windows. *European Journal of Operational Research, 176*(3), 1478-1507.

Contardo, C., Morency, C., & Rousseau, L. M. (2012). *Balancing a Dynamic Public Bike-sharing System.* Technical Report CIRRELT-2012-09, Montréal, Québec, Canada.

Datner, S., Raviv, T., Tzur, M., & Chemla, D. (2017). Setting inventory levels in a bike sharing network. *Transportation Science, 53*(1), 62-76.

Dell'Amico, M., Hadjicostantinou, E., Iori, M., & Novellani, S. (2014). The bike sharing rebalancing problem: Mathematical formulations and benchmark instances. *Omega, 45*, 7-19.

Di Gaspero, L. D., Rendl, A., & Urli, T. (2013). *A hybrid ACO+CP for balancing bicycle sharing systems*. In *Proceedings of International Workshop on Hybrid Metaheuristics*, 198-212.

Di Gaspero, L. D., Rendl, A., & Urli, T. (2016). Balancing bike sharing systems with constraint programming. *Constraints, 21*(2), 318-348.

Dias, G. M., Bellalta, B., & Oechsner, S. (2015). Predicting occupancy trends in Barcelona's bicycle service stations using open data. In *SAI Intelligent Systems Conference*, 439-445.

Erdoğan, G., Laporte, G., & Calvo, R. W. (2014). The static bicycle relocation problem with demand intervals. *European Journal of Operational Research, 238*(2), 451-457.

Farahani, R. Z., Hekmatfar, M., Arabani, A. B., & Nikbakhsh, E. (2013). Hub location problems: A review of models, classification, solution techniques, and applications. *Computers & Industrial Engineering, 64*(4), 1096-1109.

Fishman, E. (2016). Bikeshare: A review of recent literature. *Transport Reviews, 36*(1), 92-113.

Forma, I. A., Raviv, T., & Tzur, M. (2015). A 3-step math heuristic for the static repositioning problem in bike-sharing systems. *Transportation Research Part B, 71*, 230-247.

Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The Elements of Statistical Learning*: Springer, New York, USA.

Gelareh, S., & Nickel, S. (2011). Hub location problems in transportation networks. *Transportation Research Part E, 47*(6), 1092-1111.

Ghosh, S., Trick, M., & Varakantham, P. (2016). Robust repositioning to counter unpredictable demand in bike sharing systems. In *25th International Joint Conference on Artificial Intelligence*, 3096-3102.

Ghosh, S., Varakantham, P., Adulyasak, Y., & Jaillet, P. (2017). Dynamic repositioning to reduce lost demand in bike sharing systems. *Journal of Artificial Intelligence Research, 58*, 387-430.

Ghosh, S., & Varakantham, P. (2017). Incentivizing the use of bike trailers for dynamic repositioning in bike sharing systems. In *27th International Conference on Automated Planning and Scheduling*, 373-381.

Ghosh, S., Koh, J. Y., & Jaillet, P. (2019). Improving customer satisfaction in bike sharing systems through dynamic repositioning. In *28th International Joint Conference on Artificial Intelligence*, 5864-5870.

Goh, C. Y., Yan, C., & Jaillet, P. (2019). Estimating primary demand in bike-sharing systems. *Available at SSRN 3311371*.

Guido, C., Rafał, K., & Constantinos, A. (2019). A low dimensional model for bike sharing demand forecasting. In *2019 6th International Conference on Models and Technologies for Intelligent Transportation Systems*, 1-7.

Haider, Z., Nikolaev, A., Kang, J. E., & Kwon, C. (2018). Inventory rebalancing through pricing in public bike sharing systems. *European Journal of Operational Research, 270*(1), 103-117.

Ho, S. C., & Szeto, W. (2014). Solving a static repositioning problem in bike-sharing systems using iterated tabu search. *Transportation Research Part E, 69*, 180-198.

Ho, S. C., & Szeto, W. (2017). A hybrid large neighborhood search for the static multi-vehicle bike-repositioning problem. *Transportation Research Part B, 95*, 340-363.

Huang, D., Liu, Z., Fu, X., & Blythe, P. T. (2018). Multimodal transit network design in a hub-and-spoke network framework. *Transportmetrica A, 14*(8), 706-735.

Huang, D., Liu, Z., Liu, P., & Chen, J. (2016). Optimal transit fare and service frequency of a nonlinear origin-destination based fare structure. *Transportation Research Part E, 96*, 1-19.

Hulot, P., Aloise, D., & Jena, S. D. (2018). Towards station-level demand prediction for effective rebalancing in bike-sharing systems. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 378-386.

Ji, Y., Fan, Y., Ermagun, A., Cao, X., Wang, W., & Das, K. (2017). Public bicycle as a feeder mode to rail transit in China: The role of gender, age, income, trip purpose, and bicycle theft experience. *International Journal of Sustainable Transportation, 11*(4), 308-317.

Ji, Y., Ma, X., He, M., Jin, Y., & Yuan, Y. (2020). Comparison of usage regularity and its determinants between docked and dockless bike-sharing systems: a case study in Nanjing, China. *Journal of Cleaner Production*, 120110.

Jiang, R., Tang, W., Wu, X., & Fu, W. (2009). A random forest approach to the detection of epistatic interactions in case-control studies. *BMC Bioinformatics, 10*(1), S65.

Jin, J. G., Nieto, H., & Lu, L. (2019). Robust bike-sharing stations allocation and path network design: A two-stage stochastic programming model. *Transportation Letters*, 1-10.

Kadri, A. A., Kacem, I., & Labadi, K. (2016). A branch-and-bound algorithm for solving the static rebalancing problem in bicycle-sharing systems. *Computers & Industrial Engineering, 95*, 41-52.

Kaltenbrunner, A., Meza, R., Grivolla, J., Codina, J., & Banchs, R. (2010). Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system. *Pervasive and Mobile Computing, 6*(4), 455-466.

Karaboga, D., & Akay, B. (2009). A comparative study of artificial bee colony algorithm. *Applied Mathematics and Computation, 214*(1), 108-132.

Lahouar, A., & Slama, J. B. H. (2017). Hour-ahead wind power forecast based on random forests. *Renewable Energy, 109*, 529-541.

Lathia, N., Ahmed, S., & Capra, L. (2012). Measuring the impact of opening the London shared bicycle scheme to casual users. *Transportation Research Part C, 22*, 88-102.

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

Lin, J., & Yang, T. (2011). Strategic design of public bicycle sharing systems with service level constraints. *Transportation Research Part E, 47*(2), 284-294.

Li, Y., Szeto, W. Y., Long, J., & Shui, C. S. (2016). A multiple type bike repositioning problem. *Transportation Research Part B, 90*, 263-278.

Liu, Z., Wang, S., Zhou, B., & Cheng, Q. (2017). Robust optimization of distance-based tolls in a network considering stochastic day to day dynamics. *Transportation Research Part C, 79*, 58-72.

Liu, Z., Chen, X., Meng, Q., & Kim, I. (2018). Remote park-and-ride network equilibrium model and its applications. *Transportation Research Part B, 117*, 37-62.

Liu, Y., Liu, Z., and Jia, R., (2019). DeepPF: A deep learning based architecture for metro passenger flow prediction. *Transportation Research Part C, 101*, 18-34.

Liu, Z., Liu, Y., Meng, Q., and Cheng, Q., (2019). A tailored machine learning approach for urban transport network flow estimation. *Transportation Research Part C, 108*, 130-150.

Lu, C. C. (2016). Robust multi-period fleet allocation models for bike-sharing systems.

Networks and Spatial Economics, 16(1), 61-82.

Mellou, K., & Jaillet, P. (2019). Dynamic resource redistribution and demand estimation: An application to bike sharing systems. *Available at SSRN 3336416.*

Negahban, A. (2019). Simulation-based estimation of the real demand in bike-sharing systems in the presence of censoring. *European Journal of Operational Research, 277*(1), 317-332.

O'Mahony, E., & Shmoys, D. B. (2015). Data analysis and optimization for (Citi) bike sharing. In *29th AAAI Conference on Artificial Intelligence*, 687-694.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825-2830.

Rainer-Harbach, M., Papazek, P., Hu, B., & Raidl, G. R. (2013). Balancing bicycle sharing systems: A variable neighborhood search approach. In *European Conference on Evolutionary Computation in Combinatorial Optimization*, 121-132.

Rainer-Harbach, M., Papazek, P., Raidl, G. R., Hu, B., & Kloimüllner, C. (2015). PILOT, GRASP, and VNS approaches for the static balancing of bicycle sharing systems. *Journal of Global Optimization, 63*(3), 597-629.

Raviv, T., Tzur, M., & Forma, I. A. (2013). Static repositioning in a bike-sharing system: Models and solution approaches. *EURO Journal on Transportation and Logistics, 2*(3), 187-229.

Raviv, T., & Kolka, O. (2013). Optimal inventory management of a bike-sharing station. *IIE Transactions, 45*(10), 1077-1093.

Regue, R., & Recker, W. (2014). Proactive vehicle routing with inferred demand to solve the bikesharing rebalancing problem. *Transportation Research Part E, 72*, 192-209.

Rudloff, C., & Lackner, B. (2014). Modeling demand for bikesharing systems: Neighboring stations as source for demand and reason for structural breaks. *Transportation Research Record, 2430*(1), 1-11.

Ruffieux, S., Spycher, N., Mugellini, E., & Khaled, O. A. (2017). Real-time usage forecasting for bike-sharing systems: A study on random forest and convolutional neural network applicability. In *2017 Intelligent Systems Conference*, 622-631.

Ruffieux, S., Mugellini, E., & Khaled, O. A. (2018). Bike usage forecasting for optimal rebalancing operations in bike-sharing systems. In *2018 IEEE 30th International Conference on Tools with Artificial Intelligence*, 854-858.

Schuijbroek, J., Hampshire, R., & Van Hoeve, W. (2017). Inventory rebalancing and vehicle routing in bike sharing systems. *European Journal of Operational Research, 257*(3), 992-1004.

Shu, J., Chou, M., Liu, Q., Teo, C., & Wang, I. (2013). Models for effective deployment and redistribution of bicycles within public bicycle-sharing systems. *Operations Research, 61*(6), 1346-1359.

Shui, C. S., & Szeto, W. Y. (2018). Dynamic green bike repositioning problem: A hybrid rolling horizon artificial bee colony algorithm approach. *Transportation Research Part D, 60*, 119-136.

Si, H., Shi, J. G., Wu, G., Chen, J., & Zhao, X. (2018). Mapping the bike sharing research published from 2010 to 2018: A scientometric review. *Journal of Cleaner Production, 213*, 415-427

Singla, A., Santoni, M., Bartók, G., Mukerji, P., Meenen, M., & Krause, A. (2015). Incentivizing users for balancing bike sharing systems. In *29th AAAI Conference on Artificial Intelligence*, 723-729

Sohn, J., & Park, S. (2000). The single allocation problem in the interacting three-hub

network. *Networks: An International Journal, 35*(1), 17-25.
Szeto, W. Y., & Jiang, Y. (2014). Transit route and frequency design: Bi-level modeling and hybrid artificial bee colony algorithm approach. *Transportation Research Part B, 67*, 235-263.
Szeto, W. Y., Liu, Y., & Ho, S. C. (2016). Chemical reaction optimization for solving a static bike repositioning problem. *Transportation Research Part D, 47*, 104-135.
Szeto, W. Y., & Shui, C. S. (2018). Exact loading and unloading strategies for the static multi-vehicle bike repositioning problem. *Transportation Research Part B, 109*, 176-211.
Toth, P., & Vigo, D. (2002). *The Vehicle Routing Problem*: SIAM Monographs on Discrete Mathematics and Applications, Philadelphia, Pennsylvania, USA.
Urbica (2016). City Bike rebalanced. Retrieved from https://medium.com/@Urbica.co/city-bike-rebalanced-92ac61a867c7.
Vogel, P., Greiser, T., & Mattfeld, D. C. (2011). Understanding bike-sharing systems using data mining: Exploring activity patterns. *Procedia-Social and Behavioral Sciences, 20*, 514-523.
Xu, C., Ji, J., & Liu, P. (2018). The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets. *Transportation Research Part C, 95*, 47-60.