Journal of Advanced Transportation

Transitions Towards Electrification, Automation, and Shared Mobility for Urban Transport

Lead Guest Editor: Wenxiang Li
Guest Editors: Ziyuan Pu, Linchuan Yang, and Lei Wang

Transitions Towards Electrification, Automation, and Shared Mobility for Urban Transport

Journal of Advanced Transportation

Transitions Towards Electrification, Automation, and Shared Mobility for Urban Transport

Lead Guest Editor: Wenxiang Li
Guest Editors: Ziyuan Pu, Linchuan Yang, and Lei Wang

Contents

Contents

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 7772401, 15 pages
https://doi.org/10.1155/2022/7772401

Research Article

Exploring the Spatially Heterogeneous Effects of the Built Environment on Bike-Sharing Usage during the COVID-19 Pandemic

Hongtai Yang,[1] **Zishuo Guo**,[1] **Guocong Zhai**,[2] **Linchuan Yang**,[3] **and Jinghai Huo**[1]

[1]*School of Transportation and Logistics, National Engineering Laboratory of Integrated Transportation Big Data Application Technology, National United Engineering Laboratory of Integrated and Intelligent Transportation, Institute of System Science and Engineering, Southwest Jiaotong University, Chengdu 611756, China*
[2]*Department of Civil & Environmental Engineering, Old Dominion University (ODU), Norfolk, VA 23529, USA*
[3]*Department of Urban and Rural Planning, School of Architecture, Southwest Jiaotong University, Chengdu 611756, China*

Correspondence should be addressed to Zishuo Guo; 220213442@seu.edu.cn

Received 12 September 2021; Revised 9 June 2022; Accepted 11 August 2022; Published 28 September 2022

Academic Editor: Ricardo Giesen

Bike-sharing holds promise for available and healthy mobility services during COVID-19 where bike sharing users can make trips with lower health concerns due to social distancing compared to the restricted transportation modes such as public transit and ridesharing services. Leveraging the trip data of the Divvy bike-sharing system in Chicago, this study exploresspatially heterogeneous effects of built environment on bike-sharing usage under the pandemic. Results show that the average weekly ridership declined by 52.04%. To account for the spatially heterogeneous relationship between the built environment and the ridership, the geographically weighted regression (GWR) model and the semiparametric GWR (S-GWR) model are constructed. We find that the S-GWR model outperforms the GWR and the multiple linear regression models. The results of the S-GWR model indicate that education employment density, distance to subway, COVID-19 cases, and ridership before COVID-19 are global variables. The effects between ridership and the built environment factros (i.e., household density, office employment density, and the ridership) vary across space. The results of this study could provide a useful reference to transportation planners and bike-sharing operators to determine the high bike-sharing demand area under the pandemic,thus adjusting station locations, capacity, and rebalancing schemes accordingly.

1. Introduction

The outbreak of COVID-19 has seriously threatened the lives of people around the world. According to Johns Hopkins University in the United States, the cumulative number of deaths due to COVID-19 in the United States has exceeded 650,000, and the cumulative number of confirmed cases has exceeded 40.4 million as of September 8, 2021. During this period, US government agencies have implemented policies to reduce the community spread of the virus, including mandate stay-at-home and social distancing orders [1]. These orders have largely impacted residents' daily travel behaviors and further affected the urban transportation systems [2]. Given that the pandemic may last for a long time, the impactsare expectedto continue.

Consideringthe risk of exposure to COVID-19, people tend to reduce their use of public transportation modes (i.e.,subways and buses)following soical distancing guidances. However, the use of bike sharing has not been severely impacted because users could ride bikes in the open space and keep safe social distances. Studies have shown that when public transportation systemsare considered dangerous during COVID-19 [1], residents usually switch from a high-risk mode to cycling to reduce the risk of infection [3]. As a result, the demand for the use of bike-sharing has changed dramatically compared with the period before the outbreak of COVID-19 [4]. Therefore, understanding how the built

environment factors affect the usage of public bicycles under the influence of COVID-19 is necessary because it could provide an important reference to transportation planners and bike-sharing operators to determine the high-demand areas, thusmaking an adjustment to the locations,the capacity, as well as the reblancing schemes of bike-sharing stations.

There has been a rich body of work on the use of public bicycles before the pandemic. Two types of models have been widely used in previous studies, namely global regression models [5, 6] and local regression models [7]. Global models, such as linear regression models and negative binomial regression models, assume that the coefficients of all predictors do not change across space. Despite their wide use, they do not capture the spatial variation in the relationship between predictor and response variables, especially in the case of large study areas [8, 9]. Therefore, the spatial latent class model [10] and the geographically weighted regression (GWR) models [7] are some of the methods adopted to capture this spatial variation. In this study, the GWR model is used. Since GWR models assume that all variables have a spatially varying relationship with the response variable, which may not be true, semiparametric GWR (S-GWR) models have been developed that allow some variables to be global and others to be local.

As a result, this study investigates the spatially varying relationship between the built environment and the bike-sharing ridership during COVID-19 while controlling for the ridership before the pandemic. We intend to answer the following four questions.

(1) Does the bike-sharing ridership increase or decrease due to the outbreak of COVID-19?

(2) If bike-sharing ridership changes, will the change of ridership of each station in proportion to the total ridership change?

(3) If the change of ridership of each station is not in proportion to the total ridership change, what factors, including built environment and demographics, result in this difference?

(4) How do these factors contribute to this difference?

The rest of the paper is structured as follows. The second part is a summary of relevant studies on the usage of bike-sharing. The third part describes the data used in this study. The fourth part presents the model results. The fifth part concludes this study by summarizing the main findings and the limitations.

2. Literature Review

The literature review of related studies is composed of two parts: the influencing factors of public bicycle ridership before the outbreak of COVID-19 and the impact of COVID-19 on bike-sharing usage.

2.1. Factors Influencing Bike-sharing before the Outbreak of COVID-19. Scholars have used different data and models to explore the factors that significantly influence bike-sharing usage. The factors can be divided into two categories: external and internal.

External factors mainly refer to built environment factors, including density, diversity, and design [7, 11–14]. Additionally, demographic factors are also regarded as external factors, including age, private car ownership, and income[15–18]. Other external factors are also included, such as weather conditions (e.g, temperature, humidity, and wind speed), substitution mode, and holidays [6, 17, 19–24]. In addition to this, in the context of the epidemic, some studies have also considered COVID-19-related factors, such as the number of cases and the number of deaths directly related to COVID-19 [25–27]. Internal factors mainly refer to personal preferences and service levels. For example, some studies explored how users' intention to use and fares affect ridership [5, 28–30].

In terms of model selections, many studies used multiple linear regression (MLR) models to determine significant influencing variables [6, 12, 31]. Since ordinary least square models cannot account for multicollinearity,capture spatial autocorrelations, and accurately estimate regression coefficients, Hu and Chen used partial least square to deal with the multicollinearity between explanatory variables. They found that the influence of the independent variables like household income on ridership at most stations was spatially different [25]. To further investigate the spatialimpacts, Cox and Hurtubia used spatial regression models to count for the spatial autocorrelation [10, 32–34]. And one of the studies concluded that the usage of dockless bike stations was spatially autocorrelated in commercial areas and road intersections [31]. Singhvi et al. used generalized linear regression models to deal with skewed distribution of the response variable [35–37]. The positive effect of station-CBD distance and the number of entertainment venues on the number of bike-sharing trips was found [33]. Hu et al. adopted a different model, the generalized mixed-effects model, by adding random effects to the generalized linear regression model [26, 38, 39]. Researchers found that areas with more COVID-19 cases, high income, and more educational employment had less human mobility under the impact of COVID-19 [26].

The predicting variables, response variables, and models of the most relevant studies are summarized in Table 1.

2.2. Impact of COVID-19 on Bike-sharing Ridership. The studies related to the impact of COVID-19 on the ridership of bike-sharing are summarized in this section. Bucsky studied the changes in human mobility and travel mode shares in Hungary during COVID-19 [4]. Public bicycles had the smallest decrease in ridership. The studypointed to public bicycles as an alternative to public transportation under COVID-19, giving a stronger rationale for the government to promote cycling. Buehler and Pucher studied the impact of COVID-19 on public bicycle usage through national surveys [41]. They revealed the general trends and changes over time in bicycling in different cities in Europe and the United States from 2019 to 2020. In order to explore the mechanisms of changes in public bicycle use, several

Table 1: Summary of studies on the relationship between the built environment and bike-sharing ridership.

Author	Predicting variable															Response variable	Model
	Density		Diversity		Destination accessibility		Distance to public transit				Design						
	Population	Employment	Mix land use	Percent of different land use types	Commuting distance	Distance to the city center	Number of bus stops	Number of subway stations	Distance to the nearest bus station	Distance to the nearest subway station	Bike-station characteristics	Bike lane characteristics	Road network density	Primary and secondary road density	Intersection density		
Yang et al. [7]	✓	✓			✓	✓					✓		✓	✓		Hourly ridership	MLR; GWR; S-GWR
El-Assi et al. [20]	✓	✓		✓							✓	✓	✓			Hourly ridership	Linear mixed model
Lin et al. [40]																Annual ridership	MLR
Noland et al. [28]	✓	✓	✓	✓							✓	✓				Monthly ridership	Negative binomial regression
Wang and Chen [33]	✓	✓		✓			✓	✓			✓	✓			✓	Monthly ridership	SEM
Hyland et al. [39]	✓	✓		✓		✓	✓	✓		✓	✓	✓				Monthly ridership	Mixed multilayer linear model

scholars have started to study the influencing factors using different methodologies.

Hu and Chen used Bayesian structural time series models and partial least square regression to study the temporal evolution of the impact of COVID-19 on transit ridership in terms of land use, COVID-19 related features, and sociodemographic variables [25]. The number of COVID-19 cases/deaths in the study was positively associated with a decline in ridership of public transportation, opposed toeducational level and income. Hu et al. used the generalized additive mixed model to explore the relationship between trips and influencing factors, including COVID-19-related features, demographic, and employment [26]. The results showed that the number of COVID-19 cases, income level, and educational employment were negatively associated with trips. The nonlinear temporal interactions between various independent variables and bike-sharing usage change were also explored by the same model [27]. The paper illustrated that residential is positively correlated with bike-sharing usage, while car ownership is negatively correlated with it.

In summary, the existing studies addressed the evolution of the effects in the time dimension and handled mixed effects of linearity and nonlinearity. However, the spatial nonstationary relationship has not been considered, which may lead to estimation bias. Therefore, this paper uses GWR and S-GWR models to deal with this issue.

3. Data Description

Chicago is one of the most populous cities in North America, with a large and energetic downtown, which attracts many commuters and visitors. This study uses two data sets, namely, the trip data of the Chicago Divvy bike-sharing system and the Smart Location Database (SLD) developed by the U.S. Environmental Protection Agency. The bike-sharing system of Chicago covers the unrban area of Chicago and two neighboring suburbs with around 600 stations and over 6,000 bikes. The data of the Divvy bike-sharing system include the start/end time of each trip, the start/end station of each trip, the type of membership, and user information. The spatial distribution of Divvy bicycle stations in Chicago is shown in Figure 1. SLD provides the built environment and demographic information aggregated at the level of census block group (CBG).

3.1. Changes in the Spatial Distribution of Usage during COVID-19. Considering that the spread of COVID-19 began on 2020/2/26 in the United States, the bike-sharing trip data of the eight weeks before and after the spread week (2/27/2020-3/4/2020) are used in this study to represent the pre-COVID-19 period and peri-COVID-19 period.

The weekly usage of the Divvy system in the eight weeks before and during COVID-19 is plotted in Figure 2. The usage is defined as the sum of the pickup and dropoff trips.

As shown in Figure 2, the usage had fluctuated around 30,000 before COVID-19. During COVID-19, a short rise was observed, followed by a quick decline from nearly 50,000

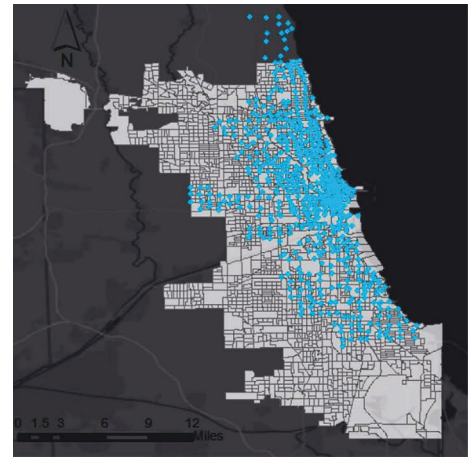


Figure 1: Divvy bicycle stations in the city of Chicago.

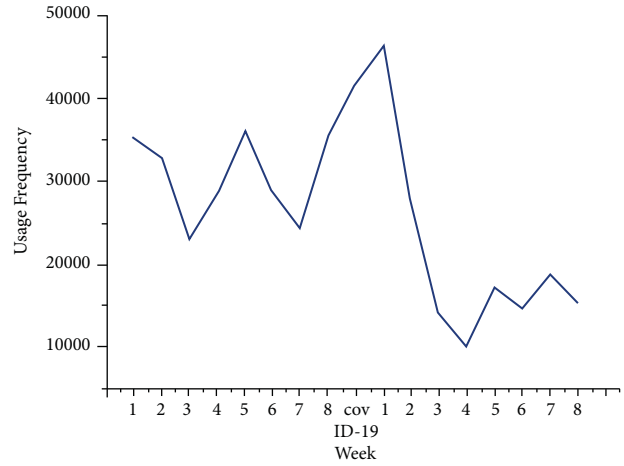


Figure 2: Weekly usage of bike sharing during the period of eight weeks before and after the outbreak of COVID-19.

to 10,000. After that, the usage became stable, being around 18,000. Therefore, the research period after the epidemic in this study is from the next three weeks to the next six weeks (3–6 weeks to the right of the spread week in Figure 2), and the research period before the epidemic is selected symmetrically from the first six weeks to the first three weeks as the research interval (3–6 weeks to the left of the spread week in Figure 2). According to statistics, the average weekly usage before the epidemic was 29,131.5 times, and the average weekly usage during the pandemic was 13,970.25 times. Overall, the usage of bike-sharingduring COVID-19 decreases by 52.04%.

If the ratio of bike-sharing usage during COVID-19 to that before COVID-19 is roughly the same for each station, the change in usage is only caused by the epidemic and has nothing to do with other factors. If this ratio differs greatly from station to station, it shows that the change is not only affected by COVID-19 but also affected by the characteristics of the station and surrounding environment such as built environment and demographics. The histogram of this ratio is shown in Figure 3.

Figure 3 shows that the ridership of most stations decreased during COVID-19 while that of some stations

increased and the ratio varies greatly. Figure 4 presents the spatial distribution of the ratio. The two figures indicate that the change of usage is affected not only by the outbreak of COVID-19 but also by other factors such as the built environment.

Figures 5 and 6 present the spatial distributions of ridership. It shows that the stations with high ridership were centered around the CBD before COVID-19 and centered around the stations in the north during COVID-19. From these two figures, we can infer that bike-sharing operators should pay more attention to the possible shortage of bicycles or docks of the stations in the north during COVID-19.

3.2. Response and Predicting Variables. The response variable in this study is the total usage (sum of pickup and drop-off trips) of each station for the four weeks after COVID-19 (3/19/2020–4/15/2020), as shown in Figure 1.

In order to study the factors that affect the change of usage during COVID-19, it is necessary to control the usage before COVID-19. As a result, it is included in the model as a control variable. When selecting built environment factors, this article refers to the "5D" variables (5 types of built environment variables whose names start with D, including density, design, diversity, distance, and destination accessibility) proposed by previous studies as potential variables that may have an impact on the bike-sharing ridership under COVID-19[17]. In the end, a total of 20 predictors, including built environment, demographics, COVID-19-related cases and deaths, and ridership before COVID-19, are selected.

Seventeen variables in the built environment and demographics are derived from SLD. A circular buffer with a radius of 300 meters is drawn around each station. The radius of 300 meters is determined based on the common walking distance between the origin or the destination and the public bicycle station [7]. Values of the predicting variables are extracted based on the buffer.

The number of COVID-19 cases and deaths are obtained from the City of Chicago. Considering that the COVID-19 cases and deaths may have a wider influencing area [27], a circular buffer zone with a radius of 500 meters is drawn around each station to extract the number of COVID-19 cases and deaths.

3.3. Data Processing. The normal distribution of the dependent variable is an assumption of the classical linear regression model that ensures that the parameter regression results are unbiased [42]. The histogram of the response variable follows a skewed distribution, different from the normal distribution. In response to this problem, a logarithmic transformation of the response variable is performed. The transformation has also been adopted by other studies. The transformed results are also shown in Figure 7.

Previous studies have found that demographical and built environment variables would affect the travel patterns of residents during COVID-19 [25, 26]. For example, many people may work remotely or study at home due to the mandated work-from-home order, following social distance guidelines. Most people's home-based trips have a lower decline than office-based trips. Therefore, controlling for pre-epidemic ridership, household density is assumed to be positively correlated with peri-epidemicusage, while employment density is assumed to be negatively correlated with post-pandemic usage. In addition, variables such as distance to public transportation and proximity to the city center are considered to influence bike sharing usage. So both variables are negatively correlated with peri-pandemic ridership. Thus, this paper includes built environment, socioeconomic, and COVID-19-related variables that may impact peri-epidemic use as the response variable.

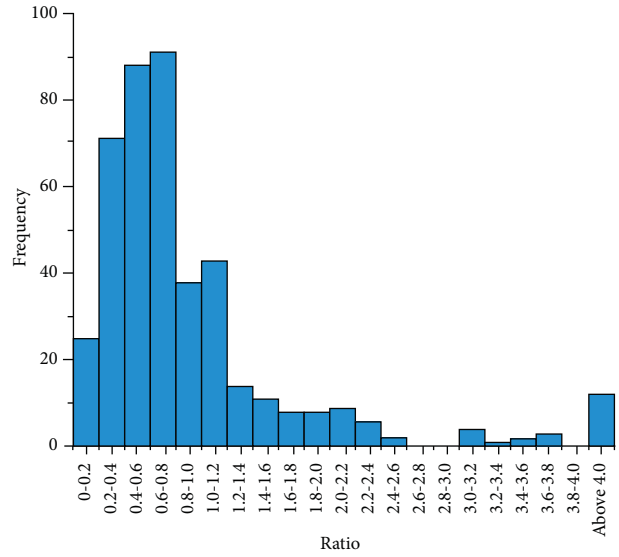


Figure 3: Histogram of the ratio of the ridership during COVID-19 to the ridership before COVID-19.

In order to eliminate the large difference in the magnitude of the explanatory variables and facilitate result interpretations, the explanatory variables are also logarithmically transformed. The modeling result will represent the elasticity of the response variable to the explanatory variables, which is expressed as the percentage of change in the response variable caused by a 1% change in the explanatory variable. The descriptive statistics of all variables in this study are shown in Table 2.

The formula for employment entropy is as follows [43]:

$$\text{Employment Entropy} = -\frac{\sum_{i=1}^{N}(p_i)\ln(p_i)}{\ln(N)}, \tag{1}$$

where N represents the number of employment types and p_i is the proportion of employment type i.

4. Methods

4.1. Multiple Linear Regression (MLR). This study establishes the MLR model to analyze factors that influence bike-sharing usage during COVID-19. The model assumes that the relationship between the predictor variables and the response variable is linear and homogeneous across space. Its function is as shown in the following equation.

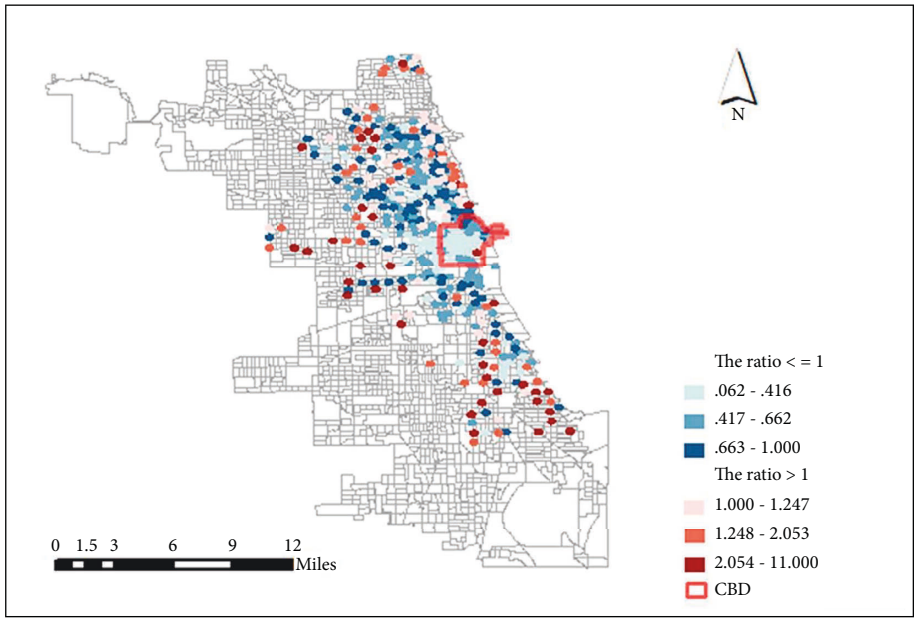


Figure 4: Spatial distribution of the ratio of ridership during COVID-19 to that before COVID-19.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon, \quad (2)$$

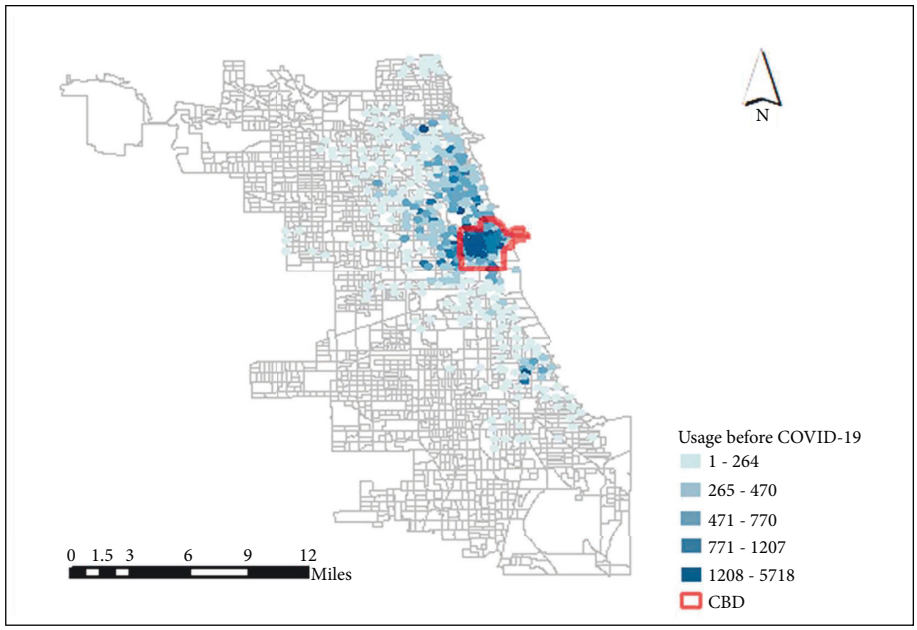


Figure 5: Spatial distribution of Divvy usage before COVID-19.

where y is the response variable; x_1, x_2,..., x_k are the predictors; $\beta_0, \beta_1, \ldots, \beta_k$ are the coefficients of the predictors; and ε is the random error, which has an expected value of zero, follows the normal distribution, and is independent of each other [42].

In the MLR model, the parameters are mainly estimated using Ordinary Least Squares (OLS) methods. The objective function is as follows:

$$\min \sum_{i=1}^{n} \left[Y_i - \left(\widehat{\beta}_0 + \widehat{\beta}_1 X_{i1} + \widehat{\beta}_2 X_{i2} \ldots + \widehat{\beta}_k X_{ik} \right) \right]^2, \quad (3)$$

where Y_i is the true value of the i-th response variable; $X_{i1}, \ldots, X_{ik}$ are the k-th predictor of the i-th response variable's predictors; and $\widehat{\beta}_0, \widehat{\beta}_1, \ldots, \widehat{\beta}_k$ are the estimates of the parameters.

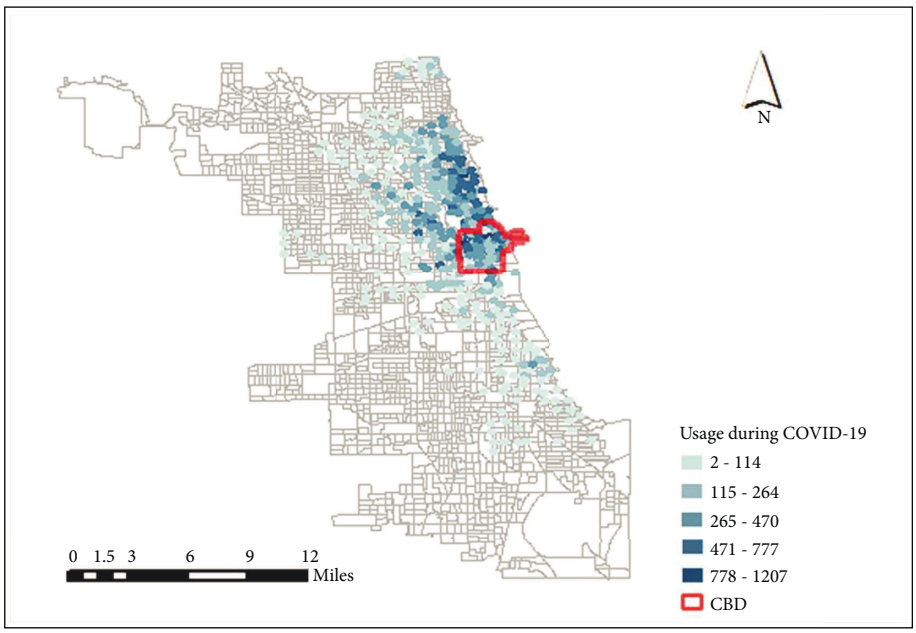


Figure 6: Spatial distribution of Divvy usage during COVID-19.

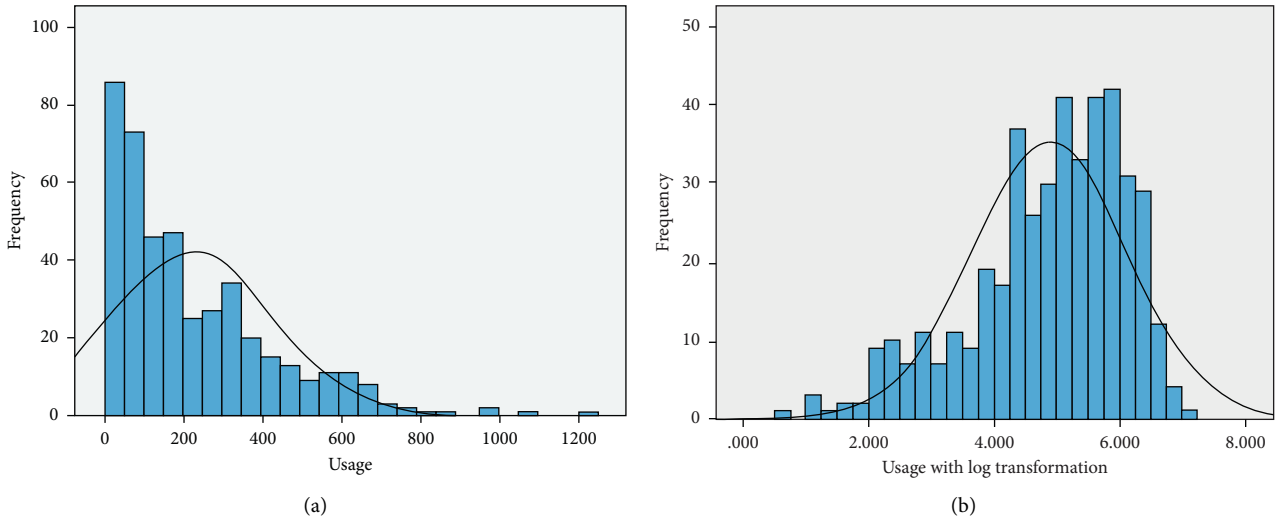


Figure 7: Histogram of the bike-sharing usage with and without log transformation: (a) histogram of the bike-sharing usage and (b) histogram of the bike-sharing usage with log transformation.

4.2. Geographically Weighted Regression (GWR). In the case of a large study area, the relationship between the response variable and explanatory variables may vary across space. So the study needs to use a local regression model such as GWR. The GWR model improves the traditional MLR model by allowing the relationship to vary across space. It establishes local regression equations at each station and thus allows the regression parameters to vary with spatial location. Its function is as follows [7]:

$$y_i = \beta_{i0}(u_i, v_i) + \sum_{k=1}^{n} \beta_{ik}(u_i, v_i) x_{ik} + \varepsilon_i, \quad (4)$$

where y_i represents the bike-sharing usage of station i, β_{ik} is the coefficient of the predictor k of station i, x_{ik} is the predictor k of station i, ε_i is the random error term of station i, and (u_i, v_i) represents the latitude and longitude of station i.

There are a set of coefficients at each public bicycle station for the GWR modeling results. It indicates that the effect of the predicting variables on the response variable varies across space. When estimating the coefficients of each station, weight w_i is assigned based on the distance from other stations to the target station.The coefficients of explanatory variablesare estimated by minimizing the

Table 2: Descriptive statistics of variables.

Category	Variable	Meaning	Source	Mean (log-transformed)	Variance (log-transformed)
Usage	Usage after COVID-19	Usage of each station from 2020/3/19 to 2020/4/15	Divvy's official website	4.872	1.524
Density	Population density	Population density (people/acre)	Smart Location Database	3.336	0.510
	Household density	Household density (households/acre)	Smart Location Database	2.700	0.711
	Entertainment employment density	Entertainment employment density (jobs/acre)	Smart Location Database	0.202	4.540
	Education employment density	Education employment density (jobs/acre)	Smart Location Database	−1.135	8.962
	Retail employment density	Retail employment density (jobs/acre)	Smart Location Database	−0.051	2.962
	Office employment density	Office employment density (jobs/acre)	Smart Location Database	0.396	4.699
	Healthcare employment density	Healthcare employment density (jobs/acre)	Smart Location Database	0.228	3.616
Diversity	Employment entropy	Employment entropy using the formula for eight types of jobs such as retail jobs, factory jobs, and service jobs	Smart Location Database	−0.625	0.342
Design	Auto-oriented links	Length of links only for automobiles per square mile	Smart Location Database	1.316	0.721
	Multimodal links	Length of links for autos and pedestrians per square	Smart Location Database	−3.462	20.677
	Pedestrians-oriented links	Length of links only for pedestrians per square	Smart Location Database	3.027	0.081
Distance from public transit	Distance to subway	Distance from population-weighted centroid to the nearest subway stop	Smart Location Database	5.112	0.432
Destination accessibility	Job accessibility	Jobs within 45 minutes auto travel time	Smart Location Database	12.969	0.189
	Working-age population accessibility	Working age population within 45 minutes auto travel time	Smart Location Database	13.228	0.120
	Distance to CBD	Distance to CBD	Own calculation	1.533	0.825
Demographic	Percentage of HH with no vehicles	Percentage of zero-car households in CBG	Smart Location Database	−1.367	0.377
	Percentage of low-income population	Percentage of workers earning $1250/month or less	Smart Location Database	−0.995	0.898
COVID-19	COVID-19 cases	Cumulative COVID-19 cases from 2020/3/19 to 2020/4/15	Chicago data portal	6.690	0.190
	COVID-19 deaths	Cumulative deaths caused by COVID-19 from 2020/3/19 to 2020/4/15	Chicago data portal	2.419	0.584
Control	Usage before COVID-19	Divvy usage from 2020/1/16 to 2020/2/12	Divvy's official website	5.254	2.751

weighted sum of squares. The objective function for the GWR model is as follows:

$$\min \sum_{j=1}^{n} w_{ij} \left(y_i - \beta_{i0} - \sum_{k=1}^{n} \beta_{ik}(u_i, v_i) x_{ik} \right)^2. \quad (5)$$

The spatial weight reflects the importance of the position. Many ways can be used to calculate the spatial weight. The simplest one is the distance threshold function. The specific function is as follows:

$$w_{ij} = \begin{cases} 1, & d_{ij} \le D, \\ 0, & d_{ij} > D, \end{cases} \quad (6)$$

where D represents the distance threshold and d_{ij} represents the distance between station i and the target station j. To solve the problem of weight discontinuity, the Gaussian function is also often used to express the relationship between weight and distance

$$w_{ij} = e^{-1/2\left(d_{ij}/b\right)^2}, \quad (7)$$

where w_{ij} represents the weight between stations i and the target station j, d_{ij} represents the distance between stations i and the target station j, and b is the bandwidth.

4.3. Semiparametric Geographically Weighted Regression (S-GWR). GWR models assume that the coefficients of all predictors vary across space. However, the relationship between some predictors and the response variable may not vary across space. The S-GWR model, as an extension of the GWR model, allows some predictors to be global and others to be local. The expression of the S-GWR model is as follows [7], and the symbols in the equation are the same as the GWR model.

$$y_i = \beta_{i0}(u_i, v_i) + \sum_{k=1}^{n} \beta_{ik} x_{ik} + \sum_{k=1}^{n} \beta_{ik}(u_i, v_i) x_{ik} + \varepsilon_i. \quad (8)$$

The objective function for the S-GWR model is as follows:

$$\min \sum_{j=1}^{n} w_{ij} \left(y_i - \beta_{i0} - \sum_{k=1}^{p} \beta_{ik} x_{ik} - \sum_{k=1}^{n} \beta_{ik}(u_i, v_i) x_{ik} \right)^2. \quad (9)$$

5. Model Results

This section establishes MLR, GWR, and S-GWR models to explore the relationships between explanatory variables and the response variable.

5.1. Results of MLR Model. There are two ways to construct regression models. One way is to treat the ridership during COVID-19 as the dependent variable and the ridership before COVID-19 as well as other variables as the independent variables. The other way is to treat the ratio of the ridership during COVID-19 to the ridership before COVID-19 as the dependent variable and other variables as the independent variables. Both ways are explored in this study.

The first way to construct the model, which is to treat the ridership during COVID-19 as the dependent variable and the ridership before COVID-19 as well as other variables as the independent variables, is first explored. The backward variable selection method is adopted to select variables. After this method, six explanatory variables are significantly related to the usage of Divvy bike-sharingduring COVID-19 at the5% level. Moreover, the variance inflation factor VIF is less than 5 for these six variables, which indicates there are no multicollinearity issues. The formula for VIF is as follows:

$$\mathrm{VIF} = \frac{1}{\left(1 - R_i^2\right)}, \quad (10)$$

where R_i^2 is the coefficient of determination for the mode using the i-th explanatory variable as the response variable and the rest of the explanatory variables as explanatory variables.

Table 3 presents the model results. The significant variables are household density, education employment density, office employment density, distance to the nearest subway station, COVID-19 cumulative cases, and usage before COVID-19.

Explanations for the relationships between significant predictors and response variables are given below.

During COVID-19, the household density is positively correlated with ridership. This may be due to work-at-home policies that keep residents at home. As a result, more trips are home-based trips. There is also a positive association between the cumulative number of cases and the ridership. This result may be because areas with more bike-sharing trips are those with high travel demand. More people gathering around the area leads to a higher risk of infection. On the other hand, the increase in the number of COVID-19 infections may also make people use bike-sharing to replace other public transportation modes, such as the subway and bus. Moreover, the usage of Divvy bike-sharing before COVID-19 is positively correlated with the usage of Divvy bike-sharing during COVID-19. This shows that stations with a high usage rate before COVID-19 continue to have high usage during COVID-19. The coefficient is 0.626, indicating that if the pre-COVID-19 usage increases by 1%, the peri-COVID-19 usage will increase by 0.626% on average.

The education employment density is negatively correlated. It could be due to that most schools require students and faculty members to stay at home and to take or teach classes online. Similarly, the office employment density is also negatively correlated. This result may also be due to the stay-at-home order and work-from-home policy in Chicago in response to COVID-19. And the number of ridership is negatively correlated with the distance to the nearest subway station. The possible reason is that residents who are close to the nearest subway station after the COVID-19 outbreak have switched from using the subway to using bike-sharing.

However, bike-sharing usage before COVID-19 may weaken the effect of time invariant explanatory variables on bike-sharing usage during COVID-19. To demonstrate the explanatory power of these predicting variables for bike-

Table 3: MLR model results.

Variable	Coefficient	Standard error	P	VIF
Household density	0.288	0.041	0.000	1.953
Education employment density	−0.028	0.011	0.013	1.865
Office employment density	−0.043	0.021	0.047	3.629
Distance to subway	−0.165	0.047	0.000	1.580
COVID-19 cases	0.224	0.069	0.001	1.522
Ridership before COVID-19	0.626	0.022	0.000	2.133

Table 4: MLR model results using the ratio of bike-sharing usage during/before COVID-19 as the response variable.

Variable	Coefficient	Standard error	P	VIF
Household density	0.133	0.045	0.004	1.324
Entertainment employment density	−0.126	0.024	0.000	2.632
Education employment density	−0.072	0.014	0.000	1.930
Working age population accessibility	−0.663	0.221	0.003	1.349
Percentage of HH with no vehicles	0.165	0.066	0.012	1.245

sharing usage during COVID-19, the ratio of bike-sharing usage during COVID-19 to bike-sharing usage before COVID-19 is also developed as the response variable. Considering there are huge differences in the ratio, the log transformation is performed on the dependent variable. The same backward variable selection method is used to screen for significant variables. The results of the model are shown in Table 4.

The model results, shown in Table 4, are different from those of the previous model. The goodness of fit of the model is 0.347, which is not as good as that of the previous model. This is probably because the variable of ridership before COVID-19 accounts for a large proportion of the time invariant components of the response variable, ridership during COVID-19, in the previous model. Household density and education employment density are significant at both mdels. Though there are some differences in the model results in terms of significant variables, it should be noted that the difference in model results is quite common in empirical studies. Sometimes, simply extending the time period of the data could lead to different model results.

Specifically, the model results show a positive correlation between household density and the dependent variable. The work-from-home policy made many people stay at home where office-based trips decreased more dramatically than home-based trips. The proportion of households without a car is also positively correlated with the dependent variable. During the epidemic, the automobile would be travelers' first preference. For households without an automobile, people are less willing to use the metro or bus, which could lead to disease transmission and thus turn to bike-sharing. As a result, bike-sharing riderships in areas with higherhousehold density and lower vehicle ownershipis not affected by the epidemic too much.

In contrast, entertainment employment density and education employment density are negatively correlated with the dependent variable. This is probably because the epidemic severely impacted the entertainment business, leading to employment loss. Regarding the effect of education employment density, both employees and students are instructed to give lectures or study at home, which reduces trips to and from schools. The working-age population accessibility is the working-age population that can be accessed by driving an automobile in 45 minutes. It is also negatively related to the dependent variable. It may be because, by dividing the population into a working population and a nonworking population, the trips performed by the working population have a higher percentage of decrease compared to those made by the nonworking population due to the work-from-home policy. Thus, by controlling for the household density, the working-age population accessibility is negatively related to the dependent variable.

Each of the two models has its own advantage. In this paper, we adopt the first model to explore the spatially varying relationship between explanatory variables and the dependent variable.

5.2. GWR Model Results. The modeling results of the GWR model are shown in Table 5.

The overall R-squared of the GWR model is 0.865. The R-squared of each station is visualized in Figure 8.

The R-squared for each station is above 0.5, with a considerable proportion being above 0.8, which indicates that the GWR model fits the data well. The generally lower goodness-of-fit for downtown areas may be due to the fact that ridership at these stations is more variable than at other stations. In addition, the variables selected do not accurately explain the dynamic changes in ridership because of the work-at-home policy.

5.3. S-GWR Model Results. The S-GWR model is also constructed.Table 6 presents the modeling results of the S-GWR model. To determine whether each predicting variable is global or local, all variables are first assumed to be global variables. The difference (DIFF) of criterion of each variable is calculated. The DIFF indicates the difference in comparison metrics (AICc) between the GWR and S-GWR models. In general, a positive value for the DIFF of criterion

Table 5: Coefficients of each variable in the GWR model.

Variable	Min	Lower quartile	Median	Upper quartile	Max
Intercept	4.650583	4.796953	5.013346	5.089241	5.200212
Household density	−0.086092	0.193541	0.277111	0.305725	0.373394
Education employment density	−0.174001	−0.086815	−0.050423	−0.017703	0.182649
Office employment density	−0.245413	−0.150852	−0.040037	0.046748	0.177302
Distance to subway	−0.286487	−0.115407	−0.097054	−0.072675	0.084666
COVID-19 cumulative cases	−0.136410	0.006409	0.035964	0.059660	0.180760
Ridership before COVID-19	0.818073	0.935148	0.982256	1.030879	1.144220

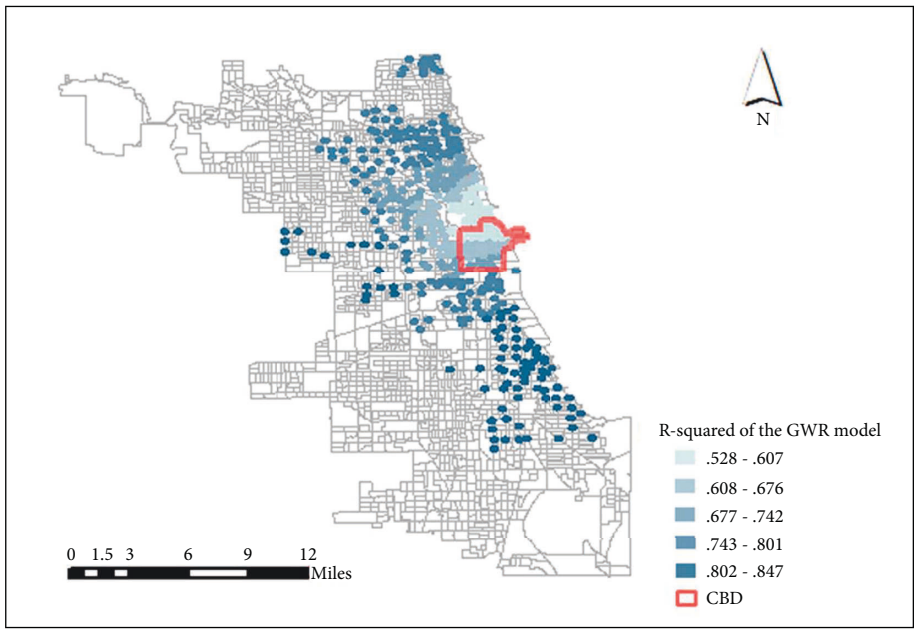


Figure 8: Spatial distribution of the R-squared of the GWR model.

Table 6: Variable type test.

Variable	F	DOF for F	DIFF
Intercept	8.902237	2.268	−15.953818
Household density	5.598979	3.301	−11.813178
Education employment density	1.722706	4.106	1.953052
Office employment density	2.630903	3.698	−1.785874
Distance to subway	2.141936	3.503	0.113963
COVID-19 cases	1.213646	2.911	2.971311
Ridership before COVID-19	1.358139	3.954	3.411965

indicates that there is no spatial variability [7]. In other terms, if DIFF > 0, the predicting variable is regarded as a global variable; otherwise, it is a local variable.

Compared with the MLR modeling results, the intercept for GWR or S-GWR is a local variable that varies with the geographical location, which could be observed in previous studies [7, 41]. As summarized in Table 6, the household density and the office employment density are local variables in the model, while the others are global variables.

The coefficients of the two local variables will be presented below. The interpretations of the results will also be given.

The spatial distribution of the coefficients of household density is shown in Figure 9. When household density is a significant factor, it usually has a positive correlation with the usage of public bicycles during the epidemic, which is consistent with the conclusions of the previous study [27]. The coefficients of household density in the center and eastern regions were relatively small, ranging from 0.161 to 0.286. This could be attributed to many high-income residents who are more likely to commute by private cars instead of public bikes in these two regions.

The spatial distribution of the coefficients of office employment density is shown in Figure 10. When office employment density is a significant factor, it is usually negatively correlated with ridership. It is probably because the stay-at-home order during COVID-19 reduced the number of people working in the office and reduced the usage of bike-sharing in areas with high office employment density. Therefore, this negative impact is obvious in the city center. Also, there is an area in the south where office employment density is significant because of the proximity of this area to the location of the University of Chicago. The office employment in this area is mainly the faculty members and staff of the university. During the epidemic, most of the faculty members and staff of the university are required to work from home. The rate of faculty members and staff who work from home is higher than that of other types of jobs. Thus, the office employment density is negative at a

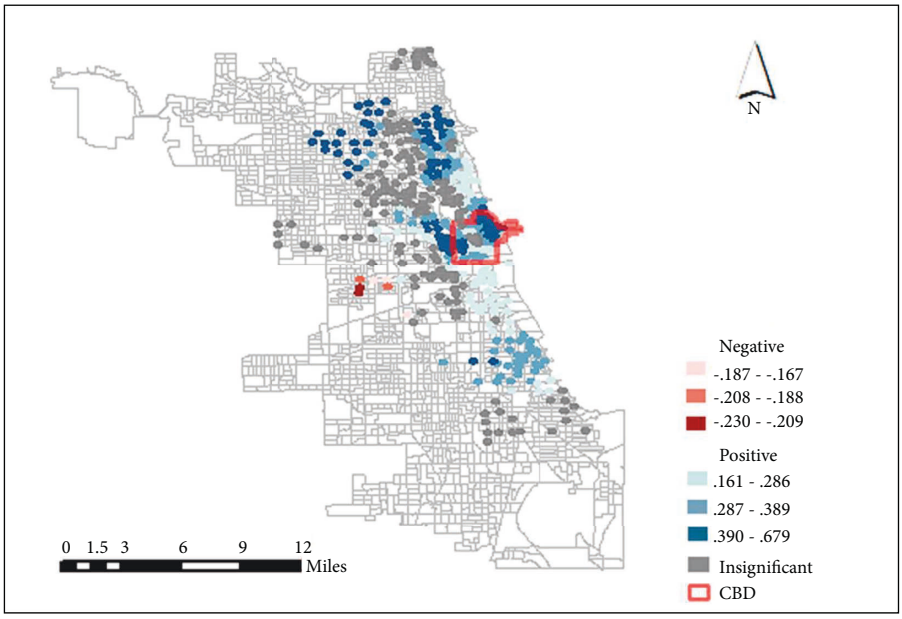


Figure 9: Distribution of household density regression coefficients.

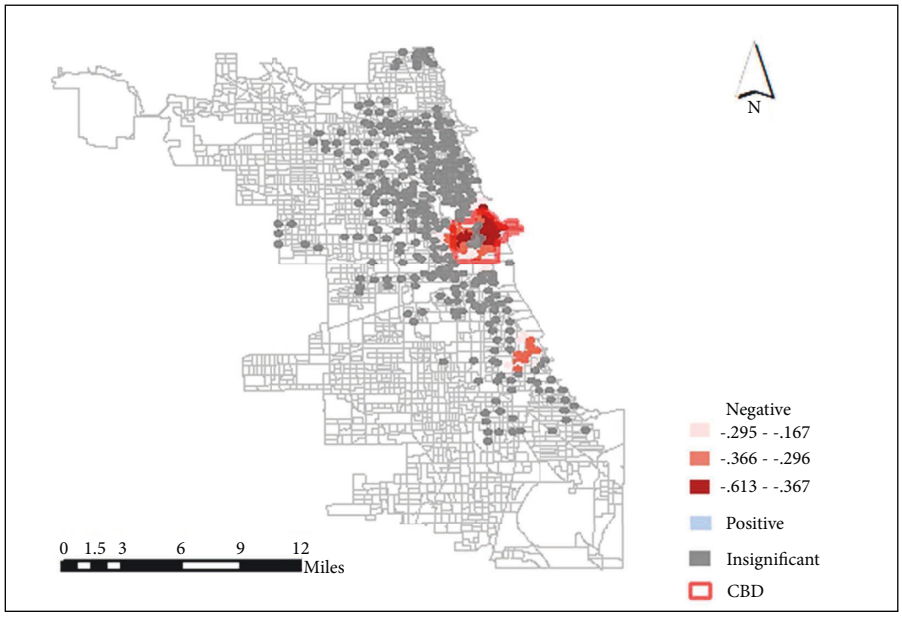


Figure 10: Distribution of office employment density regression coefficients.

significant level. Overall, the magnitude of the negative coefficients of household density and office employment density is larger in downtown areas than in the southern regions. Such relationships may be due to the higher percentage of downtown shutdowns than on the south side.

The overall R-squared of the S-GWR model is 0.886, which is higher than that of the GWR model. The spatial distribution of the R-squared of the S-GWR model is shown in Figure 11. The goodness-of-fit of the stations in the city center is still not high, which shows that the dramatic change in ridership in these areas caused by COVID-19 is difficult to capture.

5.4. Model Comparison. The results of the MLR, GWR, and S-GWR models are compared and shown in Table 7.

Since the models with a smaller sum of squares of residuals, −2 log-likelihood, AIC, AICc, and higher R-squared are regarded as better models, the S-GWR model is the best among the three models.

6. Conclusion

This study investigates the built environment factors that influence the bike-sharing ridership of the Chicago Divvy system during COVID-19 while controlling for the ridership

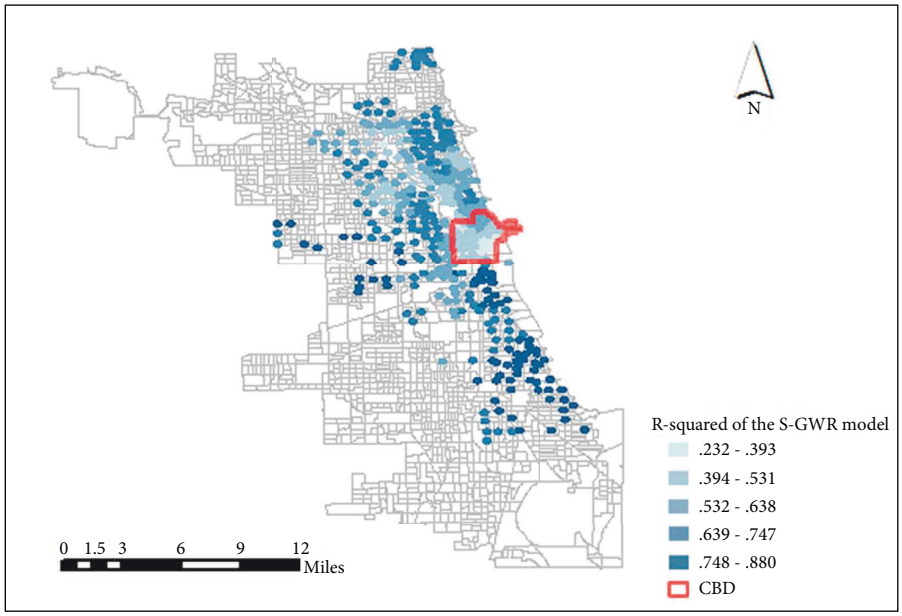


Figure 11: Spatial distribution of the R-squared of the S-GWR model.

Table 7: Comparison of model results.

Indicator	MLR	GWR	S-GWR
Residual sum of squares	112.160	89.645	75.429
−2 log-likelihood	645.349	547.657	472.374
Classic AIC	661.349	612.292	591.082
AICc	661.687	617.640	610.155
R-squared	0.831	0.865	0.886
Adjusted R-squared	0.828	0.851	0.862

before COVID-19. To capture the spatially varying relationship between built environment and ridership, GWR and S-GWR models are established.The MLR model is also developed to be comparable. We found that S-GWR has the highest goodness-of-fit from many perspectives.

We also found that the total bike-sharing ridership declined by half after the outbreak of COVID-19. The decline of the ridership of each station is different; the spatial distribution of usage of bike-sharing during COVID-19 is different from that before COVID-19. This observation indicates that transportation planners and bike-sharing operators should pay attention to this change and could adjust the capacity and location of the stations as well as the rebalancing scheme according to the current ridership pattern.

In terms of the relationship between the built environment and change in ridership, some variables are local variables (i.e., household density and office employment density), whileother variables are global variables such as education employment density and distance to the nearest subway station. The complex relationship should be fully considered when estimating the change in ridership of bike-sharing stations.

There are also some limitations in this study. First, although the results obtained from this study may not be applied to all cities, the analysis framework could be applied to other cities. Each city should develop policies based on its own condition. Secondly, because we used cross-sectional data, the revealed relationship between the independent variables and the response variable should be regarded as a correlation instead of a causal relationship. Although some causal relationships could be inferred from the results, this inference should be made with caution. In the future, panel data could be used to deal with this issue. Of course, traditional Poisson and negative binomial models are designed for count variables. However, it would be more appropriate to use linear regression models when the values of the response variable do not contain zero or small values [44, 45]. Moreover, when the area of the intersection of the buffer and the CBG is not the whole block, we assume that the independent variables are uniformly distributed in the CBG. However, the ground truth may not be the case. Another limitation of this study is that there are different ways to construct the regression models, and no theoretical justification for which model is more suitable. Each of the two ways has its own advantage. The model using the ridership during COVID-19 as the dependent variable is adopted because the results are more intuitive, and the goodness of fit is better. But it should be noted that the high goodness of fit is probably because the ridership before COVID-19 is highly correlated with the dependent variable, and this high correlation could overshadow the effect of other independent variables. As such, modeling other related variables, such as the ratio of peri-COVID-19 and pre-COVID-19 usage, is also meaningful because it could avoid this issue and is worth investigating. Finally, we use four-week bike-sharing ridership data during COVID-19 as the response variable. Though the travel volume is relatively stable during COVID-19 (as indicated in Figure 2)makes it possible to capture bike-sharing usage patterns under COVID-19, using a longer time horizon may generate more reliable modeling results.

Data Availability

The data used in this paper are available at https://www.divvybikes.com/system-data, https://www.epa.gov/smartgrowth/smart-location-mapping, https://data.cityofchicago.org/Health-Human-Services/COVID-19-Cases-Tests-and-Deaths-by-ZIP-Code/yhhz-zm2v, and https://data.cityofchicago.org/browse?tags=gis.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The authors confirm their contribution to the paper as follows. Hongtai Yang contributed to conceptualization, methodology, formal analysis, and writing, particularly review and editing. Zishuo Guo contributed to data processing, formal analysis, and writing, particularly original draft preparation. Guocong Zhai contributed to writing, review, and editing. Linchuan Yang contributed to writing, review, and editing. Jinghai Huo contributed to review and editing. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments

This study was funded by the National Natural Science Foundation of China (Grant nos. 51774241 and 71704145), China Postdoctoral Science Foundation, and Sichuan Youth Science and Technology Innovation Research Team Project (Grant nos. 2019JDTD0002 and 2020JDTD0027).

References

[1] H. Wang and R. B. Noland, "Bikeshare and subway ridership changes during the COVID-19 pandemic in New York City," *Transport Policy*, vol. 106, pp. 262–270, 2021.

[2] L. Yang, Y. Liang, B. He, Y. Lu, and Z. Gou, "COVID-19 effects on property markets: the pandemic decreases the implicit price of metro accessibility," *Tunnelling and Underground Space Technology*, vol. 125, Article ID 104528, 2022.

[3] A. Nikiforiadis, G. Ayfantopoulou, and A. Stamelou, "Assessing the impact of COVID-19 on bike-sharing usage: the case of thessaloniki, Greece," *Sustainability*, vol. 12, no. 19, 2020.

[4] P. Bucsky, "Modal share changes due to COVID-19: the case of Budapest," *Transportation Research Interdisciplinary Perspectives*, vol. 8, Article ID 100141, 2020.

[5] K. Wang, G. Akar, and Y. J. Chen, "Bike sharing differences among millennials, gen xers, and baby boomers: lessons learnt from New York city's bike share," *Transportation Research Part A: Policy and Practice*, vol. 116, pp. 1–14, 2018.

[6] R. A. Rixey, "Station-level forecasting of bikesharing ridership: station network effects in three US systems," *Transportation Research Record*, vol. 2387, no. 1, pp. 46–55, 2013.

[7] H. Yang, Y. Zhang, L. Zhong, X. Zhang, and Z. Ling, "Exploring spatial variation of bike sharing trip production and attraction: a study based on Chicago's Divvy system," *Applied Geography*, vol. 115, Article ID 102130, 2020.

[8] L. Yang, X. Chu, Z. Gou, H. Yang, Y. Lu, and W. Huang, "Accessibility and proximity effects of bus rapid transit on housing prices: h," *Journal of Transport Geography*, vol. 88, Article ID 102850, 2020.

[9] L. Yang, K. W. Chau, W. Y. Szeto, X. Cui, and X. Wang, "Accessibility to transit, by transit, and property prices: spatially varying relationships," *Transportation Research Part D: Transport and Environment*, vol. 85, Article ID 102387, 2020.

[10] T. Cox and R. Hurtubia, "Compact development and preferences for social mixing in location choices: results from revealed preferences in Santiago, Chile," *Journal of Regional Science*, vol. 62, 2021.

[11] Y. Guo, L. Yang, Y. Lu, and R. Zhao, "Dockless bike-sharing as a feeder mode of metro commute? The role of the feeder-related built environment: analytical framework and empirical evidence," *Sustainable Cities and Society*, vol. 65, Article ID 102594, 2021.

[12] A. Faghih-Imani, N. Eluru, A. M. El-Geneidy, M. Rabbat, and U. Haq, "How land-use and urban form impact bicycle flows: evidence from the bicycle-sharing system (BIXI) in Montreal," *Journal of Transport Geography*, vol. 41, pp. 306–314, 2014.

[13] J. Huo, H. Yang, C. Li, R. Zheng, L. Yang, and Y. Wen, "Influence of the built environment on E-scooter sharing ridership: a tale of five cities," *Journal of Transport Geography*, vol. 93, Article ID 103084, 2021.

[14] H. Yang, Y. Bao, J. Huo, S. Hu, L. Yang, and L. Sun, "Impact of road features on shared e-scooter trip volume: a study based on multiple membership multilevel model," *Travel Behaviour and Society*, vol. 28, pp. 204–213, 2022.

[15] A. Faghih-Imani and N. Eluru, "Analysing bicycle-sharing system user destination choice preferences: Chicago's Divvy system," *Journal of Transport Geography*, vol. 44, pp. 53–64, 2015.

[16] S. Kaviti, M. M. Venigalla, and K. Lucas, "Travel behavior and price preferences of bikesharing members and casual users: a Capital Bikeshare perspective," *Travel Behaviour and Society*, vol. 15, pp. 133–145, 2019.

[17] R. Cervero, O. L. Sarmiento, E. Jacoby, L. F. Gomez, and A. Neiman, "Influences of built environments on walking and cycling: lessons from bogotá," *International journal of sustainable transportation*, vol. 3, no. 4, pp. 203–226, 2009.

[18] H. Li, H. Ding, G. Ren, and C. Xu, "Effects of the lcslch," *Transportation Research Part A: Policy and Practice*, vol. 111, pp. 304–315, 2018.

[19] A. Faghih-Imani and N. Eluru, "Determining the role of bicycle sharing system infrastructure installation decision on usage: case study of montreal BIXI system," *Transportation Research Part A: Policy and Practice*, vol. 94, pp. 685–698, 2016.

[20] W. El-Assi, M. Salah Mahmoud, and K. Nurul Habib, "Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto," *Transportation*, vol. 44, no. 3, pp. 589–613, 2017.

[21] H. Yang, J. Huo, Y. Bao, X. Li, L. Yang, and C. R. Cherry, "Impact of e-scooter sharing on bike sharing in Chicago," *Transportation Research Part A: Policy and Practice*, vol. 154, pp. 23–36, 2021.

[22] H. Yang, G. Zhai, L. Yang, and K. Xie, "How does the suspension of ride-sourcing affect the transportation system and environment?" *Transportation Research Part D: Transport and Environment*, vol. 102, Article ID 103131, 2022.

[23] H. Yang, Y. Liang, and L. Yang, "Equitable? Exploring ridesourcing waiting time and its determinants," *Transportation Research Part D: Transport and Environment*, vol. 93, Article ID 102774, 2021.

[24] H. Yang, J. Huo, R. Pan, K. Xie, W. Zhang, and X. Luo, "Exploring built environment factors that influence the market share of ridesourcing service," *Applied Geography*, vol. 142, Article ID 102699, 2022.

[25] S. Hu and P. Chen, "Who left riding transit? Examining socioeconomic disparities in the impact of COVID-19 on ridership," *Transportation Research Part D: Transport and Environment*, vol. 90, Article ID 102654, 2021.

[26] S. Hu, C. Xiong, M. Yang, H. Younes, W. Luo, and L. Zhang, "A big-data driven approach to analyzing and modeling human mobility trend under non-pharmaceutical interventions during COVID-19 pandemic," *Transportation Research Part C: Emerging Technologies*, vol. 124, Article ID 102955, 2021.

[27] S. Hu, C. Xiong, Z. Liu, and L. Zhang, "Examining spatiotemporal changing patterns of bike-sharing usage during COVID-19 pandemic," *Journal of Transport Geography*, vol. 91, no. 11, Article ID 102997, 2021.

[28] R. B. Noland, M. J. Smart, and Z. Guo, "Bikeshare trip generation in New York City," *Transportation Research Part A: Policy and Practice*, vol. 94, pp. 164–181, 2016.

[29] J. Wang and G. Lindsey, "Neighborhood socio-demographic characteristics and bike share member patterns of use," *Journal of Transport Geography*, vol. 79, Article ID 102475, 2019.

[30] M. Verma and A. Awasthi, "Evaluating bikesharing service quality: a case study for BIXI, Montreal," *International Journal of Productivity and Quality Management*, vol. 29, no. 1, pp. 45–61, 2020.

[31] T. D. Tran, N. Ovtracht, and B. F. d'Arcier, "Modeling bike sharing system using built environment factors," *Procedia Cirp*, vol. 30, pp. 293–298, 2015.

[32] Y. Zhang, T. Thomas, and M. Brussel, "Exploring the impact of built environment factors on the use of public bikes at bike stations: case study in Zhongshan, China," *Journal of Transport Geography*, vol. 58, pp. 59–70, 2017.

[33] K. Wang and Y. J. Chen, "Joint analysis of the impacts of built environment on bikeshare station capacity and trip attractions," *Journal of Transport Geography*, vol. 82, Article ID 102603, 2020.

[34] Y. Shen, X. Zhang, and J. Zhao, "Understanding the usage of dockless bike sharing in Singapore," *International Journal of Sustainable Transportation*, vol. 12, no. 9, pp. 686–700, 2018.

[35] D. Singhvi, S. Singhvi, and P. I. Frazier, "Predicting bike usage for New York city's bike sharing system," in *Proceedings of the Workshops at the twenty-ninth AAAI conference on artificial intelligence*, Austin, TX, USA, January 2015.

[36] X. Wang, G. Lindsey, J. E. Schoner, and A. Harrison, "Modeling bike share station activity: effects of nearby businesses and jobs on trips to and from stations," *Journal of Urban Planning and Development*, vol. 142, no. 1, Article ID 04015001, 2016.

[37] R. An, R. Zahnow, D. Pojani, and J. Corcoran, "Weather and cycling in New York: the case of Citibike," *Journal of Transport Geography*, vol. 77, pp. 97–112, 2019.

[38] F. Sun, P. Chen, and J. Jiao, "Promoting public bike-sharing: a lesson from the unsuccessful Pronto system," *Transportation Research Part D: Transport and Environment*, vol. 63, pp. 533–547, 2018.

[39] M. Hyland, Z. Hong, H. K. R. D. Pinto, and Y. Chen, "Hybrid cluster-regression approach to model bikeshare station usage," *Transportation Research Part A: Policy and Practice*, vol. 115, pp. 71–89, 2018.

[40] P. Lin, J. Weng, Q. Liang, D. Alivanistos, and S. Ma, "Impact of weather conditions and built environment on public bikesharing trips in beijing," *Networks and Spatial Economics*, vol. 20, no. 1, pp. 1–17, 2020.

[41] R. Buehler and J. Pucher, "COVID-19 impacts on cycling, 2019-2020," *Transport Reviews*, vol. 41, no. 4, pp. 393–400, 2021.

[42] J. M. Wooldridge, *Introductory Econometrics: A Modern Approach*, Cengage Learning, Massachusetts, MA, USA, 2015.

[43] M. G. Turner, R. H. Gardner, and R. V. O'neill, *Landscape Ecology in Theory and Practice*, Springer, New York, NY, USA, 2001.

[44] G. Zhai, K. Xie, and D. Yang D, "Assessing the safety effectiveness of citywide speed limit reduction: A causal inference approach integrating propensity score matching and spatial difference-in-differences," *Transportation Research Part A: Policy and Practice*, pp. 94–106, 2022.

[45] J. M. Wooldridge, "Distribution-free estimation of some nonlinear panel data models," *Journal of Econometrics*, vol. 90, no. 1, pp. 77–97, 1999.

[46] J. M. C. S. Silva and S. Tenreyro, "The log of gravity," *The Review of Economics and Statistics*, vol. 88, no. 4, pp. 641–658, 2006.

[47] H. Yang, X. Lu, C. Cherry, X. Liu, and Y. Li, "Spatial variations in active mode trip volume at intersections: a local analysis utilizing geographically weighted regression," *Journal of Transport Geography*, vol. 64, pp. 184–194, 2017.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 8686584, 13 pages
https://doi.org/10.1155/2022/8686584

Research Article

Mixed Multinomial Probit Model Accommodating Flexible Covariance Structure and Random Taste Variation: An Application to Commute Mode Choice Behavior

Ke Wang[1,2] **Xin Ye**[3] **and Hongcheng Gan**[1,2]

[1]*Business School, University of Shanghai for Science and Technology, 200093 Shanghai, China*
[2]*Center for Supernetworks Research, University of Shanghai for Science and Technology, 200093 Shanghai, China*
[3]*Key Laboratory of Road and Traffic Engineering of Ministry of Education, College of Transportation Engineering, Tongji University, Shanghai 201804, China*

Correspondence should be addressed to Xin Ye; xye@tongji.edu.cn

Received 23 December 2021; Revised 25 May 2022; Accepted 7 June 2022; Published 27 June 2022

Academic Editor: Ziyuan Pu

This paper developed a mixed multinomial probit (MMNP) model with alternative error specification and random coefficients (for both generic variables and personal attributes) to accommodate flexible covariance structure and taste variation. The MMNP model can be efficiently estimated with analytic approximations of multivariate normal cumulative distribution functions, which avoid defects of simulation-based integration in the mixed multinomial logit (MMNL) model. The integral dimension of the MMNL model increases as random coefficients increase, but it only depends on the number of available alternatives in the MMNP model. Simulation experiments and empirical analysis of Shanghai commuters' mode choice behavior were undertaken to examine the performance of MMNP models. Both simulation results and empirical results show that MMNP models can well accommodate flexible covariance structures and taste variation reflected through random coefficients being associated with both generic and personal variables. Empirical results indicate that the MMNP model performs better than traditional discrete choice models, such as the multinomial logit, the cross-nested logit, MMNL, and multinomial probit models. Random coefficients of "in-vehicle time of car" and "number of companions" indicate taste heterogeneity and the identifiability of random coefficients associated with both generic and personal attributes. Pairwise positive correlations between car/taxi, bus/metro, and bus/bus and metro are to be expected. However, the positive correlation between the car and metro modes may be unique to the Chinese city, Shanghai, because of the developed metro system. Unequal error variances reflect heterogeneities in unspecified factors in commute modes' utilities. The MMNP model will offer an alternative efficient way to accommodate taste heterogeneity and flexible error covariance structure in discrete choice models. Compared with the MMNL model, the MMNP model can accommodate more random coefficients without increasing computational complexity.

1. Introduction

In general, two methods are being applied to overcome the limitation of the IIA property in the multinomial logit (MNL) model by accommodating random taste variation and correlation among alternatives in a discrete choice model. The first method adopts the mixed multinomial logit (MMNL) model, which combines random coefficients varying across individuals to represent taste variation with random error components varying across alternatives to represent correlation among alternatives [1–4]. The second method is a more complex mixed generalized extreme value (GEV) model, such as using a nested logit (NL) model to represent correlation among alternatives and random coefficients to represent taste variation [5–7]. The flexibility of these two methods in representing correlation and random taste variation comes at a cost that choice probabilities can no longer be expressed in a closed-form function. Compared with the first method, the benefit of the second one is that it has fewer dimensions of integration. The disadvantage is that

it has a more complex log-likelihood function. Garrow [8] compared these two methods using NL mixtures to approximate an NL model because, theoretically, mixture analogs can approximate any random utility model [9, 10]. The results show that the NL model cannot be well approximated by its mixture analog because it is difficult to obtain precise correlation estimates from mixture models, particularly for nests that have small correlations, and the accuracy of NL mixture coefficients is sensitive to the number of draws and number of random coefficients in the simulation. Empirical findings in Garrow [8] recommend adopting the second method to accommodate the correlation and random taste variation.

Although the second method, the mixed GEV models, has fewer dimensions of integration, the dimension of integration is the number of random coefficients in the model. With the increase of random coefficients, the integration will be more complicated with higher dimensions. Besides, GEV estimates are biased when negative correlations appear between choices [11]. However, modelers usually do not know the covariance structure in advance. Under this condition, the use of multinomial probit (MNP) models that can accommodate negative correlation will not cause the problem of estimation bias. In addition to negative correlations or covariances, an MNP model can also naturally accommodate unequal variances in its flexible error covariance structure. As well as accounting for random taste variation, this paper considers using a mixed multinomial probit (MMNP) model with random coefficients to represent random taste variation and its error terms to represent a flexible covariance structure among alternatives.

The MMNP model is rarely applied in the transportation field because of the difficulty to evaluate its likelihood function without a closed-form expression. Even if there are some applications, random error terms are usually assumed to be independently distributed [12, 13]. They only use correlated random coefficients to capture the correlation between alternatives. If random error terms are also correlated, will these correlations, as well as correlated random coefficients, be identifiable? Generally, in a mixed model, generic variables, which vary across alternatives, are assumed to take random coefficients. Are all the elements in a covariance-variance matrix of random coefficients for a set of generic variables (e.g., in-vehicle travel times across travel modes) identifiable? Can personal attributes, which do not vary across alternatives, take random coefficients? If so, how can the elements in a covariance-variance matrix of those random coefficients be identified? Some simulation experiments will be conducted to answer those questions in this paper.

The estimation of the MMNP model depends on the computation of a multivariate normal cumulative distribution (MVNCD) function. As the evaluation method of the MVNCD function improves, to avoid the above-mentioned defects of simulation techniques in mixed GEV models, several analytic approximation techniques have been developed [14]. Bhat [14] proposed four new analytic MVNCD approximation methods and compared them with other existing analytic approximation methods and simulation-based integration methods. Regardless of the dimensionality of the MVNCD function, the two-variate bivariate screening (TVBS) approach performs best in the ability to recover MNP parameters accurately and quickly. Therefore, we employed the TVBS analytic approximation evaluation of the MVNCD function to estimate the MMNP model for this study.

This paper aims to develop an MMNP model with correlated random coefficients (for both generic variables and personal attributes) and correlated error structure to accommodate a flexible covariance structure and taste variation, which can be efficiently estimated using analytic approximations of the MVNCD function. And the MMNP model will be applied to analyze commute mode choice behavior in Shanghai, China. The rest of this paper is organized as follows. The second section details the methodology of the MMNP model accommodating random taste variation and flexible covariance structures. In the third section, simulation experiments are undertaken to validate the identifiability when correlated random coefficients are associated with generic variables and personal characteristics. In the fourth section, an empirical study is conducted to explore Shanghai residents' commute mode choice behavior. Discussions and conclusions will be presented in the last section.

2. Modeling Methodology

In this section, the MMNP model will be formulated to accommodate correlated random coefficients and a flexible error covariance structure. For an I-alternative MMNP model, the utility function of alternative i ($i = 1, 2, \ldots, I$, where I denotes the total number of alternatives) is given by

$$U_i = \alpha_i + \beta_i x_i + \varepsilon_i = \alpha_i + \sum_{k=1}^{K} \beta_{ik} x_{ik} + \varepsilon_i, \tag{1}$$

where x_i is a column vector of K explanatory variables, and β_i is a row vector of corresponding coefficients. x_{ik} and β_{ik} are the k^{th} elements of vectors x_i and β_i, respectively. $k = 1, 2, \cdots, K$. ε_i is the random error term for an alternative i. $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_I)'$ follows a multivariate normal distribution with a mean vector of zeros and variance-covariance matrix $\Sigma = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1I} \\ \rho_{12} & \rho_{22} & \cdots & \rho_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1I} & \rho_{2I} & \cdots & \rho_{II} \end{bmatrix}$, or $\varepsilon \sim N(0, \Sigma)$.

β_{ik} is assumed to be a normally distributed random coefficient with mean b_{ik} and standard deviation σ_{ik}. That is, $\beta_{ik} = b_{ik} + \sigma_{ik} z_{ik}$, and herein z_{ik} follows a standard normal distribution. Then, Equation (1) can be rewritten as

$$\begin{aligned} U_i &= \alpha_i + \sum_{k=1}^{K} (b_{ik} + \sigma_{ik} z_{ik}) x_{ik} + \varepsilon_i \\ &= \left(\alpha_i + \sum_{k=1}^{K} b_{ik} x_{ik} \right) + \left(\sum_{k=1}^{K} \sigma_{ik} z_{ik} x_{ik} + \varepsilon_i \right) \\ &= V_i + \xi_i. \end{aligned} \tag{2}$$

where $V_i = \alpha_i + \sum_{k=1}^{K} b_{ik}x_{ik}$ is the systemic utility of alternative i, $\xi_i = \sum_{k=1}^{K} \sigma_{ik}z_{ik}x_{ik} + \varepsilon_i$ which is the total random utility combining the random error term ε_i and the random utility $\sum_{k=1}^{K} \sigma_{ik}z_{ik}x_{ik}$ from the random coefficient β_i. The covariance matrix of $\mathbf{z_k} = (z_{1k}, z_{2k}, \ldots, z_{Ik})'$ can be expressed as $\Lambda^k = \begin{bmatrix} 1 & r_{12} & \ldots & r_{1I} \\ r_{12} & 1 & \ldots & r_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ r_{1I} & r_{2I} & \ldots & 1 \end{bmatrix}$.

Since $\xi_i = \sum_{k=1}^{K} \sigma_{ik}z_{ik}x_{ik} + \varepsilon_i$, mean values of z_{ik} and ε_i are zeros, the mean value of the total random utility ξ_i is also zero, $E(\xi_i) = 0$. Then, the variance-covariance matrix of $\xi = (\xi_1, \xi_2, \ldots, \xi_I)\prime$ can be calculated as

$$\Omega = \sum_{k=1}^{K} \left(\boldsymbol{\sigma}_\mathbf{k} \cdot \boldsymbol{\sigma}_\mathbf{k}'\right) * \Lambda^k * \left(\mathbf{x}_\mathbf{k} \cdot \mathbf{x}_\mathbf{k}'\right) + \Sigma, \tag{3}$$

where $\sigma_\mathbf{k} = (\sigma_{1k}, \sigma_{2k}, \ldots, \sigma_{Ik})'$ and $\mathbf{x_k} = (x_{1k}, x_{2k}, \ldots, x_{Ik})$', which are both column vectors. $\sigma_\mathbf{k}'$ and $\mathbf{x}_\mathbf{k}'$ are the transpose of $\sigma_\mathbf{k}$ and $\mathbf{x_k}$, respectively. "$\cdot$" represents the matrix multiplication, and "$*$" represents Hadamard product or element-wise multiplication. Since $\beta_\mathbf{k} = \mathbf{b_k} + \sigma_\mathbf{k}\mathbf{z_k}$ (where $\mathbf{b_k} = (b_{1k}, b_{2k}, \cdots, b_{Ik})$'), the variance-covariance matrix of random coefficients $\beta_\mathbf{k} = (\beta_{1k}, \beta_{2k}, \ldots, \beta_{Ik})$' can be expressed as $\Psi^k = (\sigma_\mathbf{k} \cdot \sigma_\mathbf{k}') * \Lambda^k$.

Based on the principle of random utility maximum (RUM), the probability that an alternative m is chosen can be expressed as

$$P_m = P(U_m > U_i) = P(\xi_i - \xi_m < V_m - V_i), \quad i = 1, 2, \ldots, I; i \neq m. \tag{4}$$

Let $\xi_{im} = \xi_i - \xi_m$, $-V_{im} = V_m - V_i$ with $\forall i \neq m$. Then, the probability in Equation (4) can be transformed into

$$\begin{aligned} P_m &= P(\xi_{im} < -V_{im}) \\ &= \int_{-\infty}^{-V_{1m}} \cdots \int_{-\infty}^{-V_{im}} \cdots \int_{-\infty}^{-V_{Im}} f(\xi_{1m}, \ldots, \xi_{im}, \ldots, \xi_{Im}) \\ &\quad d\xi_{1m} \ldots d\xi_{im} \ldots d\xi_{Im}, \quad i = 1, 2, \ldots, I;\ i \neq m. \end{aligned} \tag{5}$$

where $f(\cdot)$ is the probability density function of an $(I-1)$-dimensional multivariate normal distribution (MVN).

$E(\xi_{im}) = E(\xi_i - \xi_m) = 0$. The covariance matrix of ξ_{im} $(i = 1, 2, \ldots, I;\ i \neq m)$ is calculated as $\Omega^m = M_m \cdot \Omega \cdot M_m'$, where Ω is the variance-covariance matrix of $\xi = (\xi_1, \xi_2, \cdots, \xi_I)'$ as shown in Equation (3). And M_m is an identity matrix of $(I-1)$ size with an extra column of -1 added as the m^{th} column [15]. For example, $M_m = \begin{bmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ if $m = 2$ and $I = 3$. M_m' is the transpose of M_m.

Therefore, the likelihood function of the maximum likelihood estimation procedure or probability of choosing an alternative m can be expressed as an $(I-1)$-dimensional multivariate normal cumulative distribution (MVNCD) function of $\xi_{im} \sim N(0, \Omega^m)$ as shown in Equation (5). With the development of analytic approximations of the MVNCD function, the two-variate bivariate screening (TVBS) approach is adopted to estimate the MMNP model in this paper. The TVBS approach can recover underlying parameters accurately and efficiently, regardless of the dimensionality involved in the MVNCD function. For more details of the TVBS approach, one can refer to Bhat [14].

Additionally, in an I-alternative MMNP model, $(I(I-1)/2-1)$ elements in the covariance matrix Σ of correlated error terms can be identified [15]. If correlated random coefficients are associated with generic variables, all means b and all elements in the variance-covariance matrix Ψ^k are identifiable. If correlated random coefficients are associated with a personal attribute that appears in utility functions of all the available alternatives, similar to an alternative-specific constant, only $(I-1)$ means of random coefficients for this personal attribute are identifiable. And the covariance matrix of random coefficients β is similar to that of error terms in identifiability. That is, $(I(I-1)/2-1)$ elements in the variance-covariance matrix of random coefficients Ψ^k can be identified. Simulation experiments will be conducted in the next section to demonstrate the number of identifiable elements in covariance matrices.

3. Simulation Experiments

Correlations among alternatives may exist in both correlated error structure and correlated random coefficients. To verify the identifiability of parameters, two simulation experiment scenarios were designed when correlated random coefficients were associated with generic variables and personal characteristics. The estimation of a model with independent random coefficients across alternatives is a special case of that with correlated random coefficients. All the correlations can be fixed at 0. Random error terms are correlated across alternatives in both simulation experiments.

In simulation experiments, we assume that all alternatives are available for each individual. Explanatory variables follow IID uniform distributions between 0 and 10. Dependent variables y_i are determined based on the RUM principle:

$$y_i = D[U_i = \text{Max}(U_1, U_2, ..., U_I)], \quad (i = 1, 2, 3, ..., I). \tag{6}$$

If U_i is the maximum utility among all of the available I alternatives for an individual, $D[\cdot] = 1$; otherwise, $D[\cdot] = 0$.

3.1. Experimental Design

3.1.1. Correlated Random Coefficients Associated with Generic Variables. For a four-alternative MMNP model, we assumed that four random coefficients are associated with four independent generic variables x_1, x_2, x_3, and x_4. Utility functions of four alternatives are given as follows:

$$\begin{aligned} U_1 &= \alpha_1 + \beta_1 x_1 + \varepsilon_1 = \alpha_1 + (b_1 + \sigma_1 z_1) x_1 + \varepsilon_1, \\ U_2 &= \alpha_2 + \beta_2 x_2 + \varepsilon_2 = \alpha_2 + (b_2 + \sigma_2 z_2) x_2 + \varepsilon_2, \\ U_3 &= \alpha_3 + \beta_3 x_3 + \varepsilon_3 = \alpha_3 + (b_3 + \sigma_3 z_3) x_3 + \varepsilon_3, \\ U_4 &= \beta_4 x_4 + \varepsilon_4 = (b_4 + \sigma_4 z_4) x_4 + \varepsilon_4. \end{aligned} \tag{7}$$

where random coefficients $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)'$ are normally distributed with a mean vector $b = (b_1, b_2, b_3, b_4)'$ and a standard deviation vector $\sigma = (\sigma_1, \sigma_2, \sigma_3, \sigma_4)'$. Since random coefficients are assumed to be correlated across alternatives, z_1, z_2, z_3, and z_4 are correlated and follow a standard multivariate normal distribution with a mean vector $(0, 0, 0, 0)'$ and a covariance matrix $\Lambda = \begin{pmatrix} 1 & r_{12} & r_{13} & r_{14} \\ r_{12} & 1 & r_{23} & r_{24} \\ r_{13} & r_{23} & 1 & r_{34} \\ r_{14} & r_{24} & r_{34} & 1 \end{pmatrix}$. Then, the variance-covariance matrix of $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)'$ is calculated as $\Psi = (\sigma \cdot \sigma\prime) * \Lambda$. Random error terms ε_1, ε_2, ε_3, and ε_4 follow a standard multivariate normal distribution with a mean vector $(0, 0, 0, 0)'$ and covariance matrix $\Sigma = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{12} & \rho_{22} & \rho_{23} & \rho_{24} \\ \rho_{13} & \rho_{23} & \rho_{33} & \rho_{34} \\ \rho_{14} & \rho_{24} & \rho_{34} & \rho_{44} \end{pmatrix}$.

All elements in the covariance matrix Ψ can be estimated in the MMNP model. Since only five elements in the error term covariance matrix Σ can be identified in a four-alternative MMNP model, we fixed all the diagonal elements at 1 and the correlation ρ_{34}. Thus, parameters to be estimated include all α values, all b values, all elements in the covariance matrix Ψ, and five covariance elements $(\rho_{12}, \rho_{13}, \rho_{23}, \rho_{14}, \rho_{24})$ in the covariance matrix Σ. ρ_{13} and ρ_{23} are set as negative to validate the ability of the MMNP model in accommodating negative correlation.

3.1.2. Correlated Random Coefficients Associated with Personal Attributes. Generally, in a mixed model, random coefficients are assumed to be associated with generic variables, such as travel time, cost, or other level-of-service variables of different travel modes. The identifiability when random coefficients are associated with personal attributes has not been explored in previous studies. In this subsection, random coefficients of the same personal attribute are assumed to be correlated across alternatives and vary across individuals and alternatives. A three-alternative MMNP model with utility functions is considered as follows:

$$\begin{aligned} U_1 &= \alpha_1 + \beta_1 x_1 + \varepsilon_1 = \alpha_1 + (b_1 + \sigma_1 z_1) x_1 + \varepsilon_1, \\ U_2 &= \alpha_2 + \beta_2 x_1 + \varepsilon_2 = \alpha_2 + (b_2 + \sigma_2 z_2) x_1 + \varepsilon_2, \\ U_3 &= \beta_3 x_1 + \varepsilon_3 = (b_3 + \sigma_3 z_3) x_1 + \varepsilon_3. \end{aligned} \tag{8}$$

where random coefficients $\beta = (\beta_1, \beta_2, \beta_3)'$ follow normal distributions with a mean vector $b = (b_1, b_2, b_3)'$ and a standard deviation vector $\sigma = (\sigma_1, \sigma_2, \sigma_3)\prime$. Since random coefficients $\beta_1, \beta_2, \beta_3$ are correlated across alternatives, z_1, z_2 and z_3 follow a standard multivariate normal distribution with a mean vector $(0, 0, 0)$' and a covariance matrix $\Lambda = \begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{12} & 1 & r_{23} \\ r_{13} & r_{23} & 1 \end{pmatrix}$. Then, the variance-covariance matrix of $\beta = (\beta_1, \beta_2, \beta_3)'$ is calculated as $\Psi = (\sigma \cdot \sigma') * \Lambda$. Random error terms ε_1, ε_2, and ε_3 follow a standard multivariate normal distribution with a mean vector $(0, 0, 0)'$ and a covariance matrix $\Sigma = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{12} & \rho_{22} & \rho_{23} \\ \rho_{13} & \rho_{23} & \rho_{33} \end{pmatrix}$.

In this case, random coefficients are associated with the same explanatory variable x_1 in utility functions of three alternatives. With consideration of the identification problem, one of the mean values of random coefficients (b_1, b_2 or b_3) should be fixed as a reference, similar to an alternative-specific constant. The covariance matrix of random coefficients β is similar to that of error terms in identifiability. In a three-alternative MMNP model, only two elements in the covariance matrix of error terms Σ can be identified. Therefore, all parameters to be estimated in this scenario are all α values, two of b values (b_1 was fixed), one covariance element r_{23} and one standard deviation σ_3 in the covariance matrix Ψ, and two covariance elements (ρ_{12} and ρ_{13}) in the covariance matrix of error terms Σ. r_{23} and ρ_{13} are set as negative to validate the ability of the MMNP model in accommodating negative correlation.

3.2. Simulation Results. Simulation results of the MMNP models with correlated error structure and correlated random coefficients associated with generic variables and personal attributes are shown in Tables 1 and 2. Variances in both covariance matrices of error terms are assumed to be identical in the two experiments. Parameters are also consistently estimated for heteroscedastic error terms. For brevity, only the results of homoscedastic error terms are presented here.

According to significant t-statistics at the 0.01 level shown in Tables 1 and 2, all identifiable parameters can be estimated and are seemingly consistent with their true values, indicating that correlations among alternatives can be accommodated in both error structure and random coefficients for generic variables. In empirical studies, if correlations from only one source are considered, it may lead to bias in model estimation. Results in Table 2 show that, in addition to generic variables, personal attributes can also be specified with random coefficients. In real data, respondents may have different preferences for the same generic variables (such as travel time and cost of different travel modes), and individuals with different characteristics also have different preferences for alternative modes. And the MMNP model can well accommodate random taste variation in both generic variables and personal attributes through random coefficients. As significant t-statistics at the 0.01 level are shown in Tables 1 and 2, all negative covariance elements in covariance matrices of random error terms are correctly estimated, which indicates that the MMNP model can well accommodate various correlated error structures, even with negative correlations. It can perfectly avoid defects in previous MMNP models and mixed GEV models since

Table 1: Simulation results for a four-alternative MMNP model of 500,000 observations when random coefficients are associated with generic variables.

Parameters	True values	Estimates	Std. err.	t-stat
α_1	−0.5	−0.498	0.074	−6.776***
α_2	−0.4	−0.377	0.067	−5.640***
α_3	−0.3	−0.259	0.047	−5.551***
b_1	0.5	0.511	0.044	11.732***
b_2	0.3	0.306	0.027	11.563***
b_3	0.2	0.202	0.015	13.503***
b_4	0.1	0.099	0.007	13.431***
σ_1	0.2	0.206	0.018	11.658***
σ_2	0.4	0.404	0.035	11.722***
σ_3	0.6	0.606	0.051	11.833***
σ_4	0.5	0.528	0.045	11.667***
r_{12}	0.2	0.203	0.016	12.396***
r_{13}	0.4	0.379	0.015	25.196***
r_{14}	0.5	0.541	0.013	40.782***
r_{23}	0.6	0.608	0.008	74.205***
r_{24}	0.3	0.299	0.011	28.064***
r_{34}	0.4	0.401	0.010	42.107***
ρ_{11}	1.0	1.000	—	—
ρ_{12}	0.5	0.511	0.087	5.843***
ρ_{13}	−0.3	−0.341	0.105	−3.250***
ρ_{14}	0.6	0.524	0.080	6.521***
ρ_{22}	1.0	1.000	—	—
ρ_{23}	−0.4	−0.372	0.127	−2.932***
ρ_{24}	0.5	0.532	0.096	5.517***
ρ_{33}	1.0	1.000	—	—
ρ_{34}	0.3	0.300	—	—
ρ_{44}	1.0	1.000	—	—

Note. "—" the corresponding parameter is fixed; ***significant at 0.01 level.

Table 2: Simulation results for a three-alternative MMNP model of 5,000,000 observations when random coefficients are associated with personal attributes.

Parameters	True values	Estimates	Std. err.	t-stat
α_1	−0.5	−0.490	0.008	−63.097***
α_2	−0.4	−0.396	0.050	−7.894***
b_1	0.5	0.500	—	—
b_2	0.3	0.301	0.018	17.119***
b_3	0.2	0.205	0.007	30.905***
σ_1	0.2	0.200	—	—
σ_2	0.4	0.400	—	—
σ_3	0.6	0.592	0.029	20.669***
r_{12}	0.2	0.200	—	—
r_{13}	0.4	0.400	—	—
r_{23}	−0.6	−0.605	0.089	−6.768***
ρ_{11}	1.0	1.000	—	—
ρ_{12}	0.5	0.510	0.075	6.827***
ρ_{13}	−0.3	−0.313	0.079	−3.958***
ρ_{22}	1.0	1.000	—	—
ρ_{23}	0.4	0.400	—	—
ρ_{33}	1.0	1.000	—	—

Note. "—" the corresponding parameter is fixed; ***significant at 0.01 level.

error terms are assumed to be independent in previous MMNP models [12, 13], and negative correlations cannot be accommodated in the mixed GEV models [15], which may lead to bias in model estimation [11].

However, the sample size for ensuring the significance of all the identifiable parameters is great in MMNP models. For the four-alternative MMNP model with random coefficients associated with generic variables, the sample size is around 500,000 in our experiment to ensure the significance of all the critical coefficients. The sample size for the three-alternative MMNP model with random coefficients of personal attributes is around 5,000,000 to ensure significance. Both sample sizes are much larger than those in usual empirical studies. As per Walker [16], although model parameters are theoretically identifiable, there are issues of empirical identification, namely, whether the data can support the model specification. If the empirical data cannot allow for estimating all the theoretically identifiable parameters, some parameters may be fixed, and only parameters of particular interest need to be estimated.

To improve the efficiency of estimation and decrease the required sample size for statistically significant identification, z_1, z_2, z_3, and z_4 in the random coefficients of the four-alternative MMNP model can be set as $z_1 = z_2 = z_3 = z_4$. That is, z_1, z_2, z_3, and z_4 are linearly correlated with a correlation coefficient of 1. And they follow the standard normal distribution. Then, the required sample size for identification can be decreased to 5000. The simulation results with a sample size of 5000 observations are shown in Table 3.

4. Empirical Study

To examine the performance of the MMNP model in accommodating random taste variation and correlations among alternatives, an application is presented in this section to analyze the commute mode choice behavior of Shanghai residents.

4.1. Data and Sample Description

4.1.1. Commute Trip Data. The commute trip data used in this study were derived from a web-based travel survey of Shanghai conducted in 2018. The survey collected detailed socioeconomic and demographic information, and a complete 24-hour travel diary reported by 2033 individuals. The detailed trip data included origin and destination locations, trip beginning and ending times, travel mode, trip purpose, and the number of companions. Six travel modes are considered in this study including car, taxi, metro, bus, a combination of bus and metro (hereinafter referred to as Bus and Metro), and the nonmotorized mode (bicycling and walking). After data screening to remove records with missing data for explanatory variables of interest and travel mode choice, the final sample comprised 1743 commute trips.

4.1.2. Level-of-Service Data. The LOS characteristics of different travel modes were extracted from the zone-to-zone travel impedance matrices generated from transportation networks of Shanghai, which were integrated into TransCAD based on the GIS data of roads (links and nodes), bus

Table 3: Simulation results for a four-alternative MMNP model of 5,000 observations when random coefficients are associated with generic variables.

Parameters	True values	Estimates	std. err.	t-stat
α_1	−0.5	−0.792	0.192	−4.126***
α_2	−0.4	0.561	0.066	8.481***
α_3	−0.3	−0.714	0.204	−3.500***
b_1	0.5	0.326	0.040	8.154***
b_2	0.3	−0.455	0.218	−2.091**
b_3	0.2	0.188	0.034	5.488***
b_4	0.1	0.062	0.032	1.943*
σ_1	0.2	0.199	0.024	8.410***
σ_2	0.4	0.443	0.049	8.995***
σ_3	0.6	0.660	0.069	9.521***
σ_4	0.5	0.580	0.060	9.746***
r_{12}	1.0	1.000	—	—
r_{13}	1.0	1.000	—	—
r_{14}	1.0	1.000	—	—
r_{23}	1.0	1.000	—	—
r_{24}	1.0	1.000	—	—
r_{34}	1.0	1.000	—	—
ρ_{11}	1.0	1.000	—	—
ρ_{12}	0.5	0.476	0.159	2.992***
ρ_{13}	−0.3	−0.340	0.125	−2.733***
ρ_{14}	0.6	0.628	0.137	4.584***
ρ_{22}	1.0	1.000	—	—
ρ_{23}	−0.4	−1.148	0.441	−2.602***
ρ_{24}	0.5	0.430	0.206	2.089**
ρ_{33}	1.0	1.000	—	—
ρ_{34}	0.3	0.300	—	—
ρ_{44}	1.0	1.000	—	—

Note. "—" the corresponding parameter is fixed; ***significant at 0.01 level, **significant at 0.05 level, *significant at 0.1 level.

lines and stops, metro lines and stations, and residential committees. Travel time calculated from floating car data was added to each link of the road network [17, 18]. Travel times, service frequencies, and fares by the time of day were applied to each trip OD pair in the sample.

4.1.3. Sample Description. Table 4 provides a summary of the socioeconomic and demographic characteristics of the sample. Most commuters are 20–40 years old, married, and well educated, hold a driving license, and live with family. Personal monthly income mainly falls between 4.5k and 15k RMB Yuan. More than half (62.2%) of commuters have an available private car on the travel day. Origin TAZs (or residential areas) have greater population density and lower employment density than destination TAZs (or workplaces), as expected.

Descriptive statistics of LOS characteristics are shown in Table 5. The average trip distance of nonmotorized mode is 4.204 km, which indicates that nonmotorized mode can meet most of the short-distance commute demands. In terms of average fare, the taxi has the highest fare because Shanghai taxi has a high starting price. In terms of average in-vehicle time, car, taxi, and metro are all around 20 minutes, while the bus and Bus and Metro rise obviously to 37.999 and 38.747 minutes, respectively. In the Bus and Metro mode, the in-vehicle time of the metro is longer than that of the bus, indicating that the bus is a complementary mode to the metro. The average initial waiting time and transfer waiting time of the metro are shorter than those of bus and Bus and Metro, and they vary much less, meaning that the metro system provides a higher service frequency, and the service frequency distribution is more even among different lines. In terms of average access/egress distance, bus and Bus and Metro have substantially shorter distances than the metro because the bus has a higher station coverage rate. And metro's average access distance is slightly longer than the average egress distance presumably due to the denser distribution of metro stations in workplaces than in residential areas. The Bus and Metro mode takes a slightly greater number of transfers than bus and metro, as expected.

4.2. Empirical Results. In this section, the empirical results of models, including the MNL, CNL (cross-nested logit model, shown in Figure 1), MMNL, MNP, and MMNP models, are presented in Table 6. And then, the data fit of the MMNP model is compared with that of the MNL, CNL, MMNL, and MNP models.

The MMNP model can accommodate both random taste heterogeneity and flexible covariance structure among alternatives. Compared to the MMNP model, the MNL, the CNL, and the MNP models cannot accommodate random coefficients. Although the MMNL model can capture the taste variation across individuals by random coefficients, it cannot accommodate the variance heterogeneity of different travel modes. The following paragraph will focus on model results of the MMNP model because coefficients of explanatory variables have the same signs and similar values.

4.2.1. Model Results of the MMNP Model. (1) Random Coefficients. In the empirical study, both random coefficients of a generic variable and a personal attribute appear to be significant, indicating that there are variations among commuters' sensitivities to those variables.

As a generic variable, the in-vehicle time of the car has a mean coefficient of −2.662 and a standard deviation of 1.050. The negative mean value indicates that the utility of the car will decrease with the increase of in-vehicle time on average across individuals, as expected. The value of standard deviation is different from zero significantly, implying that random taste heterogeneity does exist among commuters in this LOS variable. And the value of standard deviation in this MMNP model is quite similar to that in MMNL models [19, 20].

"The number of companions" has a random coefficient with a mean value of 0.423 and a standard deviation of 0.828 in the utility function of the car mode. The positive mean value indicates that "the number of companions," on average, has a positive impact on the choice of car because commuters can share the cost while enjoying the flexibility and mobility of the car, which is more cost-effective than driving alone. The standard deviation is significantly different from zero. "The number of companions" is a kind of personal attribute because the variable remains the same in all alternatives' utility functions. Therefore, it verifies that the

Table 4: Sample characteristics ($N = 1743$).

Categorical variables	Categories		Percentage (%)
Gender	Male		45.439
	Female		54.561
Age	Younger than 20 years		2.811
	21–30 years		34.079
	31–40 years		46.242
	41–50 years		13.827
	Above 50 years		3.041
Marital status	Married		72.060
	Unmarried		27.940
Driving license	Own		75.961
	Does not own		24.039
Residential type	Family		87.722
	Dormitory		11.819
	Other		0.459
Education attainment	Less than high school		1.549
	High school		6.311
	Bachelor's degree		78.543
	Graduate degree		13.597
Personal monthly income (RMB yuan)	Less than 2k		2.754
	2k–4.5k		8.319
	4.5k–6k		14.974
	6k–8k		17.843
	8k–10k		20.023
	10k–15k		20.367
	15k–20k		7.975
	20k–30k		5.221
	Greater than 30 k		2.524
Private vehicles	Car		62.192
	Electric bicycle		29.260
	Bicycle		30.293
	Motorcycle		2.008
Continuous variables		Mean	S.D.
Personal income[(1)] (10k RMB Yuan/month)		1.018	0.638
Population density (thousand people per square kilometer)	Origin TAZ	2.538	2.178
	Destination TAZ	1.252	1.855
Employment density (thousand people per square kilometer)	Origin TAZ	1.794	1.994
	Destination TAZ	2.367	2.704

(1) The continuous variable "Personal Income" was generated from the categorical variable "Personal Monthly Income." The exchange rate of RMB Yuan to the US dollar is approximately 6.9 : 1.

random coefficient of a personal attribute can also be estimated in an empirical study.

(2) Error Covariance Structure. As shown in Table 6, the CNL, the MNP, and the MMNP models all reveal positive correlations for the pairwise alternatives of car/taxi, car/metro, bus/metro, and bus/bus and metro. In the context of single choice models (that is, when only one alternative can be chosen), such positive correlations imply a higher level of substitutability between the corresponding pairs of modes relative to assuming zero correlation (as in the MNL).

Of the pairwise correlations, those between car/taxi, bus/metro, and bus/bus and metro are to be expected. Cars and taxis are characterized by their mobility and flexibility, and both belong to the category of a private transportation mode. Meanwhile, bus and metro are both public transportation modes and have many aspects in common. However, the positive correlation between the car and metro modes, even though less than the positive correlations between car/taxi and bus/metro, is particularly interesting because one is a private mode, while the other is a public transportation mode. This result, though perhaps likely to be unique to many Chinese cities, is probably because the Shanghai metro has been well developed and provides a convenient, high-frequency, and punctual service. Shanghai metro operates on a strict schedule, making it attractive to most commuters. The Shanghai metro operates at a speed (almost 30 km/h) even faster than the speed of road traffic during peak hours. In addition, it is difficult to find a temporary parking lot during peak hours, and it is also highly expensive to reserve a parking lot because of limited parking spaces around workplaces. This finding suggests that improvement in metro LOS can be effective in attracting commuters away from private cars to the metro and alleviate traffic congestion during peak hours.

The variance of the bus mode is also shown in Table 6. Due to the limited sample size, not all identifiable parameters appear to be significant. Variances of the other modes (car, taxi, metro, bus and metro, and nonmotorized modes)

Table 5: Descriptive statistics of LOS (level-of-service) attributes.

	Attribute	Mean	S.D.
CAR	Fare (RMB yuan)	7.065	6.471
	In-vehicle time (min)	23.406	16.604
Taxi	Fare (RMB yuan)	35.490	28.880
	In-vehicle time (min)	23.406	16.604
Metro	Fare (RMB yuan)	3.938	0.948
	In-vehicle time (min)	20.469	14.714
	Initial waiting time (min)	2.164	0.317
	Transfer waiting time (min)	1.226	1.455
	Access distance (km)	1.826	1.127
	Egress distance (km)	1.327	0.978
	Number of transfers	0.568	0.669
Bus	Fare (RMB yuan)	3.458	1.707
	In-vehicle time (min)	37.999	29.897
	Initial waiting time (min)	4.281	2.751
	Transfer waiting time (min)	3.152	4.501
	Transfer walking distance (km)	0.110	0.240
	Access distance (km)	0.657	0.427
	Egress distance (km)	0.638	0.430
	Number of transfers	0.729	0.854
Bus and metro	Fare (RMB yuan)	3.610	1.626
	Total in-vehicle time (min)	38.747	18.981
	In-vehicle time allocated to metro (min)	22.060	14.584
	In-vehicle time allocated to bus (min)	16.687	14.436
	Initial waiting time (min)	3.599	2.658
	Access distance (km)	0.720	0.441
	Egress distance (km)	0.679	0.407
	Number of transfers	0.994	1.026
	Transfer waiting time (min)	2.833	3.556
	Transfer walking distance (km)	0.126	0.258
Nonmotorized mode	Trip distance (km)	4.204	5.611

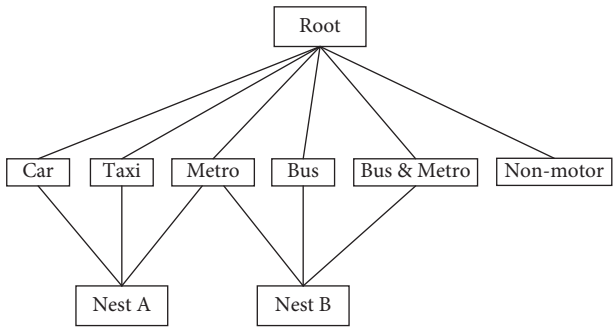


Figure 1: Nesting structure of the CNL model.

are fixed at 1 as a base. Consistent with the results in Bhat [21], the variance of the bus mode is greater than that of car, taxi, and metro, indicating that unobserved variables affecting the choice of bus have greater variances than those affecting car, taxi, and metro. For example, the comfort level is an unobserved variable whose values may vary considerably within bus modes because different kinds of bus services in Shanghai, such as customized buses (or demand-response transit), bus rapid transit (BRT), and regular buses, may have different occupancy rates. However, comfort varies little in the car, taxi, and metro. Therefore, the random component of the bus mode has greater variance than that of the other modes.

(3) Level-of-Service Variables. The in-vehicle time takes expected negative effects on the choice of car, taxi, metro, and bus. Taxi has the most negative coefficient value because the taxi fare ranks first among all the modes. Commuters can spend their in-vehicle time entertaining and do not have to focus on traffic conditions when they take public transit. As a result, coefficients of in-vehicle time for metro and bus are less than those for cars and taxis.

Access/egress distance reduces utilities of public transit (metro, bus, and Bus and Metro), indicating that longer access/egress distance will discourage commuters from choosing public transit. And the coefficient of access/egress distance for the metro is lower than that for bus, Bus and Metro. Because the bus has a higher station coverage rate than the metro, and the metro has a much more comfortable environment than the bus, commuters are more likely to accept the larger access/egress distance of the metro.

A negative coefficient of "the number of transfers" implies that the greater the number of transfers is, the less likely the commuters are to take the bus. Usually, the bus is not as punctual as the metro. The more the number of transfers is, the greater the probability that commuters will encounter bus delays. Besides, the transferring environment at public transit stations is crowded and uncomfortable. Commuters are usually anxiously waiting or in a hurry.

Table 6: Empirical results of models ($N = 1743$).

Variable	MNL		CNL		MMNL		MNP		MMNP	
	Coefficient	*t*-statistic	Coefficient	*t*-statistic	Coefficient	*t*-statistic	Coefficient	*t*-statistic	Coefficient	*t*-statistic
Alternate specific constants (nonmotorized mode is base)										
Car	−5.312	−12.38	−5.320	−12.42	−6.052	−11.23	−3.578	−5.89	−4.517	−6.58
Taxi	−4.834	−9.31	−4.840	−10.49	−4.679	−10.20	−3.023	−8.18	−2.888	−8.52
Metro	0.236	0.78	0.277	0.83	0.192	0.64	0.755	2.18	0.674	2.12
Bus	−0.226	−0.89	−0.258	−0.93	−0.255	−0.98	−0.014	−1.14	−0.607	−1.25
Bus and metro	−0.592	−1.32	−0.575	−1.19	−0.599	−1.33	−1.556	−2.47	−1.111	−1.85
Level-of-service variables										
In-vehicle time (h)										
Car										
Mean	−2.846	−6.28	−2.862	−6.68	−2.877	−6.59	−2.452	−3.88	−2.662	−4.06
Standard deviation	--	--	--	--	0.427	0.55	--	--	1.050	3.75
Taxi	−3.238	−4.90	−3.252	−4.24	−3.486	−4.28	−2.938	−4.38	−3.853	−4.31
Metro	−1.284	−2.71	−1.284	−2.78	−1.354	−2.99	−1.484	−2.53	−1.516	−2.75
Bus	−0.951	−3.05	−1.068	−3.10	−0.995	−3.28	−0.654	−1.87	−0.710	−1.84
Access/Egress distance (km)										
Metro	−0.529	−7.97	−0.552	−7.56	−0.533	−7.90	−0.501	−3.94	−0.480	−4.21
Bus	−0.550	−3.81	−0.563	−3.55	−0.561	−3.84	−0.590	−2.28	−0.690	−2.50
Bus and metro	−0.702	−3.22	−0.755	−2.86	−0.711	−2.93	−0.357	−2.14	−0.771	−2.11
The number of transfers-bus	−0.813	−5.13	−0.866	−4.71	−0.826	−5.12	−1.074	−2.86	−1.196	−2.96
Transfer waiting time (min), Bus and metro	−0.237	−4.87	−0.245	−4.71	−0.244	−4.99	−0.133	−2.66	−0.267	−3.06
Trip distance (km), nonmotorized mode	−0.256	−14.68	−0.255	−11.47	−0.261	−11.63	−0.178	−4.73	−0.185	−4.83
Sociodemographic and employment variables										
Married										
Car	0.554	3.21	0.557	3.27	0.547	2.99	0.331	2.28	0.314	1.94
Personal monthly income (10K RMB yuan)										
Car	0.812	7.32	0.828	7.13	0.841	6.97	0.620	4.09	0.718	4.73
Taxi	1.182	4.18	1.190	5.14	1.162	4.99	0.840	3.63	0.868	4.08
Metro	0.373	2.63	0.397	2.61	0.373	2.58	0.229	2.01	0.268	1.67
The number of companions										
Car										
Mean	0.234	6.15	0.232	3.62	0.494	4.69	0.149	3.43	0.423	3.75
Standard deviation	—	—	—	—	0.728	3.12	—	—	0.828	5.21
Taxi	0.341	3.07	0.339	3.46	0.174	2.17	0.191	2.48	0.150	2.07
Having a driving license										
Car	2.811	9.25	2.810	9.27	3.191	8.48	1.917	5.51	2.412	5.33
Living with family										
Car	1.315	4.29	1.310	4.44	1.556	4.58	1.026	3.70	1.323	4.70
Having available private cars on the trip day										
Metro	−0.775	−4.47	−0.819	−4.73	−0.736	−4.41	−0.896	−2.86	−0.852	−2.94
Bus	−0.861	−5.04	−0.919	−4.82	−0.818	−4.84	−1.477	−2.75	−1.509	−2.82
Having available private bikes on the trip day										
Nonmotorized mode	0.719	5.01	0.719	5.25	0.741	5.29	0.659	3.81	0.695	4.26

Table 6: Continued.

Variable	MNL		CNL		MMNL		MNP		MMNP	
	Coefficient	*t*-statistic	Coefficient	*t*-statistic	Coefficient	*t*-statistic	Coefficient	*t*-statistic	Coefficient	*t*-statistic
Workplace employment density (1k positions/km2)										
Taxi	0.156	2.35	0.157	2.90	0.165	2.99	0.101	2.82	0.079	1.94
Metro	0.119	4.05	0.118	3.99	0.126	4.36	0.103	2.70	0.103	2.85
Bus	0.109	3.49	0.102	2.84	0.116	3.68	0.063	1.98	0.082	1.92
Error covariance Structure										
Var (Bus)	—	—	—	—	—	—	2.457	4.32	2.646	2.96
Corr (Car, taxi)	—	—	0.662	3.82	—	—	0.835	2.37	0.765	2.28
Corr (Car, metro)	—	—	0.200	3.26	—	—	0.436	2.05	0.364	2.55
Corr (Bus, metro)	—	—	0.631	7.47	—	—	0.681	2.77	0.617	3.69
Corr (Bus, bus and metro)	—	—	0.662	6.94	—	—	0.643	2.84	0.629	2.54

Table 7: Measures of model performance.

Goodness-of-fit measures	MNL	CNL	MMNL	MNP	MMNP
LL (β)	−1764.631	−1757.368	−1755.445	−1743.360	−1733.051
LL (0)	−2659.259	−2659.259	−2659.259	−2659.259	−2659.259
ρ^2 (0)	0.336	0.339	0.340	0.344	0.348
Adj. ρ^2 (0)	0.326	0.327	0.328	0.332	0.335
AIC	3587.261	3580.736	3572.890	3554.720	3538.103
BIC	3745.699	3761.027	3742.254	3740.474	3734.784
Non-nested test	5.71E-17	3.25E-14	1.89E-11	2.89E-07	--

Similarly, the coefficient of "transfer waiting time" is also negative for Bus and Metro.

While the nonmotorized mode mainly serves short-distance commute trips as shown in Table 5, "trip distance" takes a significantly negative coefficient, as expected.

(4) Sociodemographic and Employment Variables. Married commuters tend to drive a car because they usually need to drive their spouse to work or drive children to school during the commute trip. Commuting by car can provide maximum flexibility and convenience. Likewise, commuters living with families also prefer to use a car. "Personal monthly income" has positive coefficients in utilities of car, taxi, and metro because high-income commuters tend to use high-quality, but high-cost services. With the emerging car-sharing services, having a driving license becomes the only requirement for commuters to drive a car as the commute mode. Therefore, having a driving license will increase the likelihood of commuting by car. Compared with using a car-sharing service, "having available private cars on the trip day" will encourage commuters to drive a car to commute because commuters do not have to walk to car-sharing stations from home and return the car to designated stations. On the other hand, "having available private cars on the trip day" will decrease the likelihood of taking public transit. "Having available private bikes on the trip day" provides a more convenient opportunity for commuters to commute by the nonmotorized mode as evidenced by a positive coefficient. Higher employment densities of workplaces indicate lower parking facility densities and higher parking fees. Commuters working in these areas will be less likely to drive to work, and the likelihood of taking a taxi, bus, and metro will increase.

4.2.2. Comparisons of Data Fit with the MNL, CNL, MMNL, and MNP Models. To evaluate the data fit of the MMNP model, other four models were also estimated: a basic six-alternative MNL model, a CNL model with nesting structure shown in Figure 1, an MMNL model with random coefficients associated with the same variables (in-vehicle time of the car, and the number of companions) as in the MMNP model, and an MNP model with the same correlated error structure as in the MMNP model.

To compare the model fitting performance of the MMNP model with those of the other four models, multiple goodness-of-fit measures are calculated, including $\rho^2(0)$, adjusted $\rho^2(0)$, Akaike Information Criteria (AIC), and Bayesian Information Criterion (BIC):

$$\rho^2(0) = 1 - \frac{\mathrm{LL}(\beta)}{\mathrm{LL}(0)},$$
$$\mathrm{adj}.\rho^2(0) = 1 - \frac{\mathrm{LL}(\beta) - R}{\mathrm{LL}(0)}. \tag{9}$$

where $LL(0)$ is the log-likelihood value of the model without any explanatory variables. $LL(\beta)$ is the log-likelihood value at convergence.

$$\mathrm{AIC} = 2(R - \mathrm{LL}(\beta)),$$
$$\mathrm{BIC} = -2\mathrm{LL}(\beta) + R \cdot \ln(N). \tag{10}$$

where R is the number of parameters estimated in the model, and N is the number of observations.

Further, to compare the logit-based models with the MMNP model using a statistical test, a nonnested test statistic is applied (see [22–24]). The adjusted log-likelihood ratio statistic of each model is first computed:

$$\overline{\rho}^2 = 1 - \frac{\mathrm{LL}(\beta) - R/2}{\mathrm{LL}(0)}. \tag{11}$$

The nonnested test statistic is

$$P\left(\overline{\rho}_1^2 - \overline{\rho}_2^2 > z\right) \leq \Phi\left(-\sqrt{-2 \cdot \mathrm{LL}(0) \cdot z}\right). \tag{12}$$

where Φ represents the CDF of standard normality, and z takes a positive value. That is, the probability that $(\overline{\rho}_1^2 - \overline{\rho}_2^2 > z)$ could have occurred by chance is no larger than $\Phi\left(-\sqrt{-2 \cdot \mathrm{LL}(0) \cdot z}\right)$. A small value for the nonnested test statistic indicates a small probability of erroneously selecting the incorrect model. And the model with a higher value for the adjusted likelihood statistic is preferred. The nonnested test is then used to compare the MMNP model with the other four models.

Based on $\rho^2(0)$, adjusted $\rho^2(0)$, the model with the highest value performs best. And the model with the smallest values of AIC and BIC has the best performance. As shown in Table 7, all the measures indicate the same result that the MMNP model performs better than the other four models with the highest values of $\rho^2(0)$ and adjusted $\rho^2(0)$, and the smallest AIC and BIC values. It indicates that the MMNP can not only accommodate random taste variation and flexible covariance structures among alternatives, but also perform better than the traditional discrete choice models. From the informal nonnested likelihood statistic values provided in Table 7, it can be inferred that the probability of erroneously selecting the MMNP model is literally zero.

5. Discussion and Conclusion

This paper developed an MMNP model with correlated error structure and random coefficients for both generic variables and personal attributes to accommodate flexible covariance structure and random taste variation. And the MMNP model can be efficiently estimated with analytic approximations of the multivariate normal cumulative distribution (MVNCD) function proposed by Bhat [14]. Both simulation experiments and empirical analysis of Shanghai commuters' mode choice behavior are carried out to examine the performance of the MMNP model.

Compared to the MMNP model, the MNL, the CNL, and the MNP model cannot accommodate random coefficients. Although the MMNL model can capture the taste variation across individuals by random coefficients, it cannot accommodate the variance heterogeneity of different alternatives. Simulation results show that the MMNP model can accommodate a flexible covariance structure, even with negative correlations that cannot be accommodated in mixed GEV models. However, the neglect of negative correlation in mixed GEV models may lead to estimation bias. Correlations among alternatives can exist in both error structure and random coefficients for generic variables, while most previous applications of MMNP models assumed that random error terms were independent of each other, which may also lead to bias in model estimation. In addition to correlated random coefficients for generic variables, correlated random coefficients for personal attributes can also be significantly identified.

Empirical results are consistent with simulation results in the fact that random coefficients can exist in both generic variables and personal attributes in MMNP models. As a generic variable, the in-vehicle time of the car has a negative mean coefficient and a standard deviation being significantly different from zero. "The number of companions" remains the same in all alternatives' utility functions and can be regarded as a kind of personal attribute. The random coefficient of "the number of companions" has a positive mean and a standard deviation being significantly different from zero. Simulation results show that all theoretically identifiable elements in variance-covariance matrices of correlated random coefficients can be estimated with a large enough sample size, while in empirical studies, the real data may not allow for estimating all these parameters due to limited sample size or little variation in explanatory variables. In these cases, some parameters may be fixed, and only parameters of interest can be estimated. The variance and correlation structures from the MMNP model provide useful insights. A greater variance for bus than metro indicates much higher variability in the quality of bus-related equipment, bus stop environment, and bus service relative to the more streamlined and less variable metro rail service. Interestingly, whenever correlations were allowed in the models, these turned out to be positive. Of the pairwise positive correlations, those between car/taxi, bus/metro, and bus/bus and metro are to be expected. However, the positive correlation between the car and metro modes, even though less than the positive correlations between car/taxi, bus/metro, and bus/bus and metro, is particularly interesting because one is a private mode, while the other is a public transportation mode. But this result, though perhaps likely to be unique to many Chinese cities, suggests that improvement in the metro level of service can be effective in attracting commuters from private cars to the metro and alleviate traffic congestion during peak hours.

Empirical results also show that the MMNP model has a better model fit than traditional discrete choice models, such as the MNL, MMNL, and MNP models, evidenced by the highest values of $\rho^2(0)$ and adjusted $\rho^2(0)$, and the smallest AIC and BIC values. And the nonnested test statistics show that the probability of erroneously selecting the MMNP model is literally zero. A model comparison could also be achieved with a synthetic population of larger size. Researchers and practitioners may not get access to large samples of available travel data in practice. To better test the performance of these models in the real world, this paper uses a small sample of real data.

Combined with analytic approximations of MVNCD functions, the MMNP model can be estimated more efficiently and quickly and avoid the high-dimensional simulation integration of using the traditional MMNL model or the mixed NL model with random coefficient to relax the response homogeneity assumption and the independence assumption of the MNL model. In summary, the MMNP model can well accommodate random taste variation and flexible covariance structure in both simulation experiments and empirical study and can be efficiently estimated by employing analytic approximations of MVNCD functions. The developed MMNP model will offer researchers and practitioners another simple and efficient way to accommodate taste heterogeneity and flexible covariance structure. In this paper, random coefficients are assumed to be normally distributed, but the possibility to develop an MMNP model based on different distributional assumptions may be explored in future research.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The authors confirm their contribution to the paper as follows. Xin Ye and Ke Wang conceptualized and designed the study; Ke Wang collected the data; Ke Wang and Xin Ye interpreted and analyzed the results; Ke Wang, Xin Ye, and Hongcheng Gan prepared the draft manuscript. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments

This research was partially supported by the general project "Study on the Mechanism of Travel Pattern Reconstruction

in Mobile Internet Environment" (no. 71671129) and "The Investigation of the Mechanism for the Impact of 'Auto/P + R' Multi-Modal Traveler Information on Mode Choice" (no. 71871143) and the key project "Research on the Theories for Modernization of Urban Transport Governance" (no. 71734004) from the National Natural Science Foundation of China.

References

[1] D. Brownstone and K. Train, "Forecasting new product penetration with flexible substitution patterns," *Journal of Econometrics*, vol. 89, no. 1-2, pp. 109–129, 1998.

[2] E. Cherchi and J. Ortuzar, "Alternative specific variables in non-linear utility functions: influence of correlation, homoscedasticity and taste variations," in *Proceedings of the Presented at the the 10th Intenational Conference on Travel Behaviour Research Luceme,* Switzerland, 2003.

[3] S. Hess and J. W. Polak, "Mixed Logit modelling of parking type choice behaviour," *Transportation Statistics*, pp. 77–102, 2009.

[4] D. Revelt and K. Train, *Customer-specific Taste Parameters and Mixed Logit (Working Paper)*, Department of Economics, University of Califomia, Berkeley, CL, USA, 2000.

[5] C. R. Bhat and J. Guo, "A mixed spatially correlated logit model: formulation and application to residential choice modeling," *Transportation Research Part B: Methodological*, vol. 38, no. 2, pp. 147–168, 2004.

[6] M. Chernew, G. Gowrisankaran, and D. P. Scanlon, "Learning and the value of information: evidence from health plan report cards," *Journal of Econometrics*, vol. 144, no. 1, pp. 156–174, 2008.

[7] S. Hess, M. Bierlaire, and J. Polak, "Capturing taste heterogeneity and correlation structure with mixed GEV models," in *Proceedings of the Infoscience. Presented at the 8th Annual Meeting of the Transportation Research Board8th Annual Meeting of the Transportation Research Board*, 2005.

[8] L. A. Garrow, *Comparison of Choice Models Representing Correlation and Random Taste Variation: An Application to Airline Passengers' Rescheduling Behavior*, Northwestern University, Evanston, IL, USA, 2004.

[9] S. R. Dalal and R. W. Klein, "A flexible class of discrete choice models," *Marketing Science*, vol. 7, no. 3, pp. 232–251, 1988.

[10] D. McFadden and K. Train, "Mixed MNL models for discrete response," *Journal of Applied Econometrics*, vol. 15, no. 5, pp. 447–470, 2000.

[11] H. Dong, E. Ben-Elia, C. Cirillo, T. Toledo, and J. N. Prashker, "On negative correlation: a comparison between Multinomial Probit and GEV-based discrete choice models," *Transportmetrica: Transportation Science*, vol. 13, no. 4, pp. 356–379, 2017.

[12] C. R. Bhat and R. Sidharthan, "A simulation evaluation of the maximum approximate composite marginal likelihood (MACML) estimator for mixed multinomial probit models," *Transportation Research Part B: Methodological*, vol. 45, no. 7, pp. 940–953, 2011.

[13] N. Eluru, V. Chakour, M. Chamberlain, and L. F. Miranda-Moreno, "Modeling vehicle operating speed on urban roads in Montreal: a panel mixed ordered probit fractional split model," *Accident Analysis & Prevention*, vol. 59, pp. 125–134, 2013.

[14] C. R. Bhat, "New matrix-based methods for the analytic evaluation of the multivariate cumulative normal distribution function," *Transportation Research Part B: Methodological*, vol. 109, pp. 238–256, 2018.

[15] K. E. Train, *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge, England, 2009.

[16] J. Walker, "Mixed logit (or logit kernel) model: dispelling misconceptions of identification," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1805, no. 1, pp. 86–98, 2002.

[17] R. Zhang, X. Ye, K. Wang, D. Li, and J. Zhu, "Development of commute mode choice model by integrating actively and passively collected travel data," *Sustainability*, vol. 11, no. 10, 2019.

[18] R. Zhang, K. Wang, J. Zhu, and X. Ye, "Development of mode choice model to understand why auto owners use rail for commute in Shanghai, China," in *Proceedings of the 17th COTA International Conference of Transportation Professionals*, 2017.

[19] C. R. Bhat, "Incorporating observed and unobserved heterogeneity in urban work travel mode choice modeling," *Transportation Science*, vol. 34, no. 2, pp. 228–238, 2000.

[20] J. Shen, Y. Sakata, and Y. Hashimoto, "A comparison between latent class model and mixed logit model for transport mode choice: evidences from two datasets of Japan," *Economics And Business*, vol. 6, 2006.

[21] C. R. Bhat, "Accommodating flexible substitution patterns in multi-dimensional choice modeling: formulation and application to travel mode and departure time choice," *Transportation Research Part B: Methodological*, vol. 32, no. 7, pp. 455–466, 1998.

[22] M. E. Ben-Akiva and S. R. Lerman, "Discrete choice analysis: theory and application to travel demand," MIT Press, Cambridge, England, 1985.

[23] K. Wang, X. Ye, and J. Ma, "An empirical analysis of post-work grocery shopping activity duration using modified accelerated failure time model to differentiate time-dependent and time-independent covariates," *PLoS One*, vol. 13, no. 11, Article ID e0207810, 2018.

[24] K. Wang and X. Ye, "Development of alternative stochastic Frontier models for estimating time-space prism vertices," *Transportation*, vol. 48, no. 2, pp. 773–807, 2021.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 5300088, 12 pages
https://doi.org/10.1155/2022/5300088

Research Article

The Car-Purchasing Intention of the Youth in the Context of Online Car-Hailing: The Extended Theory of Planned Behavior

Yihui Huang,[1] **Fei Yang**,[1] **Dong Zhang**,[2] **and Zhennan Ding**[1]

[1]*School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 611756, China*
[2]*School of Transportation and Logistics, Dalian University of Technology, Dalian 116024, China*

Correspondence should be addressed to Dong Zhang; zhangdong@dlut.edu.cn

Received 23 December 2021; Revised 14 April 2022; Accepted 4 May 2022; Published 1 June 2022

Academic Editor: Kun An

Online car-hailing services have become an integral part of people's daily travel in China. Considering that young people are the major consumer group in car purchases, it is worth investigating how the experience of online car-hailing services affects their intention to purchase a car. Based on the extended theory of Planned Behavior, this study found that the factors that negatively impact the car purchase intention of the youth are firstly the public transportation service quality, followed by the risks of private cars, and finally the online car-hailing services quality. Elevating the convenience and comfort of public transportation is conducive to reducing car purchasing intention. The indirect effect of online car-hailing services on car purchase intention is greater than its direct effect, and the most important factor is attitude. The car purchase intention is significantly heterogeneous across age and annual household income groups. Improving the convenience of public transportation will reduce the car purchase intention of people in the early youth. For middle and later youth, providing demand-responsive transit for important individuals to meet their diverse needs can reduce car purchase intention. As for online car-hailing services, youth care most about their convenience and comfort and worry most about their safety. Providing better online car-hailing services can reduce the car-purchasing intention of youth.

1. Introduction

Online car-hailing services, as a crucial form of shared mobility, are being used by an increasing number of young people in China with the popularization of online car-hailing applications such as Uber, Didi, and Gaode (AutoNavi) taxi.

There are divergences in the relationship between shared mobility and private car purchase. Some hold that shared mobility can shift individuals' dependence on traveling by private cars and reduce or postpone a car purchase [1, 2]; while some argue that shared mobility may be a compromise for individuals when they cannot afford cars, so shared mobility that appears in the transitional stage will prompt more individuals to buy private cars [3, 4]. According to China statistics, from 2016 to 2020, the number of new car registrations shows a slow downward trend from 2752 million units in 2016 to 2424 million units in 2020 [5]. At the same time, the number of people using online car-hailing services increased from 249.08 million in 2016 to 365.28 million in 2020, presenting an increase of 46% [6].

According to relevant reports, individuals aged 18 to 35 are most likely to purchase their first private car [7, 8]. The youth is at a special stage of life featured with open-mindedness and susceptibility to the influence of the external world. This period also witnesses the intensive occurrence of major life changes (such as marriage, childbirth, and migration). It has been found that major changes in life and long-term life-stage transitions impact household car ownership status [9, 10]. Youth as the core group of car buyers, their intention to purchase a car will trigger car purchase behavior. It has been found that car ownership is highly correlated with travel mode [11–13]. Then, once young people have their private cars, they are more likely to choose private cars over public transportation for travel in the future [14, 15]. However, as a populous country, China has only limited infrastructure resources. What is more, to

fulfill the objective of reducing carbon emissions, the government is prone to advocating public transportation and adjusting the travel structure. Therefore, this paper targets the young group, the major car purchasers, and explores the psychological mechanisms of their car purchase intentions in the era of online car-hailing.

In the era of online car-hailing, people regard online car-hailing as a new option in addition to traditional public transportation and private transportation. Suppose the model only considers the impact of online hailing services on car purchase intentions. In that case, the impact of online car-hailing services may be exaggerated, and the impact of improvements in other modes of transport neglected. Therefore, based on the Theory of Planned Behavior, this paper introduces the extended variables "perceived service quality of online car-hailing services," "perceived service quality of public transport," and "perceived risk of private cars" to investigate the impact of various factors on the car purchase intention.

This paper studies the factors and influencing mechanisms that affect car-purchasing intention in the era of online car-hailing services. The study has two features: one is that it is conducted among the youth, and the other is that it examines how public transportation, private transportation, and online car-hailing services affect car purchase intentions, respectively, in an environment where people can freely choose transportation modes. The rest of the paper is arranged as follows: The second section is a literature review that summarizes the previous studies; the third section constructs a model based on the Extended Theory of Planned Behavior; the fourth section displays the data collection and model modification; the fifth section reports the results of the model analysis in detail; the sixth section further concludes the study and puts forward feasible policies and measures.

2. Literature Review

Online car-hailing services, also known as carpooling, accurately match the demand and supply through the online scheduling of the platform or system [16]. Uber and Lyft Uber and Lyft, two of the most important online car-hailing service players, emerged in the United States in 2010. Subsequently, in 2016, ride-sourcing platforms such as Didi and Gaode Taxi became exceedingly popular in China. What kind of travel changes will online car-hailing services bring to Chinese people? Will it change the status of car ownership?

Most studies of shared mobility on car ownership concentrate on carsharing [17–19], and there are few studies about online car-hailing because the countries studied are mostly carsharing users. According to the US-based study by Blumenberg et al., the emergence of new technologies may reduce the dependence of some households on private cars. In contrast, for other households, it may not directly affect their decision to buy a car [20]. Wang et al. studied the situation of online car-hailing services in the United States, finding that regular and active users are more likely to own fewer cars compared to occasional users. [21]. Few studies in China have addressed the relationship between online car-hailing services and car ownership. In China, research on online car-hailing services mainly focuses on the aspect of environmental protection [22–24], safety [25, 26], and individual preferences [27].

China has different national conditions from other developed countries. In China, car ownership per capita is much lower than that of developed countries, while the cost of car purchase is much higher. China has a huge community of online car users. Online car-hailing services have become popular in China since 2016. In 2020, the users of online car-hailing services totaled 365.28 million [6]. Therefore, this paper studies the relationship between the development of online car-hailing and car purchase intention under the background of China's national conditions.

It has been found that car ownership is mainly related to individual attributes, travel demand [18, 19], car cost and performance [20], and external travel environment. The external travel environment affects car purchases from three perspectives [21]: (1) Public transportation service quality. (2) Emerging shared modes of transportation. (3) Private car travel restrictions. Few studies incorporate all the three important variables of the external travel environment into the model for consideration. This paper explores the mechanisms that influence individuals' intention to purchase a car when they are confronted with public transportation, private cars, and online car-hailing services as available modes. This study can compare the impacts of car-purchasing intention brought by different modes, which is useful for policymaking.

Previous studies on car purchase intentions mainly target all age groups, few specifically focusing on young people. Wang et al. studied how online car-hailing services affect the travel behavior of young people. The study is based on a comparison of DiDi Hitch and DiDi Express Pool, which provide sidesplitting services in Hangzhou, China [28]. Muromachi used the life-oriented approach to study the relationship between college students' previous choices of travel to school and their car purchase intentions in the future [29]. He found that the student's previous experience of going to high school by bicycle positively correlates with their future car purchasing intention. In contrast, their traveling experience via railway showed a statistically significant negative correlation with their car purchase intentions. Muromachi's study provides important insights into how young people's experiences and travel habits may influence their future car purchasing behavior. However, how do Online Car-hailing experiences and habits affect young people's future car purchase behavior? What is the mechanism of action? No clear conclusion has been drawn yet.

3. Methods

3.1. Conceptual Model. Planned Behavior is a classic structural model for studying behavior mechanisms. Ajzen proposed that the theory of Planned Behavior is a relatively open theoretical framework, and the model's explanatory power can be lifted by adding new influencing variables [27]. According to relevant studies [18] and our hypothesis, this paper adds three extended variables, namely, "Perceived Service Quality of Public Transportation," "Perceived Service

Quality of Online Car-hailing Service," and "Perceived Risk of Private Cars," to construct an extended theory of Planned Behavior (TPB).

3.2. Hypotheses. According to the TPB-related studies, behavioral intention is positively produced from a combination of attitude (ATT) toward the behavior, subjective norm (SN), and perceived behavioral control (PBC), in which subjective norm and perceived behavior control have a positive effect on attitude toward the behavior [30].

Studies have shown that as more and more people use Online car-hailing services, travelers may be less willing to buy cars [31]. According to the statistics [5, 6], the number of online car registrations negatively correlates with the number of new car registrations. Therefore, it is hypothesized that the perceived quality of Online car-hailing services can help reduce car purchase intention, and there is a negative relationship between the two. Considering that the relationship between perceived service quality of online car-hailing service and attitude, subjective norms, and perceived behavior control is not clear, they are hypothesized to be negatively correlated.

Public transportation and traveling by cars are in a competitive relationship. The higher the satisfaction with public transportation service quality, the lower the intention to buy a car [15]. Therefore, it is assumed that a high Perceived Service Quality of Public Transportation (PT) can help reduce car purchase intention. The subjective norms advocated by the society for green travel will reduce car purchase intentions. Public transportation fares and station accessibility are why travelers choose public transportation. Therefore, it is hypothesized that public transportation perception services negatively correlate with perceived behavior control. Considering that the relationship between attitude and Perceived Service Quality of Public Transportation is not clear, it is hypothesized that they are negatively correlated. Perceived risk of private cars (PR) is a factor that hinders the usage of private cars for travel and the purchase of private cars [32–36]. Attitude and perceived behavior control are related to previous travel preferences. When the living and travel environments remain unchanged, changing the familiar travel modal faces great resistance [37]. Therefore, attitude and perceived behavior control are hypothesized to negatively correlate with car purchase intentions. The relationship between the perceived risk of private cars and subjective norms is not clear, so it is hypothesized that they are negatively correlated. The hypothesized relationship in the model is shown in Table 1. The conceptual model of the young people's car purchase intention is shown in Figure 1.

Table 1: Summary of model research assumptions.

Symbol	Hypotheses
H1a	CH contributes to decreasing ATT
H1b	CH contributes to decreasing SN
H1c	CH contributes to decreasing PBC
H1d	CH contributes to decreasing PI
H2a	PT contributes to decreasing ATT
H2b	PT contributes to decreasing SN
H2c	PT contributes to decreasing PBC
H2d	PT contributes to decreasing PI
H3a	PR contributes to decreasing ATT
H3b	PR contributes to decreasing SN
H3c	PR contributes to decreasing PBC
H3d	PR contributes to decreasing PI
H4	A contributes to increasing PI
H5a	SN contributes to increasing ATT
H5b	SN contributes to increasing PI
H6a	PBC contributes to increasing ATT
H6b	PBC contributes to increasing PI

3.3. Latent Variables and Observed Variables

3.3.1. Perceived Service Quality of Online Car-Hailing (CH) Service. Perceived Service Quality of online car-hailing (CH) service denotes how travelers perceive online car-hailing services. Gilibert et al. [37] found that the biggest feature of online car-hailing service compared to traditional transportation methods is its convenience. Users are more concerned about online car-hailing services' reliability, availability, price, and time. Loa and Nurul Habib [38] researched quantitative travel characteristics (travel time, waiting time, travel expenses) and qualitative travel characteristics (weather, reliability, safety, convenience, and the availability of public transportation). He found that travelers attach more importance to comfort, reliability and safety among the qualitative characteristics. Liu et al. [25] studied the general safety perception and situational safety perception of Chinese passengers on online car-hailing services, concluding that the overall safety of online car-hailing services is significantly lower than that of traditional car renting services.

According to the above literature, five variables of economy, reliability, convenience, comfort, and safety are selected to describe CH.

3.3.2. Perceived Service Quality of Public Transportation (PT). Perceived Service Quality of Public Transportation (PT) refers to how travelers perceive the service level of public transportation. Bi et al. [39] compared people's intention of choosing between buses and online car-hailing services and found that the travel cost affects the intention of choosing online car-hailing services. The accessibility of bus stops has a significant negative impact on car-hailing behavior, and more changes in public transportation will reduce the intention of travelers to take the bus. Atombo and Dzigbordi Wemegah [40] found that affordability, availability, safety, and reliability are highly correlated with passenger satisfaction expectations when studying commuters' satisfaction with public transportation services. Björklund and Swärdh [41] estimated the value of public transport in-car comfort and reduced congestion policy, finding that all travelers tend to strongly refuse the crowding level of eight standing passengers per square meter.

Based on the above literature, five variables of economy, reliability, convenience, comfort, and safety are selected to reflect PT.

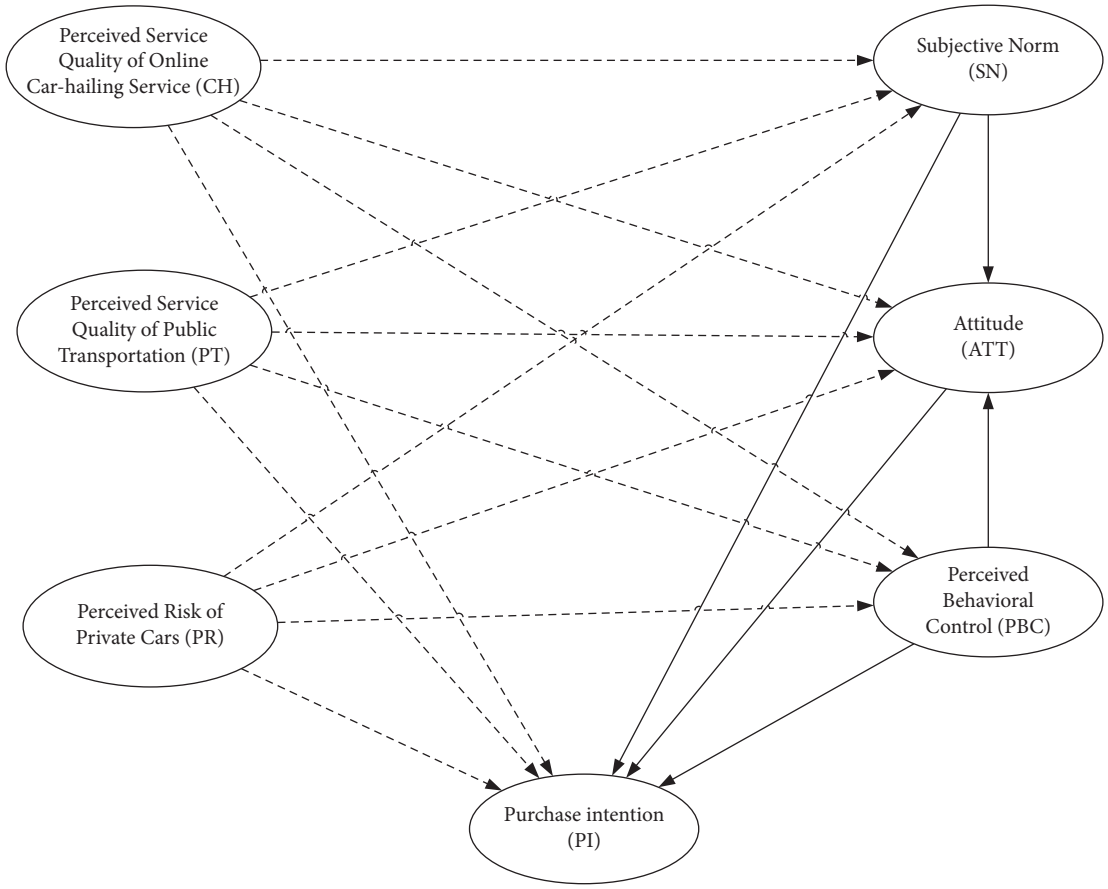


Figure 1: Conceptual model of young people's car purchase intentions (Note: The solid line represents positive correlation, and the dashed line represents negative correlation).

3.3.3. Perceived Risk of Private Cars (PR). Perceived Risk of Private Cars (PR) refers to the risks of traveling by private cars, affecting travelers' intention to use and purchase cars. Chen et al. [42] and Liu et al. [32] studied China's policy of traffic restrictions based on the last digit of license plate numbers, which can reduce congestion and affect private car ownership. Christiansen et al. [33] discovered that the use of private parking spaces or reserved parking spaces doubled the possibility of owning a car. The empirical research by Khordagui [34] shows that higher parking prices will reduce driving, and parking pricing can indeed become an effective transportation demand management tool. Jinkins [35] compared the consumption between China and the United States and pointed out that in addition to the one-time purchase of the car, car consumption also includes car insurance, car maintenance, mechanical and electrical repair, and replacement.

Two factors that may also affect car purchase are added—the risk of congestion and traffic accidents. The former is added because, during peak hours, passengers may abandon their private cars and choose buses or subways with exclusive bus lanes to ensure time reliability and avoid being late. The latter is included because individuals, when drunk, tired, or unfamiliar with the road, may choose public transportation or online car-hailing service to reduce the risk of accidents.

PR is measured by five variables: parking difficulty, traffic restriction risk, continuous expenditure risk, congestion risk, and traffic accident risk.

3.3.4. Attitude (ATT). Attitude (ATT) is an individual's evaluation of car purchasing behavior.

The evaluation consists of two parts: belief strength (strength) and result evaluation (evaluation), among which the strength refers to the degree to which people expect behavior results to be achieved, and the evaluation is people's expectation of the result regarding whether it is positive or negative [43]. Besides the strength and evaluation, the attitudes also include youth preferences for a car purchase. Belgiawan et al. [44] found that the preference for travel behavior in the past will affect future decisions on car purchases.

Attitude is measured with three variables: belief intensity, result evaluation, and preference.

3.3.5. Subjective Norms (SN). Subjective norms (SN) represent the perceived social pressure of car purchasing behavior.

Social pressure comes from important individuals or groups, such as relatives, friends, classmates, or the public [43]. The important individuals that affect the car purchase intention of the youth include their friends and classmates (peer pressure) and family members. Car purchase intention

of the youth is also related to future family activities [9, 45], which can be summarized as "motivation to take care of a family."

SN is described by three variables: peer pressure, family guidance, and motivation to take care of a family.

3.3.6. Perceived Behavior Control (PBC). Perceived Behavior Control (PBC) refers to the degree of ease/difficulty with which an individual believes that he or she can control and perform a given behavior (car purchasing behavior) [43].

Young people's perception of the influencing factors of car purchase mainly includes three aspects: perception of car cost (car purchase cost and subsequent maintenance cost); perception of the ease/difficulty of obtaining license plates; self-control (Whether it can be decided by oneself) [46].

PBC is described by three variables: car cost perception, the ease/difficulty of obtaining license plates, and self-control.

3.3.7. Car Purchase Intention (PI). Behavioral intention is the intention prior to performing a specific act. The car purchase intention of the youth is affected by their economic conditions and the external traffic environment, among which the external traffic environment includes the public transportation environment [39] and the online car-hailing environment [31, 47].

The latent variables and observed variables are summarized in Table 2.

The above is the main body of the questionnaire used to collect empirical data for the 27 observed variables. The scale for the observed variables collection was the Likert 5 level scale, which used different scores to indicate individuals' subjective orientation. A score of 5 indicates strongly agree, 4 indicates agree, 3 indicates fair, 2 indicates disagree, and 1 indicates strongly disagree.

4. Data and Results

4.1. Survey Design

4.1.1. Basic Information of the Questionnaire. A questionnaire was designed to collect data for model estimation—both the sociodemographics of the respondents and their response to the previous measurement questions constituted the questionnaire. The survey was taken in Chengdu, Sichuan Province, China, from July 11 to 18, 2019. The specific locations of the survey were universities, commercial areas, parks, and hubs in the Chengdu metropolitan area. A total of 590 sample electronic and paper questionnaires were collected, among which 451 were valid. The sample distribution in Table 3 is consistent with the actual social and economic conditions of the youth in Chengdu, indicating that the survey data are demographically representative.

4.1.2. Age Classification in the Questionnaire. Does age affect the intention to purchase a car? This paper attempts to explain the factors influencing car-purchasing intention at different ages. Age was divided into three categories for sample collection. Young people in the early youth years (18–22 years old), the middle youth years (23–28 years old), and later youth years (28–35 years old). The sample sizes are 131, 179, and 141, respectively. The reasons for this classification are as follows.

The youth group aged 18–22 is mainly a student group. Their dormitories, cafeterias and classrooms are located inside the campus, and they do not need to travel outside the campus frequently, so they often do not have a strong desire to purchase a car. The level of transportation services affects their perceptions. This stage is the "perception stage."

The youth group aged between 22 and 28 are in a transitional stage of life; some may be continuing their education, and some are just entering the workforce. One thing that young people at this stage often have in common is the "lack of funds." They lack the necessary funds to buy a car and the ability to repay the car loan. This stage is the "Preparatory stage."

The youth group aged 28–35 years old in China typically faces two important life events, marriage and childbirth. Relevant studies have shown that changes in family structure affect car ownership. As the number of family members increases and the variety of transportation events, there is a necessity for private car purchases. This stage is the "implementation stage."

4.2. Model Test and Modification

4.2.1. Reliability and Validity. Reliability refers to the degree of stability and reliability of the questionnaire measurement results. Cronbach's α, CICT (Corrected Item-Total Correlation), two indicators, are used to determine the reliability. Cronbach's α is an indicator describing internal consistency, and most studies consider a value of Cronbach's alpha above 0.8 to be acceptable [48–50]. CICT refers to the Pearson correlation coefficient of each entry aggregated with other entries. In general, if this indicator is less than 0.3, we consider that the entry is not strongly correlated with other entries and can be excluded [48–50].

Validity is an indication of the accuracy and validity of the measurement results. The structural validity was tested by factor analysis with KMO (Kaiser–Meyer–Olkin) values, factor loadings, and cumulative explained variance of variables. If the KMO value is below 0.7, it is not suitable as an analysis factor. The factor loadings represent the correlation between the observed and latent variables. The lowest value should not be less than 0.5. Otherwise, it is regarded as an insignificant correlation. The minimum cumulative variance should not be less than 50% [48–50].

SPSS 24.0 software was used to test the reliability and validity of the questionnaire data, and the results are shown in Table 4.

The overall Cronbach's α coefficient is 0.820, Cronbach's α is around 0.8, and the CITC is greater than 0.3, indicating that the reliability test is qualified. KMO (Kaiser–Meyer–Olkin) value, factor loadings, and cumulative variance are all within a reasonable range.

Table 2: Statements used to measure the latent constructs.

Latent variable	Symbol	Measurable variable	Symbol	Statements
Perceived service quality of online car-hailing service (CH)	CH	Economy	CH1	With reasonable charging, traffic costs can be accepted.
		Reliability	CH2	Online car-hailing services usually meet my time expectations.
		Convenience	CH3	Feel convenient when using online car-hailing.
		Comfort	CH4	Feel comfortable when using online car-hailing. (Automotive interior/Driver's quality)
		Safety	CH5	Feel safe when using online car-hailing. (Driving safety/Personal and property safety.)
Perceived service quality of public transportation	PT	Economy	PT1	With reasonable charging, traffic costs can be accepted.
		Reliability	PT2	Public transport usually meets my time expectations. (Waiting time/Travel time)
		Convenience	PT3	Usually feel convenient when using public transport. (Stations covering/Transfer)
		Comfort	PT4	Usually feel comfortable when using public transport. (Automotive interior/Seats/Crowding)
		Safety	PT5	Usually feel safe when using public transport. (Driving safety/Personal and property safety)
Perceived risk of private cars	PR	Parking difficulty	PR1	I worry that the shortage of parking spaces in the city will increase my parking time and cost.
		Traffic restriction	PR2	I worry that the policy of driving and traffic restrictions will increase the additional fine or the idle rate of the car.
		Continuous expenditure	PR3	I worry that fuel, insurance, maintenance, and repair will cost me more after buying my car.
		Congestion risk	PR4	I worry that the road congestion will continue to worsen after purchasing a car, making car travel more inconvenient.
		Traffic accident	PR5	I worry that traffic accidents will threaten the life and property safety of myself and my family.
Attitude	ATT	Preference	ATT1	I think buying a car will make my life more convenient and free.
		Belief intensity	ATT2	I think cars are an indispensable part of life in the future.
		Results evaluation	ATT3	I think buying a car is a wise decision.
Subjective norm	SN	Peer pressure	SN1	All my peers and colleagues have cars.
		Family guidance	SN2	My family encouraged me to buy a car.
		Motivation to take care of a family	SN3	I would consider buying a car for taking care of my family's needs (medical appointments, taking my children to school)
Perceived behavioral control	PBC	Perception of the car cost	PBC1	I fully understand the cost (purchase/insurance/maintenance) of the car.
		The ease/difficulty of obtaining license plates	PBC2	Buying a car within five years of working is easy for me.
		Self-control	PBC3	Whether or not I need to buy a car for now or future is entirely up to me.
Purchase intention	PI	Purchase a car when economic conditions are met	PI1	I will buy a car when the money allows.
		Purchase a car when the shared mobility services are well-developed	PI2	Even though shared mobility services are well-developed in the future, I will still buy a car.
		Purchase a car with better public transportation	PI3	Even though public traffic services are better in the future, I will still buy a car.

4.2.2. Check Path Significance. AMOS 24.0 was used to analyze the effects between paths in the hypothetical model. The basic fitness level mainly considers the significance of the path coefficient between latent variables from the Standard Error of Estimate (S.E.) and Critical Ratio (C.R.) to test the rationality of the hypothesis. The critical ratio is the ratio of the parameter estimate to the standard error. The path significance test results are shown in Table 5.

Table 3: Sociodemographic statistics of the youth group sample.

Character	Value	Ratio (%)	Character	State	Ratio (%)
Gender	Male	51.6	Student	33.7	
	Female	48.4	Enterprise or company employee	42.9	
Age	18–22	28.7	Occupation	Civil servant	5.9
	23–28	39.2		Individual business	9.0
	29–35	32.2		Freelancer	3.9
Personal monthly income	Under ¥2000	30.4		Others	4.6
	¥2001–5000	31.9		Under 100000	32.6
	¥5001–8000	17.9	Annual household Income (CNY)	100000–150000	32.8
	¥8001–12000	14.2		150000–200000	22.5
	Above ¥12000	5.5		200000–300000	8.8
Education	Junior high or below	2.8		Above 300000	3.3
	High school	6.1		1	13.3
	Junior college	12.9	Family members	2	23.4
	Undergraduate	49.9		3	51.2
	Postgraduate or above<	28.2		4	9.6
Marriage and childbirth status	Unmarried	64.8		Over 5	2.4
	Married without child	22.8			
	Married with child	12.5			

Table 4: Reliability and validity results of latent variables.

Latent variable	Observed variable	CITC	Cronbach's α	Factor loadings	KMO	Cumulative variance (%)
CH	CS1	0.588	0.798	0.753	0.807	55.74
	CS2	0.600		0.764		
	CS3	0.601		0.767		
	CS4	0.591		0.748		
	CS5	0.532		0.699		
PT	PT1	0.487	0.791	0.659	0.777	54.65
	PT2	0.572		0.745		
	PT3	0.653		0.806		
	PT4	0.628		0.789		
	PT5	0.512		0.685		
PR	PR1	0.601	0.811	0.769	0.833	58.22
	PR2	0.652		0.799		
	PR3	0.624		0.784		
	PR4	0.697		0.831		
	PR5	0.456		0.614		
ATT	ATT1	0.652	0.841	0.837	0.715	76.07
	ATT2	0.741		0.891		
	ATT3	0.738		0.888		
SN	SN1	0.661	0.789	0.859	0.702	70.47
	SN2	0.634		0.843		
	SN3	0.596		0.816		
PB	PBC1	0.631	0.794	0.838	0.71	70.97
	PBC2	0.643		0.846		
	PBC3	0.641		0.844		
PI	PI1	0.708	0.896	0.859	0.705	82.80
	PI2	0.825		0.925		
	PI3	0.863		0.943		

Table 5: Significance results of structural model path coefficient.

Latent variable path			Estimate	S.E.	C.R.	P	Significance
SN	←	CH	−0.185	0.078	−2.382	0.017	*
SN	←	PT	−0.295	0.073	−4.030	<0.001	***
SN	←	PR	−0.206	0.077	−2.687	0.007	**
PBC	←	CH	−0.079	0.054	−1.464	0.143	—
PBC	←	PT	−0.106	0.051	−2.078	0.040	*
PBC	←	PR	−0.200	0.079	−2.523	0.012	*
ATT	←	CH	−0.187	0.08	−2.324	0.020	*
ATT	←	PT	−0.211	0.075	−2.820	0.005	**
ATT	←	PR	−0.157	0.054	−2.924	0.003	**
ATT	←	SN	0.438	0.058	7.540	<0.001	***
ATT	←	PBC	0.236	0.048	4.900	<0.001	***
PI	←	CH	−0.102	0.050	−2.036	0.042	*
PI	←	PT	−0.126	0.048	−2.634	0.008	**
PI	←	PR	−0.135	0.051	−2.657	0.008	**
PI	←	ATT	0.431	0.103	4.173	<0.001	***
PI	←	SN	0.162	0.068	2.378	0.017	*
PI	←	PBC	0.261	0.052	5.055	<0.001	***

Note. ***indicates $P < 0.001$, **indicates $P < 0.01$, *indicates $P < 0.1$, — indicates not significant..

As the Critical Ratio absolute value of the path "PBC←CH" is lower than 1.96, which does not meet the significance requirement, this path is excluded.

4.2.3. Check Modification Indices. In order to improve the overall fitness of the model, the Modification Indices (M.I.) are used to make appropriate modifications to the model. When M.I. >3.84, it generally means that we need to free the constraints between variables (to establish a covariance) [50]. The M.I. is shown in Table 6. It is required to establish covariances between e4 and e5 (the error terms of the observed variables of CH), e6 and e7 (the error terms of the observed variables of PT), as well as e26 and e27 (the error terms of the observed variables of the PI).

4.2.4. The Revised Model. The revised model is obtained after deleting the "CH⟶PBC" path and establishing the covariation relationship. The significance test results of the revised model are shown in Table 7. All paths pass the significance test. The factor loadings of the measurement model are shown in Table 8. The factor loadings of the observed variables are all between 0.5 and 0.95 [50], implying that the basic fitness of the measurement model is good. The fitness indices before and after the modification are compared, as shown in Table 9.

5. Discussion

5.1. Path Analysis and Factor Analysis. This section discusses the factors that influence the Car purchase intention based on path relationships and explains the possible reasons.

The direct and indirect effects of latent variables are shown in Table 10. The influence of observed variables on latent variables is shown in Table 8.

Table 6: Modification index.

Associated path			M.I.	Par change
e5	←⟶	e4	18.526	0.08
e26	←⟶	e27	14.24	0.031
e7	←⟶	e6	6.97	0.044

Note. Par Change: indicates the estimated parameter change amount of the forecast.

Table 7: Significance results of structural model path coefficient.

Latent variable path			Estimate	S.E.	C.R.	P	Significance
SN	←	CH	−0.208	0.077	−2.699	0.007	**
SN	←	PT	−0.311	0.076	−4.083	<0.001	***
SN	←	PR	−0.193	0.077	−2.503	0.012	*
PBC	←	PT	−0.11	0.052	−2.115	0.036	*
PBC	←	PR	−0.187	0.081	−2.319	0.02	*
ATT	←	CH	−0.21	0.081	−2.577	0.01	*
ATT	←	PT	−0.23	0.079	−2.928	0.003	**
ATT	←	PR	0.451	0.058	7.71	<0.001	***
ATT	←	SN	0.243	0.048	5.071	<0.001	***
ATT	←	PBC	−0.179	0.052	−3.466	<0.001	***
PI	←	CH	0.156	0.072	2.158	0.031	*
PI	←	PT	0.288	0.053	5.405	<0.001	***
PI	←	PR	0.488	0.108	4.511	<0.001	***
PI	←	ATT	−0.135	0.052	−2.607	0.009	**
PI	←	SN	−0.155	0.052	−2.992	0.003	**
PI	←	PBC	−0.139	0.053	−2.6	0.009	**

Table 8: Test results of a load of measurement model factors.

Latent variable	Observed variable path	Factor loadings
CH	CH1-CH	0.790
	CH2-CH	0.828
	CH3-CH	0.841
	CH4-CH	0.780
	CH5-CH	0.707
PT	PT1-PT	0.502
	PT2-PT	0.643
	PT3-PT	0.765
	PT4-PT	0.744
	PT5-PT	0.586
PR	PR1-PR	0.838
	PR2-PR	0.827
	PR3-PR	0.803
	PR4-PR	0.855
	PR5-PR	0.694
ATT	ATT1-ATT	0.726
	ATT2-ATT	0.840
	ATT3-ATT	0.813
SN	SN1-SN	0.769
	SN2-SN	0.744
	SN3-SN	0.743
PBC	PBC1-PBC	0.742
	PBC2-PBC	0.713
	PBC3-PBC	0.759
PI	PI1-PI	0.806
	PI2-PI	0.827
	PI3-PI	0.871

Table 9: Comparison indicators between the initial model and the modified model.

	CMIN/DF	RMSEA	GFI	NFI	CFI	TLI
Superior standard	<3	<0.05	>0.9	>0.9	>0.9	>0.9
Initial model	3.164	0.069	0.872	0.912	0.911	0.898
Revised model	2.895	0.058	0.884	0.923	0.923	0.911
Fit judgment	Superior	Accept	Accept	Superior	Superior	Superior

Note. CMIN/DF (chi-square degree of freedom ratio); RMSEA (root-mean-square); GFI (goodness of fit); CFI (comparative fit index); NFI (normed fit index); TLI (Tucker–Lewis index).

Table 10: Standardized direct and indirect effects between latent variables.

Path			Direct effect	Indirect effect	Total effect
SN	←	CH	−0.160	—	−0.160
SN	←	PT	−0.362	—	−0.362
SN	←	PR	−0.235	—	−0.235
PBC	←	PT	−0.126	—	−0.126
PBC	←	PR	−0.232	—	−0.232
ATT	←	CH	−0.164	−0.077	−0.241
ATT	←	PT	−0.273	−0.205	−0.478
ATT	←	PR	−0.231	−0.171	−0.402
ATT	←	SN	0.479	—	0.479
ATT	←	PBC	0.253	—	0.253
PI	←	CH	−0.101	−0.119	−0.220
PI	←	PT	−0.159	−0.274	−0.433
PI	←	PR	−0.149	−0.252	−0.401
PI	←	ATT	0.404	—	0.404
PI	←	SN	0.137	0.194	0.331
PI	←	PBC	0.249	0.102	0.351

According to Tables 8 and 10, the factors influencing the car purchase intention of the youth could be discussed as follows.

(1) Among the positive effects of car purchase intention, ATT (0.404) > PBC (0.351) > SN (0.331), indicating that attitude is the strongest factor influencing car purchase intention. In attitude, the strength of the belief that "the car is an indispensable part of the future" has the greatest influence on car purchase intention. This consensus may be related to China's social background, as it is widely believed that "owning a house and a car is a symbol of success."

(2) Among the negative effects of car purchase intention, PT (−0.433) > PR (−0.401) > CH (−0.22), indicating that PT and PR are the major factors that affect individual purchase intentions. CH has a weak effect since the public may only use online car-hailing services to supplement public transportation or as an alternative in the absence of private car accessibility, which aligns with the relevant findings [34]. The key to reducing car purchase intention is to improve PT, especially its convenience and comfort.

(3) The level of external transportation service affects car purchase intention through psychological changes. For the effects of CH, PT, and PR on car purchase intention, the direct effect values are lower than the indirect effect values. It indicates that CH, PT, and PR suppress car purchase intention by affecting young people's ATT, SN, and PBC in most cases. (4) CH means perceived service quality of online car-hailing service. From Table 8, Convenience (0.841) > Reliability (0.828) > Economy (0.790) > Comfort (0.780) > Safety (0.707). Individuals are most satisfied with the convenience of online car-hailing services. Individuals' satisfaction with reliability stems from the fact that they can directly see the travel time on Apps, enabling them to schedule the appropriate departure time according to their needs without fear of being late. Individuals are most dissatisfied with safety, so improving the safety can prompt people to choose online car-hailing services.

5.2. Multi-Group Analysis of the Youth. We classify the features of individual attributes and perform a multi-group analysis to understand the moderating effect of attribute features. If there is a significant difference in the estimation results of the correlation coefficient, it means that the attribute variable has a moderating effect on the model [50]. Using Multiple-Group Analysis in AMOS 24.0, the analysis results are shown in Table 11.

The results in Table 11 show that the path coefficients are significant for the age groups and annual household income groups at the 0.05 level. It indicates that only age and household income significantly moderate the model among the personal attribute variables.

This paper focuses on the analysis of the age groups. The path coefficient analysis for the age group is presented in Table 12. Path correlation analysis is useful for finding out the key variables that influence the intention to purchase a car for different age groups.

5.2.1. Early Youth Years (18–22 Years Old). Young people in the early youth years (18–22 years old) are most negatively affected by PT. In China, the 18- to 22-year-old youth group mainly studies in universities with limited income. They are more willing to take public transportation. Combined with Table 8, it can be found that improving the convenience and comfort of public transportation will help young people to choose public transportation and thus reduce their intention to purchase a car.

Early youth is most positively influenced by attitudes (ATT). Attitudes can be changed by guiding perceptions. For example, promote the idea that "access is more important than ownership" and dispel the illusion that "owning a car equals a successful life."

Table 11: Results of multiple-group analysis.

Model	Gender	Age	Personal monthly income	Education	Occupation	Annual household income	Family members
CMIN/DF	1.26	1.63	1.08	1.23	1.09	1.22	1.95
P value	0.134	<0.001	0.293	0.084	0.324	0.166	<0.001

Note. $P < 0.001$ means significant.

Table 12: Path coefficient analysis of car purchase intention under different youth years.

Path	Early youth years	Middle youth years	Later youth years
PI ← CH	−0.18	−0.20	−0.34**
PI ← PT	−0.51*	−0.34	−0.44
PI ← PR	−0.24	−0.52*	−0.39**
PI ← ATT	0.41*	0.31	0.40
PI ← SN	0.30*	0.35**	0.41*
PI ← PBC	0.25*	0.24**	0.34**

Note. **indicates $p < 0.01$, *indicates $p < 0.01$.

5.2.2. Middle Youth Years (23–28 Years Old). Young people in the middle youth (23–28 years old) are most negatively affected by PR. Combined with the results of Table 8, the risk of traffic congestion is their biggest concern, followed by the inconvenience of parking and then the traffic restriction policy. Limiting car purchases requires raising the perceived risk of private cars, and in addition to effective means of traffic restrictions, regulating parking resources through price leverage is an effective and viable means.

Subjective norms are most positively affected by middle youth car purchase intention (SN). SN represents the social pressure of car purchasing behavior, which comes from important individuals. If the transportation needs of important individuals can be better met, the car-purchasing intention may be reduced from the perspective of SN.

5.2.3. Later Youth Years (29–35 years Old). Young people in the later youth years (29–35 years old) are most negatively affected by CH. It may be because 29- to 35-year-olds are more financially independent and have more opportunities to use online car-hailing services. Combined with the results of Table 8, it can thus be concluded that further improvement in the reliability and convenience of online car-hailing services will reduce the youth's willingness to purchase a car. Significant improvements in reliability and convenience rely on more advanced platform algorithms and technological innovations.

There was a significant positive effect on both PBC and SN. The effect of SN was on important individuals. The influence of important individuals is greater in later youth than in middle youth years. The reason for this is that after the age of 28, in China, young people face more changes in family structure, such as marriage and children, and important individuals from the family may have a stronger intention to purchase a car, thus enhancing the individual's willingness to purchase a car. The stronger willingness of family members to purchase a car may be because some of their travel needs cannot be met by public transportation and online car services. Perceived behavioral control was more pronounced in later youth than in early and middle youth. As youth grow older, they are more financially independent and have more control over purchasing a car.

Table 13: Abbreviations.

Term	Abbrev.
Theory of planned behavior	TPB
Perceived service quality of online car-hailing service	CH
Perceived service quality of public transportation	PT
Perceived risk of private cars	PR
Attitude	ATT
Subjective norms	SN
Perceived behavior control	PBC
Car purchase intention	PI
Standard error of estimate	S.E.
Critical ratio	C.R.
Modification indices	M.I.

6. Conclusion

The present study introduces "Perceived Service Quality of Online Car-hailing (CH)," "Perceived Service Quality of Public Transportation (PT)," and "Perceived Risk of Private Cars (PR)" to construct an extended theory of Planned Behavior and studies the influence mechanism of online car-hailing service on Chinese youth groups' car purchase intention. The results show that PT has the greatest impact on the car purchase intention of the youth. Online car-hailing service is mainly leveraged to supplement public transportation or as a substitute for private car travel. The perception of the external traffic environment (CH, PT, PR) mainly affects the youth's car purchase intention through internal psychological factors (attitude, subjective norms, perceived behavior control), of which attitude is the most important one. There are also differences in the factors that influence the willingness to purchase a car for young people of different ages.

Based on the research and analysis in the article, we offer the following suggestions to reduce car purchase intention:

(1) For the young people in the early youth years (18–22 years old), improved accessibility to public transportation will significantly improve convenience. Accessibility can be improved by establishing subway stations or bus stops at the entrances of colleges, as Muromachi has also argued [29]. Early youth is in a great habit-forming stage. By having easier access to subways and buses, college students may develop a preference for public transportation travel, greatly reducing the willingness to purchase a private car.

(2) SN (Subjective norms) significantly impacted both mid and late youth. SN comes from important

individuals, such as family members. And, our study also found that the number of family members significantly affects the car purchasing intention. Therefore, satisfying the needs of family members will have a positive effect on reducing car purchase intention. The proposal is to provide demand-responsive transit service, which can better meet the various travel needs of families. For example, it could provide transport to and from school for children in the same neighborhood, escort older people in the same neighborhood to a neighboring hospital for regular medical care, and provide transport from residential areas to rail stations during peak hours. It is worth studying exactly what kind of services is more attractive to family members. In conclusion, we hope demand-responsive transit offers a more diverse range of services in terms of flexibility, convenience, and cost-efficiency.

(3) One of the keys to improving the attractiveness of online car-hailing services is to improve safety. Safety includes two categories: driver's driving behavior and security from not being the victim of rape and homicide. Safe driving by drivers can be ensured by raising the barrier to entry to become an online driver and requiring a longer and safer driving record to ensure safe driving behavior. The security of individuals and property can be achieved through better technical means, such as recording the car, real-time positioning information of the online car, and a key alarm.

Much work could be further conducted to investigate how online car-hailing services would influence the car purchasing intention of the youth cohort. For example, the present study has found car purchasing intention heterogeneity across age groups. It would be interesting to study how such intention will change as the family lifecycle stage changes from a long-time perspective in an era of shared mobility.

The abbreviations in the text are summarized in Table 13.

Data Availability

Data are available on request to the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by China's National Nature Science Foundation (grant no. 52072313).

References

[1] A. Kolleck, "Does car-sharing reduce car ownership? Empirical evidence from Germany," *Sustainability*, vol. 13, no. 13, Article ID 7384, 2021.

[2] Y. Hui, Y. Wang, Q. Sun, and L. Tang, "The impact of car-sharing on the willingness to postpone a car purchase: a case study in Hangzhou, China," *Journal of Advanced Transportation*, vol. 2019, Article ID 9348496, 11 pages, 2019.

[3] J. Schlüter and J. Weyer, "Car sharing as a means to raise acceptance of electric vehicles: an empirical study on regime change in automobility," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 60, pp. 185–201, 2019.

[4] Y. Hui, W. Wang, M. Ding, and Y. Liu, "Behavior patterns of long-term car-sharing users in China," *Transportation Research Procedia*, vol. 25, pp. 4662–4678, 2017.

[5] MPS: 2022, https://www.mps.gov.cn/n6557563/c7647931/content.html.

[6] CAC: 2022, http://www.cac.gov.cn/2021-02/03/c_1613923423079314.htm.

[7] Jiemain, "Insights into potential car buyers in 2018," 2021, https://www.jiemian.com/article/2757593.html.

[8] Zhuanlan, "Aurora Big Data: User Portrait Data Report of Potential Automotive Consumers," 2021, https://zhuanlan.zhihu.com/p/43805264.

[9] N. A. Khan and M. A. Habib, "When do households make vehicle transactions? What type of vehicle households choose after transaction events?" *Travel Behaviour and Society*, vol. 22, pp. 32–47, 2021.

[10] B. Herrenkind, I. Nastjuk, A. B. Brendel, S. Trang, and L. M. Kolbe, "Young people's travel behavior-using the life-oriented approach to understand the acceptance of autonomous driving," *Transportation Research Part D: Transport and Environment*, vol. 74, pp. 214–233, 2019.

[11] N. Commins and A. Nolan, "Car Ownership and Mode of Transport to Work in Ireland," The Economic and Social Research Institute (ESRI), Dublin, Ireland, ESRI Working Paper No. 310, 2009.

[12] A. Nobile, C. R. Bhat, and E. I. Pas, "A random-effects multinomial probit model of car ownership choice," in *Case Studies in Bayesian Statistics*, C. Gatsonis, J. S. Hodges, R. E. Kass, R. McCulloch, P. Rossi, and N. D. Singpurwalla, Eds., vol. 121, pp. 419–434, Springer New York, New York, NY, USA, 1997.

[13] C. Ding, D. Wang, C. Liu, Y. Zhang, and J. Yang, "Exploring the influence of built environment on travel mode choice considering the mediating effects of car ownership and travel distance - ScienceDirect," 2021, https://www.sciencedirect.com/science/article/abs/pii/S0965856416303391.

[14] N. Paulley, R. Balcombe, R. Mackett et al., "The demand for public transport: the effects of fares, quality of service, income and car ownership," *Transport Policy*, vol. 13, no. 4, pp. 295–306, 2006.

[15] D. Yao, L. Xu, C. Zhang, and J. Li, "Revisiting the interactions between bus service quality, car ownership and mode use: a case study in Changzhou, China," *Transportation Research Part A: Policy and Practice*, vol. 154, pp. 329–344, 2021.

[16] S. Huang, "The rise of online ride-hailing services and their impact on urban transport in China," *Handbook on Transport and Urban Transformation in China*, Edward Elgar, vol. 451, pp. 212–228, Cheltenham, UK, 2020.

[17] P. Jochem, D. Frankenhauser, L. Ewald, A. Ensslen, and H. Fromm, "Does free-floating carsharing reduce private vehicle ownership? The case of SHARE NOW in European cities," *Transportation Research Part A: Policy and Practice*, vol. 141, pp. 373–395, 2020.

[18] F. Giesel and C. Nobis, "The impact of carsharing on car ownership in German cities," *Transportation Research Procedia*, vol. 19, pp. 215–224, 2016.

[19] M. Golalikhani, B. B. Oliveira, M. A. Carravilla, J. F. Oliveira, and A. P. Antunes, "Carsharing: a review of academic literature and business practices toward an integrated decision-support framework," *Transportation Research Part E: Logistics and Transportation Review*, vol. 149, Article ID 102280, 2021.

[20] E. Blumenberg, J. Paul, and G. Pierce, "Travel in the digital age: vehicle ownership and technology-facilitated accessibility," *Transport Policy*, vol. 103, pp. 86–94, 2021.

[21] Y. Wang, W. Shi, and Z. Chen, "Impact of ride-hailing usage on vehicle ownership in the United States," *Transportation Research Part D: Transport and Environment*, vol. 101, Article ID 103085, 2021.

[22] Y. Sui, H. Zhang, X. Song et al., "GPS data in urban online ride-hailing: a comparative analysis on fuel consumption and emissions," *Journal of Cleaner Production*, vol. 227, pp. 495–505, 2019.

[23] T. Wu, S. Wang, L. Wang, and X. Tang, "Contribution of China's online car-hailing services to its 2050 carbon target: energy consumption assessment based on the GCAM-SE model," *Energy Policy*, vol. 160, Article ID 112714, 2022.

[24] J. Gao, S. Ma, L. Li, J. Zuo, and H. Du, "Does travel closer to TOD have lower CO2 emissions? Evidence from ride-hailing in Chengdu, China," *Journal of Environmental Management*, vol. 308, Article ID 114636, 2022.

[25] Y. Liu, Q. Gao, and P.-L. P. Rau, "Chinese passengers' security perceptions of ride-hailing services: an integrated approach combining general and situational perspectives," *Travel Behaviour and Society*, vol. 26, pp. 250–269, 2022.

[26] L. Ma, X. Zhang, X. Ding, and G. Wang, "Risk perception and intention to discontinue use of ride-hailing services in China: taking the example of DiDi Chuxing," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 66, pp. 459–470, 2019.

[27] X. Hu, W. Lin, J. Wang, and J. Jiang, "Choice of ride-hailing or traditional taxi services: from travelers' perspectives," *Research in Transportation Business & Management*, Article ID 100788, 2022.

[28] Z. Wang, X. Chen, and X. M. Chen, "Ridesplitting is shaping young people's travel behavior: evidence from comparative survey via ride-sourcing platform," *Transportation Research Part D: Transport and Environment*, vol. 75, pp. 57–71, 2019.

[29] Y. Muromachi, "Experiences of past school travel modes by university students and their intention of future car purchase," *Transportation Research Part A: Policy and Practice*, vol. 104, pp. 209–220, 2017.

[30] W. T. Duan and G. R. Jiang, "A review of the theory of planned behavior," *Advances in Psychological Science*, vol. 2, pp. 315–320.

[31] K. Shi, R. Shao, J. De Vos, L. Cheng, and F. Witlox, "The influence of ride-hailing on travel frequency and mode choice," *Transportation Research Part D: Transport and Environment*, vol. 101, Article ID 103125, 2021.

[32] Z. Liu, R. Li, X. Wang, and P. Shang, "Effects of vehicle restriction policies: analysis using license plate recognition data in Langfang, China," *Transportation Research Part A: Policy and Practice*, vol. 118, pp. 89–103, 2018.

[33] P. Christiansen, N. Fearnley, J. U. Hanssen, and K. Skollerud, "Household parking facilities: relationship to travel behaviour and car ownership," *Transportation Research Procedia*, vol. 25, pp. 4185–4195, 2017.

[34] N. Khordagui, "Parking prices and the decision to drive to work: evidence from California," *Transportation Research Part A: Policy and Practice*, vol. 130, pp. 479–495, 2019.

[35] D. Jinkins, "Conspicuous consumption in the United States and China," *Journal of Economic Behavior & Organization*, vol. 127, pp. 115–132, 2016.

[36] P. Haggar, L. Whitmarsh, and S. M. Skippon, "Habit discontinuity and student travel mode choice," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 64, pp. 1–13, 2019.

[37] M. Gilibert, I. Ribas, N. Maslekar, C. Rosen, and A. Siebeneich, "Mapping of service deployment use cases and user requirements for an on-demand shared ride-hailing service: MOIA test service case study," *Case Studies on Transport Policy*, vol. 7, no. 3, pp. 598–606, 2019.

[38] P. Loa and K. Nurul Habib, "Examining the influence of attitudinal factors on the use of ride-hailing services in Toronto," *Transportation Research Part A: Policy and Practice*, vol. 146, pp. 13–28, 2021.

[39] H. Bi, Z. Ye, L. Hu, and H. Zhu, "Why they don't choose bus service? Understanding special online car-hailing behavior near bus stops," *Transport Policy*, vol. 114, pp. 280–297, 2021.

[40] C. Atombo and T. Dzigbordi Wemegah, "Indicators for commuter's satisfaction and usage of high occupancy public bus transport service in Ghana," *Transportation Research Interdisciplinary Perspectives*, vol. 11, Article ID 100458, 2021.

[41] G. Björklund and J.-E. Swärdh, "Estimating policy values for in-vehicle comfort and crowding reduction in local public transport☆," *Transportation Research Part A: Policy and Practice*, vol. 106, pp. 453–472, 2017.

[42] D. Chen, Y. Sun, and Z. Yang, "Optimization of the travel ban scheme of cars based on the spatial distribution of the last digit of license plates," *Transport Policy*, vol. 94, pp. 43–53, 2020.

[43] I. Ajzen, "The theory of planned behavior," *Organizational Behavior and Human Decision Processes*, vol. 50, no. 2, pp. 179–211, 1991.

[44] P. F. Belgiawan, J.-D. Schmöcker, and S. Fujii, "Understanding car ownership motivations among Indonesian students," *International Journal of Sustainable Transportation*, vol. 10, no. 4, pp. 295–307, 2016.

[45] E. M. F. van Sluijs, U. Ekelund, B. H. Hansen et al., "Family car ownership and activity in young people: cross-sectional and longitudinal analyses using the International Children's Accelerometry Database," *The Lancet*, vol. 392, p. S89, 2018.

[46] R. Luke, "Car ownership perceptions and intentions amongst South African students," *Journal of Transport Geography*, vol. 66, pp. 135–143, 2018.

[47] T. Lyu, P. S. Wang, Y. Gao, Y. Wang, and Y. Wang, "Research on the big data of traditional taxi and online car-hailing: a systematic review," *Journal of Traffic and Transportation Engineering*, vol. 8, no. 1, pp. 1–34, 2021.

[48] I Corp, *IBM SPSS Statistics for Windows, version 24.0*, IBM, New York, NY, USA, 2012.

[49] T. Raykov and G. A. Marcoulides, *Introduction to Psychometric Theory*, Routledge, New York, NY, USA, 2010.

[50] A. M. Arachchige, "Structural equation modelling with AMOS," 2018, http://www.morebooks.de/store/cn/book/structural-equation-modeling-with-amos/isbn/978-613-9-82333-8.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 5664004, 16 pages
https://doi.org/10.1155/2022/5664004

Research Article

Analysis of the Relationship between Dockless Bicycle-Sharing and the Metro: Connection, Competition, and Complementation

Yuru Wu,[1] **Weifeng Li**,[1] **Qing Yu**,[1,2] **and Jian Li**[1]

[1]*Key Laboratory of Road and Traffic Engineering of the Ministry of Education, College of Transportation Engineering, Tongji University, Shanghai 201804, China*
[2]*Center for Spatial Information Science, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 2778563, Japan*

Correspondence should be addressed to Weifeng Li; liweifeng@tongji.edu.cn

Received 16 December 2021; Revised 5 April 2022; Accepted 12 April 2022; Published 9 May 2022

Academic Editor: Wenxiang Li

Dockless bicycle-sharing (DLBS) is one of the novel transportation modes emerging in recent years. As a newly arisen mode, dockless bicycle-sharing inevitably has influence on the existing components of the public transportation system, especially the metro system. A large number of scholars have explored the integration relationship between the two. However, through the evaluation and quantification of the dockless bicycle-sharing data and the metro automatic fare collection data, we find that the relationship between the two is not unique. Based on the location of origin and destination, the travel duration, and the travel distance, the dockless bicycle-sharing trips closely related to the metro were identified and categorized into three different temporal-spatial relationships: competition trips, connection trips, and complementation trips. Three indicators were proposed to characterize the relationship between the two systems. A case study was carried out in Shanghai, China. The proposed method was applied to investigate when, where, and to what extent the dockless bicycle-sharing trips compete with, integrate with, and complement the metro. The results show that dockless bicycle-sharing mainly integrates with and complements the metro. It is where the dockless bicycle-sharing trip takes place and the trip significantly determines its relationship with the metro. The findings provide significant implications regarding the design and management of dockless bicycle-sharing and the metro.

1. Introduction

Since the end of the 2000s, bicycle-sharing systems have been rapidly developed and adopted worldwide [1] and have become an important component of the public transportation system for short-distance travel in cities. The Internet-based dockless bicycle-sharing (DLBS) systems are particularly well received by the public and universally recognized by the market [2]. The DLBS enhances urban mobility, promotes the accessibility of public transport, and reduces the use of motorized transport [3]. As a newly arisen mode, DLBS inevitably has influence on the public transit systems, including the bus system and the metro. The DLBS can positively or negatively impact public transit meaning it can be either a collaborator or a competitor for public transit. [4] As a solution to last-mile problems, the first- and last-mile connections to and from the metro have gotten wide attention. In Shanghai, there are 1.07 million DLBS trips on average whose origins or destinations are associated with metro stations, accounting for about 34% of the daily DLBS trips [5]. Statistics in Shenzhen [6] also show that 47% of the DLBS trips are used to reach metro stations. However, the relationship between the DLBS and the metro is mixed and complex. Except for the collaboration (connection), the other relationship, such as the complementation and competition relationship between the two systems have not been well substantiated and qualified.

The lack of data integration and the lack of integrated analytic skills are the two reasons for the research gaps in deciphering the relationship between the two systems. The understanding of the relationship between the two systems requires the traceability of the entire journey with multiple transportation modes. Some existing studies collected data based on traditional inquiry surveys which cannot fully trace

the entire travel process. Some literature applied large-scale datasets from the two systems, respectively, but lacked the ability to incorporate the different data sources.

To fill the research gaps in the existing literature, this study introduces the DLBS data and the metro automatic fare collection (AFC) data to evaluate and quantify the relationship between DLBS and the metro. Based on the location of origin and destination of DLBS trips, the DLBS travel duration, and the DLBS travel distance, five scenarios are proposed to establish the connection between DLBS and the metro and enable to analyze the travel patterns of the two systems as a whole. DLBS trips closely related to the metro are identified and categorized into three groups: competition trips, connection trips, and complementation trips. To investigate when, where, and to what extent the DLBS trips compete with, integrate with, and complement the metro, three indicators are proposed to characterize the relationship between the two systems. Finally, the proposed method is carried out in the case study of Shanghai, China. Several significant implications regarding the design and management of DLBS and the metro are summarized from the results.

The contributions of this paper are as follows:

(i) A method is proposed to identify the relationship between DLBS and the metro based on the data integration of the large-scale dataset collected from the two systems, respectively. A set of indicators is proposed to integrate the travel patterns of the two systems and quantify the relationship between the two systems.

(ii) Through the example verification of Shanghai, it is confirmed that in addition to the connection, there are also overlapping (competition relationship) and supplement (complementation relationship) between dockless bicycle-sharing and the Metro.

The remainder of this paper is organized as follows. Section 2 presents a review of the related works. Section 3 introduces the multi-source dataset. Section 4 presents the methodology of the relationship identification between the two systems. Section 5 presents the case study in Shanghai, China. Finally, conclusions and future directions are given.

2. Related Works

In this section, the related work is summarized in three aspects, which are the DLBS usage characteristics analysis, the relationship between the BSS and the public transit system, the relationship between the DLBS and the metro.

2.1. The DLBS Usage Characteristics Analysis. Sufficient studies have analyzed the characteristics of the usage of bicycle-sharing. Compared to docked bicycle-sharing, DLBS have no fixed stations and are usually private, start-up, and venture capital projects. Most DLBS users are mobile Internet users and most DLBS units are equipped with GPS devices for easy redistribution and maintenance. [7] Many studies [8–12] have proposed a series of methods to analyze the temporal and spatial heterogeneity of DLBS usage characteristics. For example, Zhuang et al. [8] applied cluster analysis to a DLBS dataset to automatically identify typical cyclic patterns in the spatiotemporal dimension. Song et al. [9] analyzed the demand of BSs and used global and local Moran's I index and community detection to model the spatiotemporal dynamics of circular mobility. Bao et al. [11] applied clustering methods and text mining-based potential dirichlet assignment methods to identify DLBS trip distributions and trip purposes. Zhang et al. [13] studied the temporal and spatial heterogeneity from the perspective of penetration theory, and the results revealed the oversupply of bikes in urban centers and the imbalance between supply and demand on a larger scale. In addition, a series of studies have analyzed the factors influencing usage characteristics, showing that conditions such as land use, socio-economics, population density, roadway designs, transportation facilities, cycling infrastructures, weather (including temperature, humidity, air quality) [3, 14–18], and changes in human mobility due to the COVID-19 pandemic [19–21] can have an impact on the mobility characteristics of bike-sharing (either DBS or DLBS).

2.2. The Relationship between BS and the Public Transit. With the gradual increase of research on the regularity of bicycle-sharing movement, the relationship between bicycle-sharing and the public transit is being further explored. Fuller *D* pointed out that bicycle-sharing is associated with transportation mode transfer in his research on the impact of BS on collision accidents in Canada [22]. Zhang [23] investigated the relationship between the frequency of public transportation use and the probability and frequency of BS use using survey data from the United States. The results showed that for each unit increase in the frequency of public transportation use, the probability of using BS increased by 4.0% and the frequency of using BS increased by 1.4%. Zhu et al. [24] surveyed public bicycle users in Shanghai and found that most users used shared bicycles to replace their original public transportation. Martin et al. [25] focused on Washington, DC and Minneapolis bicycle-sharing users' travel behavior. They found that in low-density urban areas, bike-sharing users used it more to connect to public transportation services; in high-density urban centers, more bike-sharing trips were used to replace public transportation. Fishman [26] surveyed bikeshare users in five cities, Melbourne, Brisbane, Washington D.C., London, and Minneapolis, and ultimately found a substitution relationship between bikeshare, walking, and public transportation. Fuller et al. [27] used telephone survey data to investigate bikeshare use in Montreal and found that bikeshare had a substitution effect on public transportation. Campbell [28] developed a two-difference model using data from transit and bikeshare systems to test the substitution relationship between BS and transit in New York City. Kong et al. [29] deciphered the relationship between bicycle-sharing and public transit and divided them into three categories: Modal substitution, integration, and complementation, which have great reference value.

2.3. The Relationship between BS and the Metro. In practical applications, bicycle-sharing is one of the important ways to solve the "last-mile" problem of public transportation; therefore, in recent years, a wealth of research on the relationship between bike-sharing and metro systems has focused on the integration between the two. Bocker et al. [30] found that when other factors such as distance, elevation, travel time, and urban form are consistent, bike-sharing traffic will be significantly higher if the trip origin and destination are related to a rail transit station. Part of the literature [10, 18] mined the impact of bike-sharing on rail transit stations and used hierarchical clustering to cluster rail transit stations into morning and evening peak bike-gathering stations, morning peak bike-gathering stations, and evening peak bike-gathering stations. Yu et al. [31] used bike-sharing data and subway travel data to extract indicators and used a confidence ellipse approach to make an analysis of the service area of shared bicycles around metro stations. Yan et al. [32] extracted shared bicycle operation order data related to Shanghai rail transit line 9, visualized and analyzed the differences in the usage patterns of dockless shared bicycles around the metro on weekdays and holidays, and revealed the characteristics of the service area of dockless shared bicycles around the metro. In addition, some studies have investigated the factors influencing the integration between bike-sharing and metro. For example, Muhammad et al. [33] investigated and analyzed user preferences for last-mile feeder using a mixed choice model and found the influence of user preferences on metro-BS integration. Hu et al. [34] fitted a set of generalized additive models considering marginal nonlinear interactions and examined the relationship between the external environment (including land use, sociodemographics, roadway designs, transportation facilities, metro station features, and DLBS operator features) and the activity characteristics of metro-integrated bicycles. And, some bike-sharing systems that primarily serve city centers (e.g., the bike-sharing system in downtown Dublin) or serve suburban areas (e.g., the system in Jiangning District, Nanjing, China) show integration relationships with the metro as users need to commute between central and suburban areas. [35–37].

Apart from integration, other relationships between bike-sharing and metro have been less studied. Ma Knaap [38] compared the use of bicycle-sharing and metro in Washington, DC in 2010 and 2015, and found that for metro stations in the downtown area, there are bicycle-sharing stations within a 1/4 mile (about 402 m) range, which reduces the number of metro travels, and for the metro stations in the fringe area, the stations with bicycle-sharing increase the metro passenger flow. This demonstrates the substitution and complementarity of bike-sharing for the metro.

2.4. Literature Summary. In summary, the research on the relationship between bike-sharing systems and public transportation: bike-sharing and public transit (bus) can be summarized as substitution, integration, and complementation. In contrast, research on the relationship between dockless bicycle-sharing systems (DLBS) and subways focuses on integration, and there is a gap in research on other relationships between the two. In addition, it can be seen from the previous studies that most of the studies are based on questionnaire surveys, studying the relationship between bicycle-sharing and public transportation and trying to find out the influence factors. However, traditional survey methods have many drawbacks. With the development of information and communication technology, multi-source big data contain a wealth of crowd movement information and has the characteristics of large samples and low cost [39], which can provide more effective tools for bicycle-sharing-related research. Therefore, to fill the research gap, this study aims to fuse the travel data of the two systems, quantify the relationship between the DLBS system and the metro, and investigate the spatiotemporal patterns of DLBS trips that replace, integrate, and complement the metro.

3. Datasets

3.1. The Dockless Bicycle-Sharing Data. In this study, the DLBS data were provided by Shanghai Transportation Commission, including dockless bicycles registered through all formal channels in Shanghai from May 1st to May 12th, 2019. It is composed of more than 1.7 million bicycles and generates 1.07 million average daily trips. The information of each data includes the bicycle ID, lock status, timestamps, and longitude and latitude coordinates. In addition, most of the data processing in this paper is supported by the Python package TransBigData. [40].

Several filters were applied to the data before further data processing. First, the OD information was extracted. The data were generated when the lock was opened and closed, so by which record the starting and ending positions of each single bicycle travel. The change of lock status of each bicycle with the same ID can correspond to the beginning and end of a riding order. Thus, the OD information of each order can be extracted. And, through the calculation and sorting, the bicycle ID, the duration and distance of travel, the departure and arrival time, and the departure and arrival position are explicitly recorded in the dataset. Therefore, by monitoring the DLBS data, the DLBS system can be turned into a virtual sensor network for sensing mobility in the city. Second, there existed some outliers with too short or long travel time. In this study, trips with durations less than 1 min or greater than 120 min were identified as outliers and were dropped. Third, to reflect the actual demand of DLBS, May 6th (Monday) and 12th (Sunday) were selected as two typical days to represent the travel demand on weekdays and weekends. After data cleaning and data filtering, 1,086277 and 853237 DLBS trips were obtained on May 6th and 12th, respectively (Figure 1).

3.2. Metro Automatic Fare Collection Data. The metro automatic fare collection (AFC) data were collected from the metro system in Shanghai from May 1st to May 12th, 2019. Shanghai metro network is composed of 17 lines and 387 stations, with an average daily ridership of 10.6 million. The

metro AFC system stores the inbound and outbound records, and the metro ridership at the station-to-station level. The metro AFC data used in this paper provide the information of hourly passenger flows in 387 metro stations, including the metro line ID, the metro station name, hours of the day, the hourly inbound passenger flow, and the hourly outbound passenger flow.

Figure 2 shows the temporal variation in hourly ridership of the metro network on May 6^{th} and 12^{th}. The temporal variation of hourly ridership shows significant differences between weekdays and weekends. On weekdays, the temporal variation presents obvious characteristics of morning and evening peak within a day. While on weekends, the temporal distribution of ridership tends to be relatively uniform.

4. Methodology

4.1. Methodology Framework. Figure 3 shows the methodology framework of this study. The DLBS data, metro GIS data, and metro AFC data were used to evaluate and analyze the relationship between the DLBS system and the metro system. The methodology can be briefly described as follows:

(1) Spatial matching. The origin and destination of each DLBS trip are extracted from the DLBS data. With the help of metro GIS data, spatial matching is carried out to search and match the origin and destination with their nearest metro stations.

(2) Buffer area segmentation. To distinguish the DLBS demand around metro stations, a data-driven method is proposed to segment the coverage area of metro stations.

(3) DLBS trip classification. To establish the connection between DLBS and the metro, DLBS trips closely related to the metro are identified based on the location of DLBS origin and destination, the DLBS travel duration, and the DLBS travel distance. According to the relationship with the metro, the identified trips are categorized into three groups: competition trips, connection trips, and complementation trips.

(4) DLBS-metro Relationship Characterization. To analyze the travel patterns of the two systems as a whole, three indicators are proposed to characterize the relationship between the two systems, including the share of connection around the metro station, the intensity of competition along the metro route, and the demand of complementation outside the metro coverage area.

4.2. Identification of Relationship between DLBS and the Metro. The large-scale dataset of DLBS trips applied in this study enables to collect the travel information of DLBS trips with smaller bias and less labor intensity. By comparing and fusing the two types of data, the relationship between their travel demands at the spatial and temporal levels can be found and thus their relationship can be explored.

First, trips with a relationship to the metro are identified, the distance between DLBS origin/destination and metro station is the main factor to be considered. In previous studies, most scholars often take a definitional approach to identify the interchange trips: a bike-sharing trip is defined as an interchange trip if its origin or destination is located in a catchment area around a metro station. Some literature defines the radius of the catchment area as 300 m [41, 42] or 500 m [5, 31, 43, 44], and the smaller the radius, the higher the percentage of interchange trips within the catchment area [45]. In a face-to-face questionnaire survey conducted by Li et al. [46] on the trip purpose of bike-sharing users around the metro, most respondents (>75%) indicated that their bike-sharing trips were related to the metro when the origin or destination was located within a 500 m buffer zone of a Shanghai metro station. Zhao et al. [47] proposed a data matching method based on association rules and verified it. The results show that 573 pairs of smart cards are matched within 300 m around the subway station, and the accuracy is 100%. Therefore, we can assume that by judging the distance between DLBS origin/destination and metro station, we can judge the relationships of most trips.

Then, the time of DLBS trip occurrence and DLBS travel distance are considered to make it closer to the real situation and thus increase the accuracy. Finally, the trips are categorized into three modes and they are defined as follows:

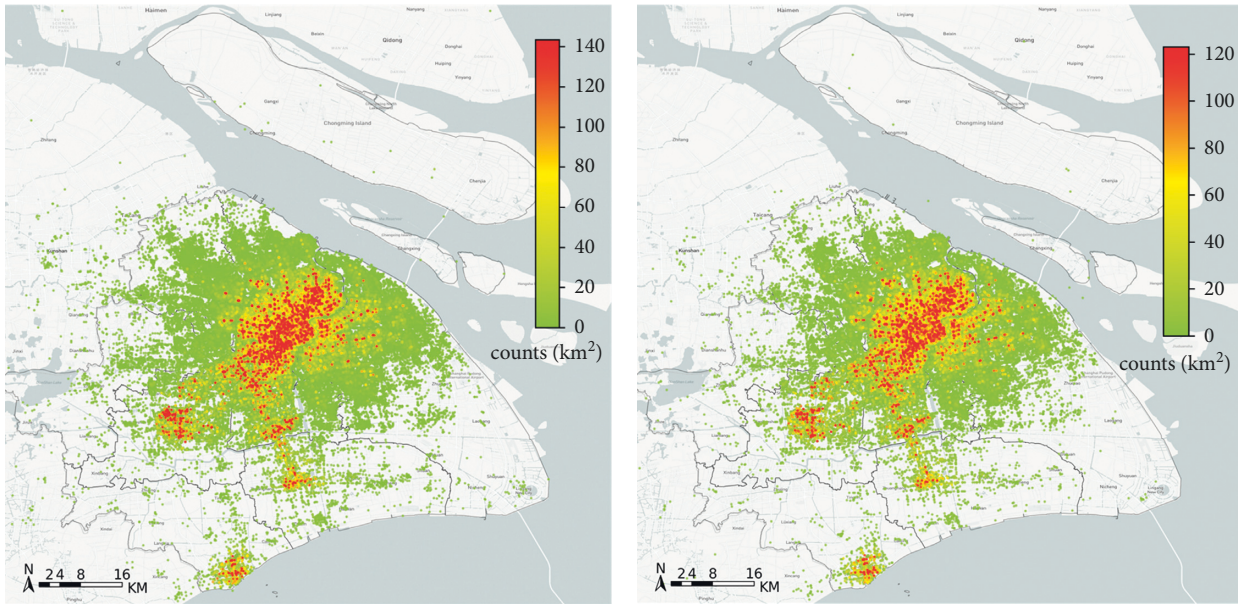
(1) Competition: Competition refers to the situation in which DLBS is used as a substitute for the metro. Within the service time of the metro system, if both the origin and the destination of DLBS trip are close to metro stations, the user is supposed to take the metro to complete the trip without too much walking, waiting, and transferring. In this case, DLBS is considered an alternative to the metro.

(2) Connection: Connection refers to the situation where the metro is the core of the entire journey, and DLBS is used to get access to the metro. In this case, both the spatial relationship and the temporal relationship between DLBS and the metro should be considered. The spatial relationship between the two systems requires that either the origin or destination of the DLBS trip should be close to metro stations. For temporal relationship, the DLBS trip should arrive shortly before the departing time of metro trip or depart shortly after the arriving time of metro trip. In addition, the duration of the DLBS trip should not be too long.

(3) Complementation: Complementation refers to the situation where DLBS is used in areas with insufficient metro coverage. In other words, if either the origin or destination, or both, of the DLBS trip is far away from metro stations, DLBS is considered a complementary mode to the metro.

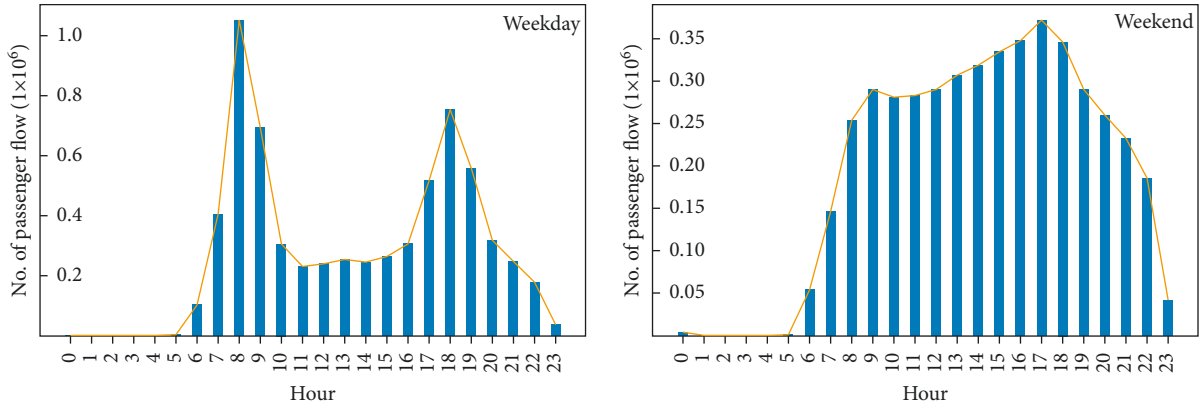
Based on the definition of relationships between the two systems, a three-step method is proposed to distinguish the relationship between DLBS and the metro.

4.2.1. Spatial Matching. The first step is spatial matching. To identify the relationship between DLBS and the metro, it is important to understand the spatial relationship between the



(a) (b)

Figure 1: Spatial distribution of DLBS trips' origin point (unit: counts/km^2). (a) May 6th (Weekday). (b) May 12th (Weekend).



(a) (b)

Figure 2: The temporal variation in hourly ridership of the metro network. (a) May 6th (Weekday). (b) May 12th (Weekend).

origin and destination of each DLBS trip and the metro stations. To improve the speed of matching, this paper uses The KDtree algorithm to search the origin and destination of DLBS trips, and matches them with the nearest metro stations.

(1) Extract the departure and arrival information of each DLBS trip from the DLBS data, and determine the geographic coordinates of the origin and destination.

(2) Divide the space of the metro stations, establish a spatial index, and perform a binary tree search on the origin and destination of each DLBS trip to find the nearest metro station.

(3) Calculate the closest distance between the origin or destination and the metro station.

$$dist = 2 \arcsin\sqrt{\sin^2\frac{y_1 - y_2}{2} + \cos\ y_1 \times \cos\ y_2 \times \sin^2\frac{x_1 - x_2}{2}} \times 6378.137, \tag{1}$$

where x_1, y_1 denote the longitude and latitude of DLBS trip origin or destination, x_2, y_2 denote the longitude and latitude of the nearest metro station, and *di st* represents the distance measured by the kilometer.

4.2.2. Buffer Area Segmentation. Buffer area segmentation is applied to divide the coverage area of the metro system. Previous studies mostly generated the buffer area of public

transportation based on questionnaire surveys or empirical experience, which cannot accurately reflect the buffer relationship between bicycle-sharing and public transportation. Hawas et al. [48] proposed that a place can be considered to be covered by public transportation, if it gets access to a public transit station within a comfortable walking distance. Wu [49] determined this comfortable walking distance to be 400 meters. Jin et al. [50] introduced 100 meters as another threshold to assist the analysis of the relationship between Uber and public transit. Kong et al. [29] proposed "traffic coverage" to measure the spatial distribution of public transportation services, and divided the urban space into three regions based on the threshold of 100 meters and 400 meters. And, Hu et al. [43] proposed a spatial network density-based method to determine the optimal size of the parking ring (buffer area) based on large volume of bike-sharing travel data in determining the bike-and-ride (BnR) trips, where the size of the parking ring is measured by the network distance and its boundary is obtained from the concave surface of all accessible nodes, which can better reflect the real cycling scenario. (Figure 3)

This study proposes a data-driven method for buffer area segmentation based on large-scale DLBS data. Given the origin and destination of DLBS trips, the distances between the origin or destination and the nearest metro station can be obtained. Applying Kernel Density Estimation to find the estimated probability density function of distance, the spatial relationship between DLBS and the metro can be visualized in Figure 4.

Most DLBS trips start or end within two kilometers away from metro stations. It can be seen from Figure 4 that there are two peaks in the probability density function. The first peak is around 80–100 meters, the second peak is around 400–600 meters, and a trough can be found around 240–280 meters. The first peak is likely to associate with the bicycle parking facilities around the metro stations. The second peak is the outcome of the origins or destinations of users' first-mile or last-mile trips. The trough between the two peaks represents the weakened spatial agglomeration of DLBS demand. It indicates that 260 meters can be used as a threshold to distinguish the DLBS demand around metro stations.

In addition, to distinguish whether the DLBS trip is associated with the metro, set the second threshold. Without considering the connection transportation, the service range of the metro station is that people can walk (750 m) to the entrance and exit of the station within 10 minutes, and it should be the superposition of the reasonable reachable range of multiple single starting points [51]. Therefore, considering the distribution of multiple entrances and exits of the metro station and surrounding roads, it is set as 1000 m. The two thresholds divide the coverage area around a metro station into three parts:

(1) Near metro station area (Area A): Buffer areas within 260 meters away from the metro stations. Area A is within a few minutes' walk from the metro stations. If the DLBS trip starts or ends in Area A, the user is likely to pick up the bicycle after alighting the train or drop off the bicycle before boarding. 21.20% of the weekday DLBS trips and 23.23% of the weekend DLBS trips start or end in Area A around the metro stations.

(2) Metro radiation area (Area B): Buffer areas within the range of 260–1000 meters away from the metro stations. Area B is the transition area; it lies in the coverage area of the metro station but outside the comfortable walking distance which indicates the service coverage of the subway system. 53.45% of the weekday DLBS trips and 59.78% of the weekend DLBS trips start or end in Area B around the metro stations.

(3) Far metro station area (Area C): Areas outside the buffer areas, more than 1000 meters away from the metro stations. Trips in these areas will be unrelated to the metro system, meaning that when the OD of the trip is all far from the metro, the bike becomes a transportation option.

4.2.3. DLBS Trip Classification. Allocating the origin and destination of each DLBS trip to the areas as mentioned above, the cases shown in Table 1 are proposed to represent the relationship between DLBS and metro, which can contain all valid DLBS trip orders in the dataset.

According to the spatial distribution of the seven ODs obtained in the above table, to further improve the accuracy of the relationship classification, attention should be paid to the time of trip occurrence, and orders that exceed the Shanghai metro's operation hours (5 : 25 to 23 : 00) should be excluded from the potential connection and substitution relationships and categorized in the supplementary relationships. Finally, combined with the DLBS trip distances, five key scenarios are proposed to represent the relationship between DLBS and the metro which are shown in Figure 5.

Scenario 1. Short-distance DLBS trips between Area A and Area B/C

In Scenario 1, Considering that the distance of DLBS trips for connection should not be too long [4], we mainly focus on the DLBS trip shorter than or equal to 2 kilometers. The DLBS trips either start or end very close to the metro stations. DLBS is used as the first-mile/last-mile connection of the metro. In this case, DLBS trips are categorized as connection trips.

Scenario 2. DLBS trips between Area A around different metro stations

In Scenario 2, both the origin and destination of DLBS trips are located very close to the metro stations. The DLBS trips in Scenario 2 are likely to replace the trips formerly made by metro. In this case, DLBS trips are categorized as competition trips.

Scenario 3. DLBS trips between Area B around the same metro station

In Scenario 3, there is a certain distance between the metro station and the origin and destination of DLBS trips.

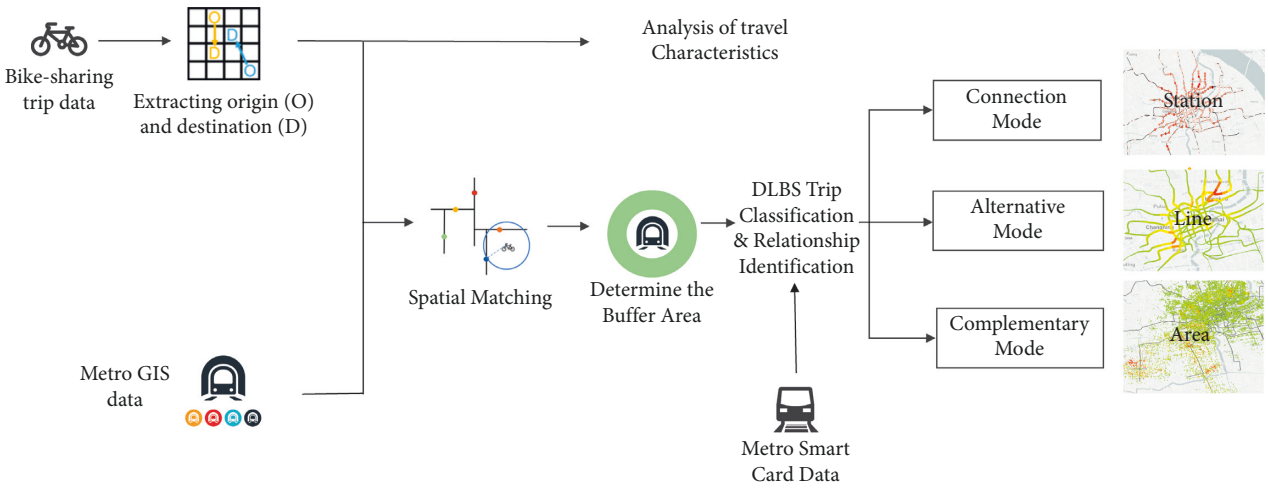


Figure 3: Methodology framework.

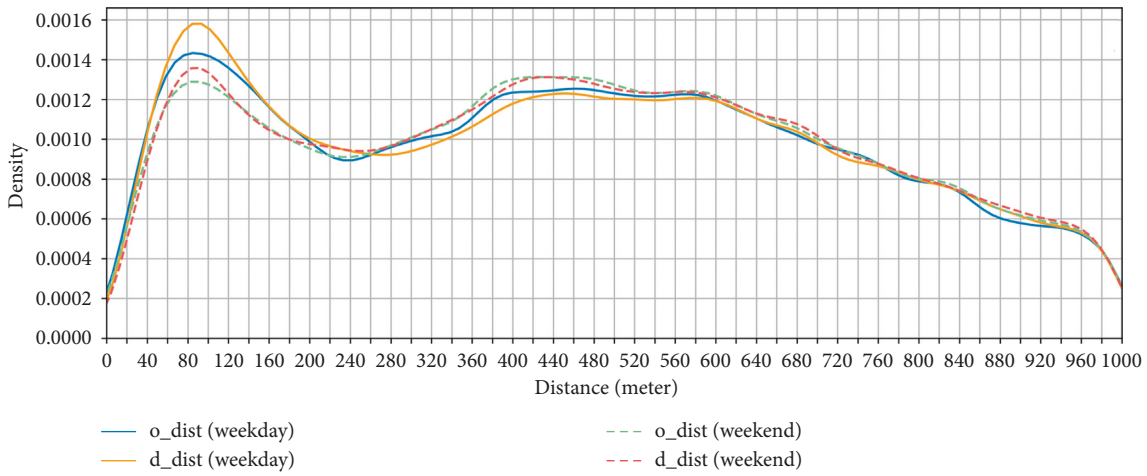


Figure 4: The estimated probability density function of the distance between the origin or destination of DLBS trips and metro stations.

Table 1: Spatial relationship between DLBS trips OD and buffer zone.

Cases	Description	Relationship
Case 1	Only one of O and *D* is in area A	Potential connection
Case 2	Area A to area A around different metro stations	Potential competition
Case 3	Area B to area B around different metro stations	Potential competition
Case 4	Area A to area A around the same metro station	Excluded due to short-distance
Case 5	Area B to area B around the same metro station	Complementation
Case 6	Area B/C to area C/B	Complementation
Case 7	Area C to area C	No relation

Are located very close to the metro stations. In this case, since both the origin and destination are in the Area B around the same metro station, DLBS is a complementary mode to the metro. Therefore, DLBS trips are labelled as complementation trips.

Scenario 4. DLBS trips between Area B around different metro stations

In Scenario 4, users prefer to use DLBS to complete the entire journey between Area B around different metro stations. Both the origin and destination of DLBS trips are a certain distance away from the metro stations. Compared with the metro, DLBS provides a convenient, inexpensive, and door-to-door service. In this case, DLBS is more likely to be considered as a substitute for the metro. Therefore, DLBS trips are categorized as competition trips.

Scenario 5. DLBS trips between Area B and Area C

In Scenario 5, either the origin or destination of DLBS trips is outside the coverage of the metro. These DLBS trips

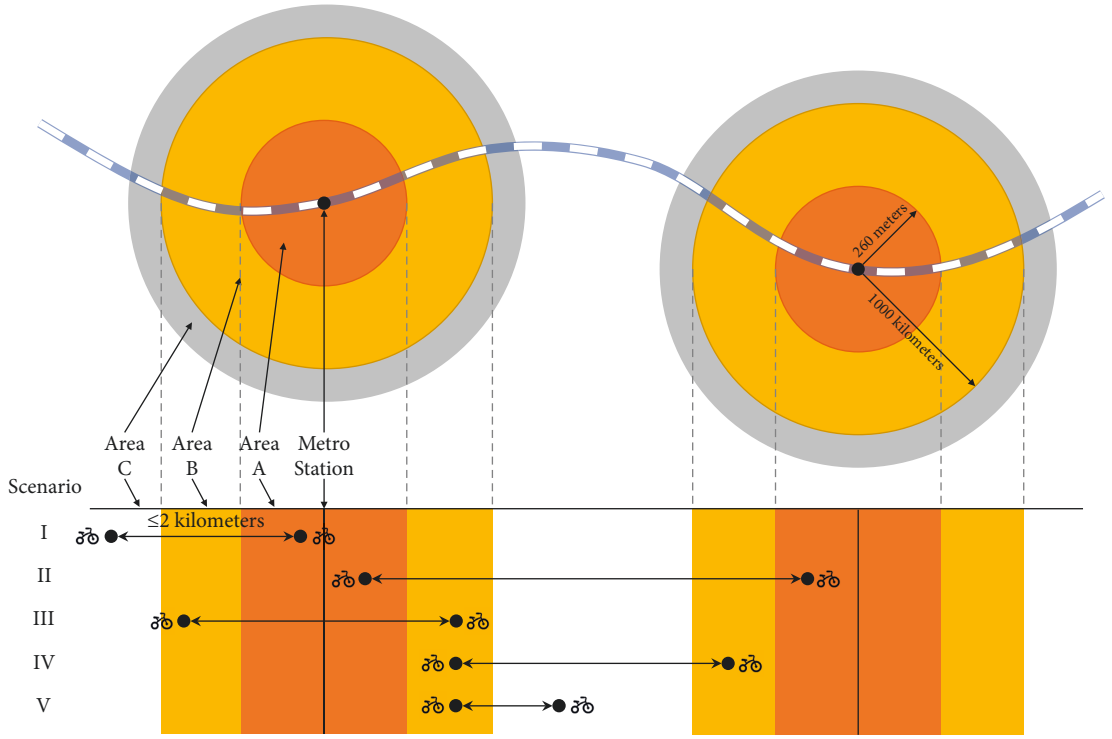


Figure 5: Illustration of the five scenarios.

cannot replace or be combined with the rail transit. In this case, DLBS acts as a complementary mode to enhance the urban mobility in areas with insufficient metro coverage. In this case, DLBS trips are labelled as complementation trips.

4.2.4. DLBS-Metro Relationship Characterization. The five scenarios mentioned above establish the connection between DLBS and the metro and enable to analyze the travel patterns of the two systems as a whole. Three indicators are proposed to characterize the relationship between DLBS and the metro. The three indicators correspond to the travel patterns of three types of DLBS trips, respectively, including the share of connection around the metro stations, the intensity of competition along the metro lines, and the demand of complementation outside the metro coverage area.

(1) The share of connection. The share of connection refers to the share of DLBS connection in the passengers of the metro station. This indicator is proposed to analyze the DLBS connection demand of metro passengers at the station level. Given a time window Δt, let A_i denote the Area A around the metro station i, in the DLBS data, there are o_i DLBS trips departing from A_i and d_i trips arriving at A_i. Meanwhile, there are o_i' boarding passengers in the metro station i and d_i' alighting passengers in this station. Then, the share of connection at metro station i can be calculated as :

$$S_i = \frac{d_i + o_i}{o_i' + d_i'}. \tag{2}$$

(2) The Intensity of Competition. The intensity of competition refers to the metro passenger flow that has been diverted to DLBS. This indicator is proposed to restore the distribution of diverted metro passenger flow from the O-D demand of DLBS competition trips. Given a time window Δt, let A_i and B_i denote the Area A, B, respectively, around the metro station i, in the DLBS data, there are n_{ij} DLBS competition trips from A_i to A_j and m_{ij} competition trips from B_i to B_j. Assuming that all the DLBS competition trips were formerly made by metro, the diverted metro O-D demand from station i to station j can be calculated as (3):

$$r_{ij} = n_{ij} + m_{ij}. \tag{3}$$

Based on the diverted metro O-D matrix $R = (r_{ij})$, traffic assignment is carried out to allocate trips between metro stations to the metro network. For the two adjacent metro stations i, j, the intensity of competition I_{ij} refers to the allocated passenger flow in the segment from station i to j, that is, the segment passenger flow diverted by DLBS.

(3) The Demand of Complementation. The demand of complementation refers to the number of DLBS complementation trips in the areas with insufficient metro coverage. This indicator is proposed to analyze the demand of DLBS complementation trips at the area level. Given a time window Δt and a 500 m × 500 m grid covering the study area, there are u_i DLBS competition trips departing from the grid cell i and w_i trips arriving at this cell. Then, the demand of complementation trips in the grid cell i can be calculated as:

$$D_i = u_i + w_i. \tag{4}$$

5. Results and Discussion

5.1. Travel Characteristics of DLBS. The travel characteristics of DLBS are analyzed based on the travel duration, the travel distance and the turnover rate obtained from DLBS data.

As for the travel duration, the average travel duration is 13.12 minutes on May 6th and 11.05 minutes on May 12th. The average travel duration on working days is slightly longer than that on weekends.

As for the travel distance, the average travel distance is 1250.9 meters on May 6th and 1199.3 meters on May 12th. The estimated probability density function of DLBS travel distance is shown in Figure 6. There is no great difference in travel distance between weekdays and weekends. The distribution of travel distance has an obvious peak and fat tail. The DLBS travel distance is mostly smaller than 5000 meters and reaches a peak at around 600–700 meters (Figure 7).

The daily turnover rate of bicycles can reflect the overall situation of bicycle resource usage. The average turnover rate of active bicycles is 5.68 on May 6th and 4.97 on May 12th. However, according to the statistics published by the government, the average turnover rate of all the shared bicycles in Shanghai was 1.1 in 2019. The dramatic difference indicates that there are a large number of inactive bicycles in the city. The horizontal axis of the bar chart in Figure 7 represents the daily turnover rate, while the vertical axis is the frequency, showing the frequency distribution of the different daily turnover rates of active bikes in the dataset. The figure shows that most of the bike-sharing bikes have less than ten completed orders per day. And, compared with the turnover rate on weekends, there are more bicycles used for more than three times on weekdays.

5.2. Analysis of the Relationship between DLBS and the Metro

5.2.1. General Patterns. Given the preprocessed DLBS data on May 6th and 12th, 2019, the relationship between DLBS and the metro was identified, and the proportion of different kinds of relationship was illustrated in Figure 8. The share of three kinds of trips is significantly different between weekend trips and weekday trips. Figure 8(a) shows that the connection trips dominate weekday trips, followed by complementation trips and competition trips. Figure 8(b) shows that the share of complementation trips is the largest in weekend trips, followed by connection trips and competition trips. The share of connection trips and complementation trips are considerable in both weekdays and weekends, accounting for about 75% of DLBS trips. The results show that DLBS mainly integrates with and complements the metro.

5.2.2. Temporal Patterns. In addition to the difference between weekend and weekday trips, DLBS trips in different times of day might also be different. Figure 9 illustrates the number of connection, complementation, and competition trips in different times of day, and Figure 10 illustrates the temporal variation of the share of different kinds of trips. Several observations can be made from the two figures. First, the hourly patterns of DLBS trips are clearly different between weekday and weekend. All the three kinds of DLBS trips show sharp morning and evening peak on weekdays, whereas the weekend patterns are featured with relatively smooth summits. These results suggest that all the three kinds of DLBS trips have a large percentage of commuting trips, whereas the purposes of weekend trips are more diversified. Second, the temporal patterns illustrated in Figures 8 and 9 indicate that commute plays an important role in the composition of DLBS trips. Connection trips dominate DLBS trips during the morning and evening peak on weekdays, the share of which ranges from 40% to 48%. DLBS has been widely accepted as the seamless access to metro stations and widely used in the peak commuting hours. Complementation trips make up the majority of DLBS trips during the off-peak hours of weekdays and the whole weekends. The differences in the share of connection trips and complementation trips between peak hours and off-peak hours indicate that commute as a trip purpose might be a significant determining factor in the relationship between DLBS and the metro.

5.2.3. Spatial Patterns. Figure 11 shows the DLBS connection shares around Shanghai metro stations, using the color and size of the dots to indicate the proportion of passengers entering and exiting the station at that metro station who chose DLBS for their connection. The results resonate with our previous findings that the share of connection trips is much larger on weekdays. The mean value of the share of connection is 7.5% on May 6th and 5.4% on May 12th.

From the spatial distribution of the share of connection in Figure 11, significant spatial heterogeneity can be observed. The value of the share of connection is smaller in the city center and suburban areas. The city center is covered with dense public transportation networks and well-developed metro networks. There barely exists the demand of medium- or long-distance connection in the city center. As for the suburban areas, the insufficient DLBS supply cannot meet all the demand of connection with the metro.

Fuxing Island Station, Tongji University Station, Guoquan Road Station, Sanmen Road Station, and Wuwei Road Station are the five metro stations with the highest share of connection. The five stations and other stations with high share of connection have some common features that they are located between city center and suburban areas and surrounded with residential districts or industrial districts.

These findings imply that for connection trips, the public transportation networks, the supply of DLBS service, and the building environment around the metro stations are determining factors (Figure 11).

The intensity of competition along the metro segments is calculated and illustrated in Figure 12. The color and the width of lines represent the intensity of competition with DLBS, that is, the segment passenger flow diverted by DLBS.

The spatial distribution of the intensity of competition on weekdays coincides with that on weekends. Most of the metro sections with a large intensity of competition are the sections before the first transferring stations from the

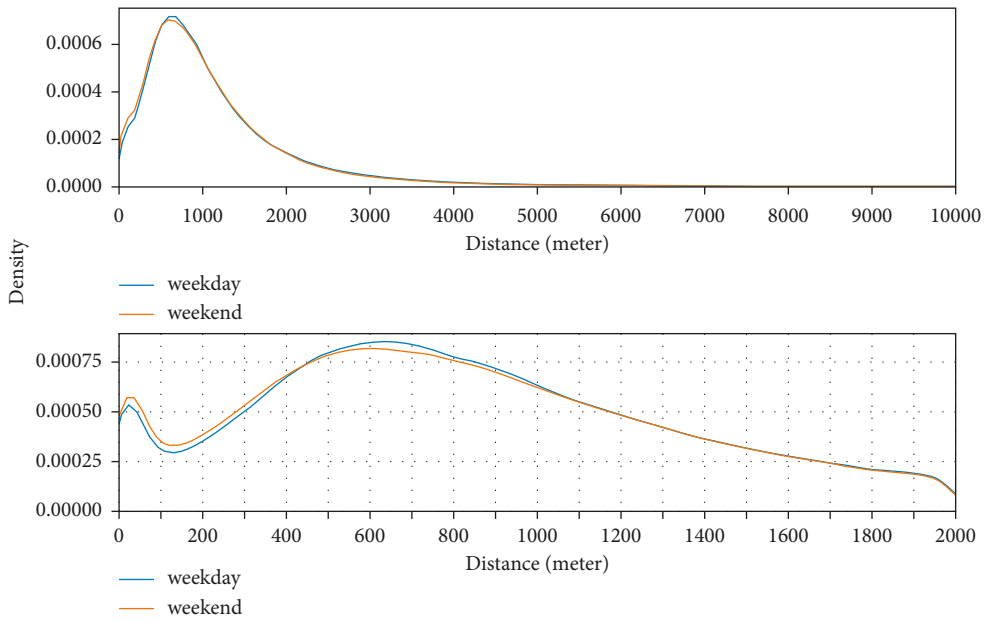


Figure 6: The estimated probability density function of DLBS travel distance.

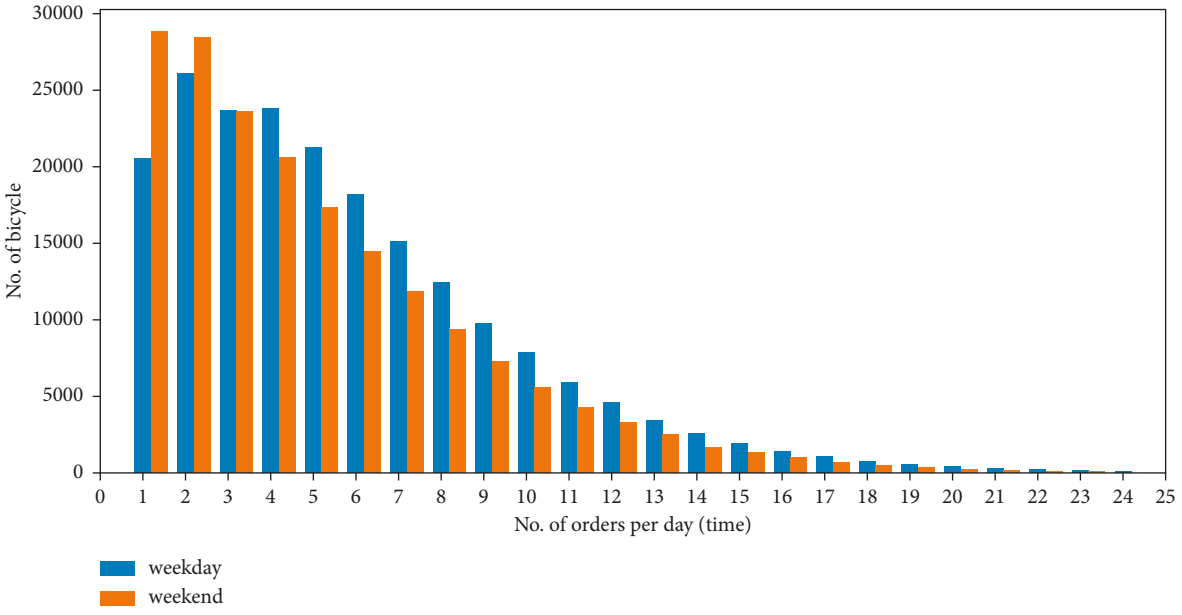


Figure 7: The distribution of the daily turnover rate of active bicycles.

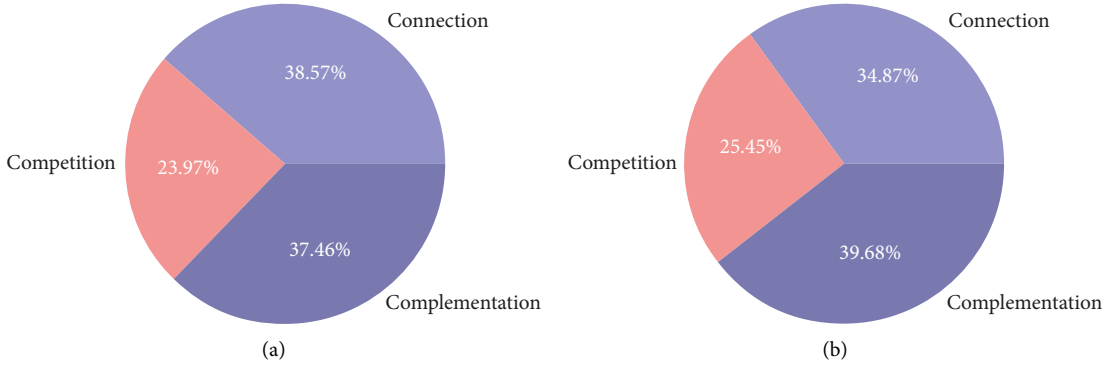


Figure 8: The share of connection, complementation, and competition trips. (a) Weekday. (b) Weekend.

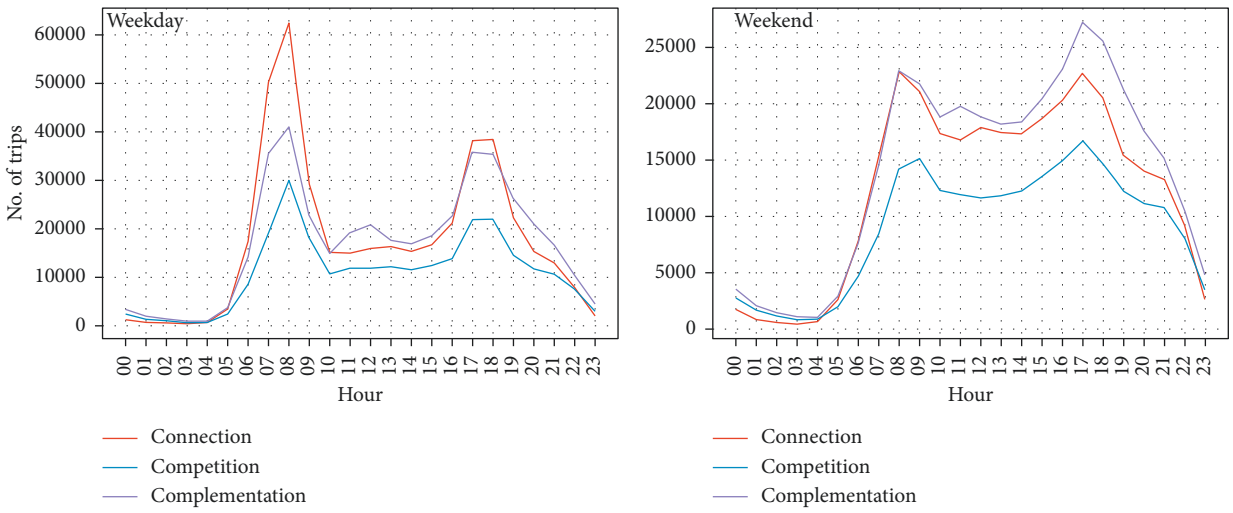


Figure 9: The hourly variation of the number of connection, complementation, and competition trips.

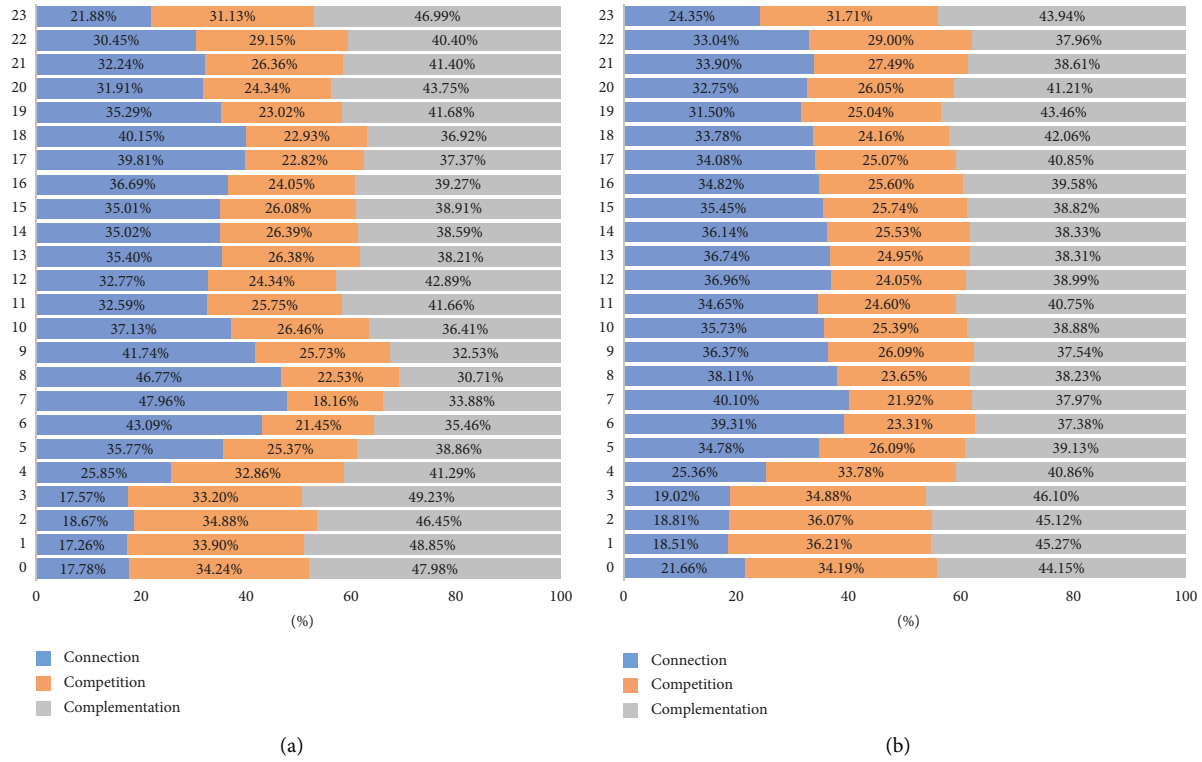


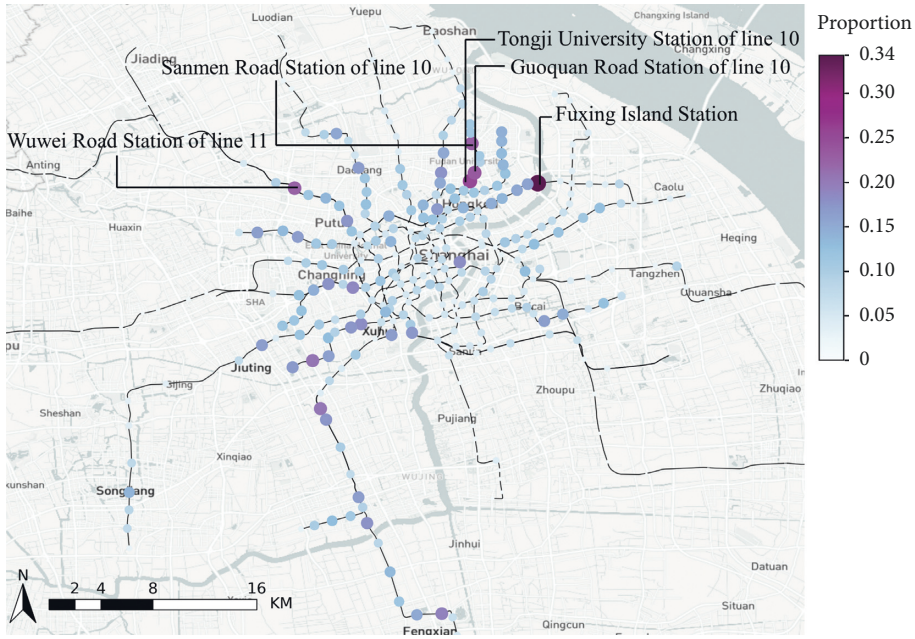
Figure 10: The hourly variation of the share of connection, complementation, and competition trips. (a) Weekday. (b) Weekend.

suburban areas to the city center. These sections are usually the maximum section of passenger flow of the metro lines.

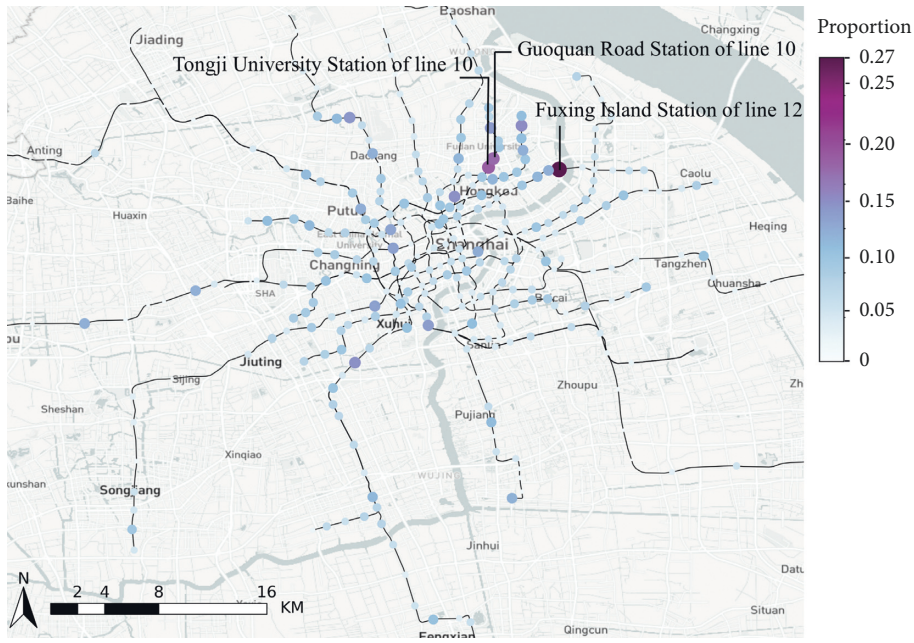
These results suggest that for competition trips, what matters is not how far a place is away from the city center, but rather how crowded the metro is. DLBS is regarded as an alternative solution to escape from the crowded metro.

With the rapid development of the city and the continuous expansion of urban lands, DLBS can be a useful tool to complement public transportation networks in areas with insufficient metro coverage. The demand of complementation is calculated and plotted in Figure 13.

As can be seen from the figure, complementation trips are more widely distributed on weekends, which might be because of the more diversified travel demand on weekends. The areas with huge demand of complementation trips are mainly located around the terminal stations in the suburban end of Metro Line 1, 5, 9, and 12. In these areas, the public transportation is not well-developed and short in supply. Preference has been given to DLBS to complete the medium- or long-distance trips. As a supplementary to public transportation services, DLBS enhances the mobility in the area and promotes the accessibility of public transportation networks.



(a)



(b)

Figure 11: The share of connection around the metro stations. (a) Weekday. (b) Weekend.

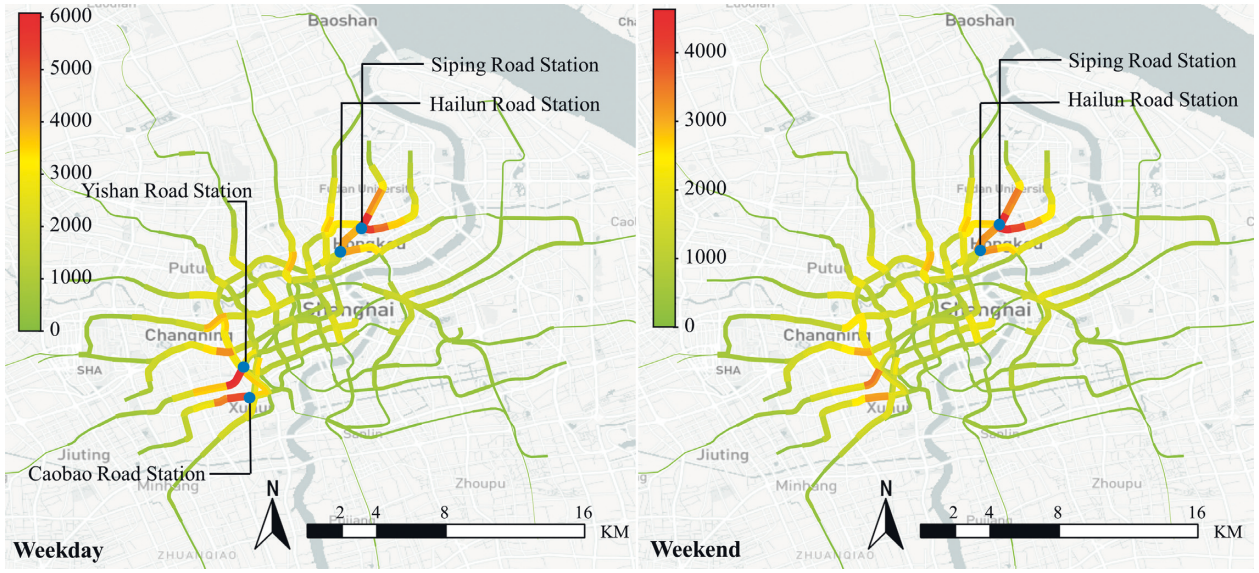


Figure 12: The intensity of competition along the metro segments.

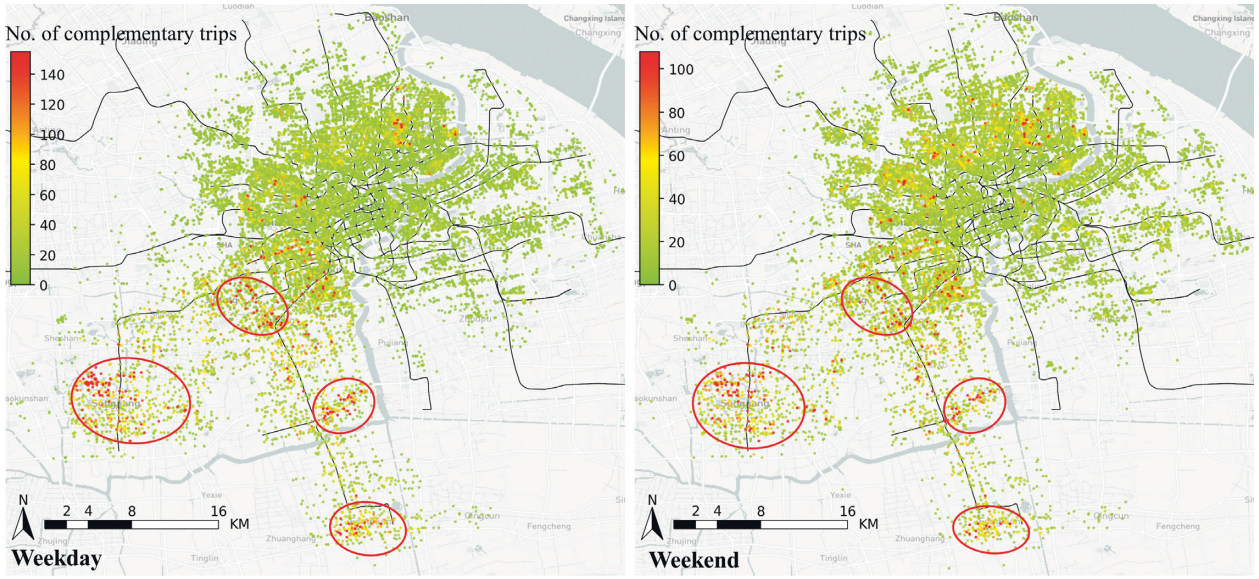


Figure 13: The demand of complementation trips.

6. Conclusions and Future Directions

In recent years, DLBS has gradually been warmly welcomed by the public. In the urban transportation system, as an environment-friendly short-distance shared travel mode, DLBS enhances the mobility of the city and promotes the accessibility of the transportation network. The development of new transportation modes will inevitably have an impact on the existing modes, especially the metro system. This study applies the DLBS data and the metro AFC data to quantify the internal relationship and the interaction between the two systems. Using the large-scale dataset collected from the two systems, respectively, a method is proposed to identify the relationship between the two systems. Three indicators are proposed to integrate the travel patterns of the two systems and quantify the relationship between the two systems.

(i) The relationship between DLBS and the metro can be defined as competition, connection, and complementation. DLBS trips mainly integrate with and complement the metro.

(ii) Corresponding to the three kinds of relationships, the relationship between the two systems can be characterized by the share of connection around the metro stations, the intensity of competition along the metro lines, and the demand for complementation outside the metro coverage area.

The findings from the case study of Shanghai suggest that both where the dockless bicycle-sharing trip takes place and when the trip happens significantly determine its relationship with the metro. Commute as a trip purpose might be a significant determining factor in the relationship between DLBS and the metro.

Our study provides a framework for analyzing the relationship between DLBS and other transportation modes. According to the research conclusion, the accuracy of the two systems can be recommended for the construction and management of stations, lines, and regions to improve the operation efficiency and service level of the two systems, promote the reasonable integration and development between the two systems, and effectively promote the overall optimization of the urban transportation system. We believe that not only the proposed methods but also the problem-solving ideas and management framework apply to other cities with different conditions.

However, this paper only uses longitude and latitude coordinates, which may not reflect street layout or terrain, in the distance calculation of shared single vehicle travel, which has certain limitations. In addition, it does not know more about the relationship between DLBS companies and the planning entity. It should continue to be discussed in future research.

Meanwhile, based on the spatiotemporal heterogeneity that we found for different categories of DBL trips, deeper research can continue, for example, what are the potential underlying mechanisms that present these spatiotemporal patterns. Hu et al. [43] fitted a set of generalized additive models considering marginal nonlinear interactions to the association between metro-related trips and the external environment in their study, including land use, socio-demographics, road design, transportation facilities, metro station characteristics, and DBS operator characteristics, which is of great relevance to our subsequent study. However, since the existing research data do not contain information on the built-up urban environment and individual attributes, very accurate modeling of regression and willingness models cannot be performed at this stage.

In future research, we will endeavor to obtain urban built environment data such as POI and conduct spatial regression models based on the known spatial heterogeneity. Secondly, we are already working on the design of a survey on travel intention of bike-sharing users, which will explore the degree of influence of weather environment, personal factors, cycling environment, subway environment, travel characteristics, environmental protection philosophy, and epidemic on people's willingness to use bike-sharing to connect, replace, and supplement the metro.

Data Availability

The dockless sharing bicycle data and metro automatic fare collection data used to support the findings of this study have not been made available because of the privacy policy.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 71734004) and the National Key R&D Program of China (No. 2018YFB1601100).

References

[1] C. S. Shui and W. Y. Szeto, "A review of bicycle-sharing service planning problems," *Transportation Research Part C: Emerging Technologies*, vol. 117, Article ID 102648, 2020.

[2] C. Xu, J. Ji, and P. Liu, "The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 47–60, 2018.

[3] M. Hatzopoulou, S. Weichenthal, H. Dugum et al., "The impact of traffic volume, composition, and road geometry on personal air pollution exposures among cyclists in Montreal, Canada," *Journal of Exposure Science and Environmental Epidemiology*, vol. 23, no. 1, pp. 46–51, 2013.

[4] Y. Lv, D. Zhi, H. Sun, and G. Qi, "Mobility Pattern Recognition Based Prediction for the Subway Station Related Bike-Sharing Trips," *Transportation Research Part C: Emerging Technologies*, vol. 133, Article ID 103404, 2021.

[5] Q. Yu, W. Li, D. Yang, and Y. Xie, "Policy zoning for efficient land utilization based on spatio-temporal integration between the bicycle-sharing service and the metro transit," *Sustainability*, vol. 13, no. 1, p. 141, 2020.

[6] 任逸帆 and 李开国, "共享单车对轨道接驳交通的影响分析. 交通与港航," 2018.

[7] J. Bachand-Marleau, B. H. Lee, and A. M. El-Geneidy, "Better understanding of factors influencing likelihood of using shared bicycle systems and frequency of use," *Transportation Research Record*, vol. 2314, no. 1, pp. 66–71, 2012.

[8] D. Zhuang, J. G. Jin, Y. Shen, and W. Jiang, "Understanding the bike sharing travel demand and cycle lane network: The case of Shanghai," *International journal of sustainable transportation*, vol. 15, no. 2, pp. 111–123, 2021.

[9] J. Song, L Zhang, L. Zhang, Z. Qin, and M. A. Ramli, "A spatiotemporal dynamic analyses approach for dockless bike-share system," *Computers, Environment and Urban Systems*, vol. 85, p. 101566, 2021.

[10] Y. Du, F. Deng, and F. Liao, "A model framework for discovering the spatio-temporal usage patterns of public free-floating bike-sharing system," *Transportation Research Part C: Emerging Technologies*, vol. 103, no. JUN, pp. 39–55, 2019.

[11] J. Bao, C. Xu, P. Liu, and W. Wang, "Exploring bikesharing travel patterns and trip purposes using smart card data and online point of interests," *Networks and Spatial Economics*, vol. 17, no. 4, pp. 1231–1253, 2017.

[12] Y. Zhang, T. Thomas, M. J. G. Brussel, and M. F. A. M. Van Maarseveen, "The characteristics of bike-sharing usage: case study in zhongshan, China," *International Journal of Transport Development and Integration*, vol. 1, no. 2, pp. 255–265, 2017.

[13] Y. Zhang, D. Lin, and X. C. Liu, "Biking islands in cities: an analysis combining bike trajectory and percolation theory," *Journal of Transport Geography*, vol. 80, p. 102497, 2019.

[14] R. B. Noland, M. J. Smart, and Z. Guo, "Bikesharing trip patterns in New York City: associations with land use, subways, and bicycle lanes," *International journal of sustainable transportation*, vol. 13, no. 9, pp. 664–674, 2019.

[15] K. Schimohr and J. Scheiner, "Spatial and temporal analysis of bike-sharing use in Cologne taking into account a public transit disruption," *Journal of Transport Geography*, vol. 92, no. 1, p. 103017, 2021.

[16] J. Mattson and R. Godavarthy, "Bike share in fargo, north Dakota: keys to success and factors affecting ridership," *Sustainable Cities and Society*, vol. 34, pp. 174–182, 2017.

[17] A. A. Campbell, C. R. Cherry, M. S. Ryerson, and X. Yang, "Factors Influencing the Choice of Shared Bicycles and Shared Electric Bikes in Beijing," *Transportation research, Part C. Emerging technologies*, vol. 67, pp. 399–414, 2016.

[18] K. Kim, "Investigation on the effects of weather and calendar events on bike-sharing according to the trip patterns of bike rentals of stations," *Journal of Transport Geography*, vol. 66, no. jan, pp. 309–320, 2018.

[19] W. L. Shang, J. Chen, H. Bi, Y. Sui, Y. Chen, and H. Yu, "Impacts of COVID-19 pandemic on user behaviors and environmental benefits of bike sharing: a big-data analysis," *Applied Energy*, vol. 285, Article ID 116429, 2020.

[20] S. Hu and C. Z. L. Xiong, "Examining spatiotemporal changing patterns of bike-sharing usage during COVID-19 pandemic," *Journal of Transport Geography*, vol. 91, no. 11, Article ID 102997, 2021.

[21] X. Chai, X. Guo, J. Xiao, and J. Jiang, "Analysis of Spatial-Temporal Behavior Pattern of the Share Bike Usage during COVID-19 Pandemic in Beijing," arXiv e-prints, 2020, https://arxiv.org/abs/2004.12340.

[22] D. Fuller, L. Gauvin, P. Morency, Y. Kestens, and L. Drouin, "The impact of implementing a public bicycle share program on the likelihood of collisions and near misses in Montreal, Canada," *Preventive Medicine*, vol. 57, no. 6, pp. 920–924, 2013.

[23] Y. Zhang and Y. Zhang, "Associations between public transit usage and bikesharing behaviors in the United States," *Sustainability*, vol. 10, no. 6, p. 1868, 2018.

[24] W. Zhu, Y. Pang, D. Wang, and H. Timmermans, "Travel behavior change after the introduction of public bicycle systems: case study in minhang district, Shanghai," in *Proceedings of the 92nd Annual Meeting Of the Transportation Research Board*, Washington, D.C, January 2013.

[25] E. W. Martin and S. A. Shaheen, "Evaluating public transit modal shift dynamics in response to bikesharing: a tale of two U.S. cities," *Journal of Transport Geography*, vol. 41, no. Dec, pp. 315–324, 2014.

[26] E. Fishman, S. Washington, and N. Haworth, "Bike share's impact on car use: evidence from the United States, Great Britain, and Australia," *Transportation Research Part D: Transport and Environment*, vol. 31, no. 2, pp. 13–20, 2014.

[27] D. Fuller, L. Gauvin, Y. Kestens, and P. L. Morency, "The potential modal shift and health benefits of implementing a public bicycle share program in Montreal, Canada," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 10, no. 1, pp. 66–6, 2013.

[28] K. B. Campbell and C. Brakewood, "Sharing riders: how bikesharing impacts bus ridership in New York City," *Transportation Research, Part A. Policy and Practice*, vol. 100, pp. 264–228, 2017.

[29] H. Kong, S. T. Jin, and D. Z. Sui, "Deciphering the relationship between bikesharing and public transit: modal substitution, integration, and complementation," *Transportation Research Part D: Transport and Environment*, vol. 85, Article ID 102392, 2020.

[30] L. Böcker, E. Anderson, T. P. Uteng, and T. Throndsen, "Bike sharing use in conjunction to public transport: exploring spatiotemporal, age and gender dimensions in Oslo, Norway," *Transportation Research Part A: Policy and Practice*, vol. 138, no. 1, pp. 389–401, 2020.

[31] Q. Yu, W. Li, and D. Yang, "Policy zoning for efficient land utilization based on spatio-temporal integration between the bicycle-sharing service and the metro transit," *Sustainability*, vol. 13, no. 1, p. 141, 2020.

[32] Q. Yan, K. Gao, L. Sun, and M. Shao, "Spatio-temporal usage patterns of dockless bike-sharing service linking to a metro station: a case study in Shanghai, China," *Sustainability*, vol. 12, no. 3, p. 851, 2020.

[33] M. Adnan, S. Altaf, T. Bellemans, A. U. H. Yasar, and E. M. Shakshuki, "Last-mile travel and bicycle sharing system in small/medium sized cities: user's preferences investigation using hybrid choice model," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 12, pp. 4721–4731, 2018.

[34] A. Sh, S. Hu, M. Chen, Y. Jiang, W. Sun, and C. Xiong, "Examining Factors Associated with Bike-And-Ride (BnR) Activities Around Metro Stations in Large-Scale Dockless Bikesharing Systems," *Journal of Transport Geography*, vol. 98, Article ID 103271, 2021.

[35] P. Jiménez, M. Nogal, B. Caulfield, and F. Pilla, "Perceptually important points of mobility patterns to characterise bike sharing systems: the Dublin case," *Journal of Transport Geography*, vol. 54, pp. 228–239, 2016.

[36] J. Zhao, J. Wang, and W. Deng, "Exploring bikesharing travel time and trip chain by gender and day of the week," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 251–264, 2015.

[37] 李配配 and 崔珩, "公共自行车与轨道交通的接驳与换乘研究. 交通科技," no. 1, p. 4, 2013.

[38] T. Ma and G.-J. Knaap, "Estimating the impacts of capital bikeshare on metrorail ridership in the Washington metropolitan area," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 7, pp. 371–379, 2019.

[39] Q. Yu, H. Zhang, W. Li, X. Song, D. Yang, and R. Shibasaki, "Mobile phone GPS data in urban customized bus: dynamic line design and emission reduction potentials analysis," *Journal of Cleaner Production*, vol. 272, p. 122471, 2020.

[40] Y. Qing and Y. Jian, "TransBigData: A Python Package Developed for Transportation Spatio-Temporal Big Data Processing, Analysis and Visualization," *Journal of Open Source Software*, vol. 7, no. 71, p. 4021, 2022.

[41] X. Ma, Y. Ji, Y. Jin, J. Wang, and M. He, "Modeling the factors influencing the activity spaces of bikeshare around metro stations: a spatial regression model," *Sustainability*, vol. 10, no. 11, p. 3949, 2018.

[42] Y. Ji, X. Ma, M. Yang, Y. Jin, and L. Gao, "Exploring spatially varying influences on metro-bikeshare transfer: a geographically weighted Poisson regression approach," *Sustainability*, vol. 10, no. 5, p. 1526, 2018.

[43] S. Hu, M. Chen, Y. Jiang, W. Sun, and C. Xiong, "Examining factors associated with bike-and-ride (BnR) activities around metro stations in large-scale dockless bikesharing systems - ScienceDirect," *Journal of Transport Geography*, vol. 98, p. 103271, 2022.

[44] Z. Wang and L. Y. Z. Cheng, "Spatiotemporal characteristics of bike-sharing usage around rail transit stations: evidence from beijing, China," *Sustainability*, vol. 12, no. 4, p. 1299, 2020.

[45] Y. Guo, L. Yang, Y. Lu, and R. Zhao, "Dockless bike-sharing as a feeder mode of metro commute? The role of the feeder-related built environment: Analytical framework and empirical evidence," *Sustainable Cities and Society*, vol. 65, p. 102594, 2020.

[46] W. Li, S. Chen, J. Dong, and J. Wu, "Exploring the spatial variations of transfer distances between dockless bike-sharing systems and metros," *Journal of Transport Geography*, vol. 92, p. 103032, 2021.

[47] D. Zhao, W. Wang, G. P. Ong, and Y. Ji, "An association rule based method to integrate metro-public bicycle smart card data for trip chain analysis," *Journal of Advanced Transportation*, vol. 2018, pp. 1–11, Article ID 4047682, 2018.

[48] Y. E. Hawas, M. N. Hassan, and A. Abulibdeh, "A multi-criteria approach of assessing public transport accessibility at a strategic level," *Journal of Transport Geography*, vol. 57, pp. 19–34, 2016.

[49] C. Wu and A. T. Murray, "Optimizing public transit quality and system access: the multiple-route, maximal covering/shortest-path problem," *Environment and Planning B: Planning and Design*, vol. 32, no. 2, pp. 163–178, 2005.

[50] S. T. Jin, H. Kong, and D. Z. Sui, "Uber, Public Transit, and Urban Transportation Equity: A Case Study in New York City," *Professional Geographer*, vol. 71, no. 2, pp. 315–330, 2019.

[51] 李孟冬, "步行可达性与地铁车站服务范围的研究," 生态文明视角下的城乡规划——*2008*中国城市规划年会论文集, 2008.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 7489897, 23 pages
https://doi.org/10.1155/2022/7489897

Research Article

Factors Influencing Users' Willingness to Adopt Connected and Autonomous Vehicles: Net and Configurational Effects Analysis Using PLS-SEM and FsQCA

Gang Li,[1] Yikai Liang,[2] Haiqing Wang,[3] Jiali Chen,[4] and Xiangbo Chang[2]

[1]*Shandong Computer Science Center, University of Technology, Jinan 250353, China*
[2]*School of Management Science and Engineering, Shandong University of Finance and Economics, Jinan 250014, China*
[3]*Graduate Business School, UCSI University, Kuala Lumpur 56000, Malaysia*
[4]*School of Management Engineering, Shandong Jianzhu University, Jinan 250101, China*

Correspondence should be addressed to Haiqing Wang; sunnyhaiqingwang@qq.com

Received 23 November 2021; Revised 5 January 2022; Accepted 16 April 2022; Published 2 May 2022

Academic Editor: Wenxiang Li

To accelerate the widespread adoption of connected and autonomous vehicles (CAVs) and take full advantage of CAVs' transportation safety, efficiency, and pro-environment, a deep understanding of CAVs acceptance is needed. However, little is known about the combined effects of factors influencing CAVs acceptance using traditional statistical methods. We developed an integrated model to explore how various antecedent factors work together on CAVs' acceptance. The symmetric (Structure Equation Modeling) and asymmetric (Qualitative Comparative Analysis) techniques were utilized for analyzing data from 362 Chinese. PLS-SEM assesses the net effect of each antecedent on CAVs' adoption, while fsQCA provides supplementary analysis by revealing the configurations of causal conditions associated with CAVs' adoption. PLS-SEM results show that perceived usefulness, perceived ease of use, and initial trust directly influence users' willingness to adopt CAVs, while perceived risk, social influence, and facilitating conditions do not. Meanwhile, automation, ubiquitous connectivity, structural assurance, and corporation reputation indirectly influence CAVs adoption, while environmental performance and technological uncertainty have no statistically significant indirect effect. Interestingly, fsQCA reveals five configurations resulting in a high level of CAVs' acceptance, and seven configurations leading to the negation of CAVs' acceptance. The complementary analysis results provide insights into both theory and practice.

1. Introduction

The rapid adoption of shared mobility services and electric vehicles, coupled with the prospect of driverless vehicles, has the potential to radically transform transportation around the world. Three revolutions in transportation (3 Revolutions), including electrification, automation, and shared mobility, stand as a promising alternative to improve transport efficiency and reduce costs, times, and carbon emissions [1]. Connected and autonomous vehicles (CAVs), as advanced technology of automation, have become the forefront of the vehicle market and attracted widespread attention. Various connected and autonomous technologies and systems, e.g., intelligent technology products, body electronic control products, and vehicle-mounted electronic devices, could make transportation safer (fewer accidents), cleaner (lower fuel consumption, emissions, and environmental pollutants), more efficient (reduced congestion and driving time), and more comfortable [2–7]. CAVs may bring not only substantial benefits but also tough challenges, e.g., users' safety and data safety [7]. Therefore, there is still much uncertainty in the transitions toward CAVs for transport. Thus, CAVs are slowly gaining market share, but less than expected, particularly under the Covid-19 pandemic. Recent surveys have shown that the level of consumers' actual purchases of CAVs is generally low. For example, a survey in China reported that 45% of the respondents were familiar with CAVs. However, 24% and 12% of respondents are

willing to purchase high and fully automated vehicles, respectively. Another survey including more than 8,500 consumers by Kantar, one of the world's leading data and consulting companies, shows that 40% of consumers in Europe and 32% of consumers in North America accept connected car technologies. Ordinarily, innovations sometimes meet resistance or refusal in the marketplace [4]. Any innovative technology is hard to become a reality without the adoption of its targeted customers [8]. Market penetration rates depend largely on public opinions regarding benefits, concerns, and the adoption of CAVs. Hence, it is important to understand consumers' willingness to adopt CAVs [9].

Besides the concept, benefits, and concerns of CAVs [10], plenty of researchers have identified determinant factors influencing users' willingness to adopt or purchase CAVs from different perspectives of primary technology adoption theories. On the one hand, most empirical researches identify the critical factors of CAVs adoption, drawing on a single theory, e.g., theory of planned behavior (TPB) [11, 12], behavioral reasoning theory [13], technology acceptance model (TAM) [2], unified theory of acceptance and use of technology (UTAUT) [4, 14], diffusion of innovation (DOI), behavioral reasoning theory [13], cognitive appraisal theory [15], and innovation resistance model [16]. As a single theoretical perspective is insufficient for comprehensive analysis and a better understanding of the adoption decision-making process [17], it is necessary to integrate various theories or additional factors according to the specific background to improve the model's explanatory power [8, 17–20]. On the other hand, the dominant methodology employed in empirical studies for analyzing the relationships between independent and dependent variables is the conventional symmetric-based approaches, e.g., multiple regression analysis [12, 14, 21–23], covariance-based structural equation modeling (CB-SEM) [11, 17, 24–28], and partial least square (PLS-SEM) [6, 16, 17, 29–32]. The traditional methodologies, which are only centered on the net effects of individual variables, are criticized for multicollinearity and symmetry [33–35]. At the same time, the relationship between antecedents and consequences, in reality, is highly asymmetric [36, 37]. Therefore, as a holistic approach, qualitative comparative analysis (QCA) has been recommended for dealing with the disadvantages based on causal complexity theory [38–42].

To fill these gaps, this study responds to answer the calls for applying integrated theoretical models and a holistic approach to understanding consumers' intention to accept CAVs. First, based on the detailed literature review related to studies on CAVs adoption, the primary models and constructs that have been extensively used in a majority of studies are selected as a theoretical foundation to develop a systematic and integrated model, including the factors that influence CAVs adoption, e.g., perceived usefulness, perceived ease of use, perceived risk, initial trust, social influence, facilitating conditions, and technological (automation, ubiquitous connectivity, environmental performance, and technological uncertainty) and environmental characteristics (structural assurance, corporation reputation). Subsequently, using a sample of 362 respondents in China, we apply not only symmetric (PLS-SEM) but also asymmetric (fsQCA) to analyze the net effect and combined effect of the above factors on CAVs adoption. As our theoretical models involve several factors from different models, it starts with assessing the net effects of each factor on CAVs' adoption by PLS-SEM. What is more, fsQCA is used to offer a deeper insight into the complex reality of CAVs' adoption that PLS-SEM cannot find. FsQCA can explain how the factors combine together to produce numerous causal equifinality pathways that could result in CAVs' adoption or the negation of CAVs' adoption (causal asymmetry). The results of PLS-SEM and fsQCA provide complementary insights, which are of double significance both in theory and practice.

2. Literature Review

2.1. The Definition and Acceptance of CAVs. From the view of the technology development pathway, intelligent vehicles can be chopped up into three solutions: connected vehicles (CV), autonomous vehicles (AV) (also called self-driving cars, driverless cars, or automated vehicles) [2, 3, 25, 43], and the fusion of the first two, namely, connected and autonomous vehicles [9], or intelligent and connected vehicles. As a node of the IoV system, CAV refers to a new generation vehicle equipped with advanced technologies and intelligent cooperative driving capabilities, aiming to make sure the safety, comfort, and efficiency of drivers.

The levels of automation (SAE International Standard) include no automation, driver assistance, partial, conditional, high, and full automation [44]. While the reality of fully autonomous cars on the road is still some way off, many CAVs have already existed. CAVs are believed to produce substantial benefits and shift people's lives in the short run by improving traffic safer, more efficient, and pro-environmental. Nevertheless, a broad range of CAVs issues still exist, e.g., the slightly higher price [10], and vehicle safety concerns in complex conditions, which could hinder CAVs' acceptance [11]. Therefore, to facilitate the widespread adoption of CAVs, many scholars with different disciplines and perspectives have studied the user's adoption aspects of CAVs and investigated what factors drive the user to accept and purchase a CAV. Table 1 provides a summary of previous empirical studies on CAVs acceptance.

Studies on consumer preferences toward CAVs can be divided into economic and psychological studies [59]. For economic studies, discrete choice modeling has been the prevailing approach for understanding various aspects of the demand for CAVs. Choice models try to capture decision-makers' preferences amongst a set of available alternatives [9]. The estimated model can be used for calculating CAVs' market share or adoption rate [60], calculating willingness-to-pay for CAVs [7, 61–66], or developing agent-based models or market penetration models to simulate market penetration when the overall demand is known [67].

For psychological studies, based on psychological or behavioral theories, many conceptual models have been proposed to study the adoption of different CAVs, e.g.,

Table 1: Empirical studies on CAVs acceptance.

Author	Theory	Types of CAVs	Methodology (collection & analysis)	Antecedent conditions (IV) and outcome (DV)
[22]	TAM	Autonomous vehicles	Online survey (318 participants in Korea) & multiple regression	PU*, PEOU*, SI, and FC* on intention to use AVs
[26]	UTAUT and DOI	Automated shuttles	Questionnaire (340 individuals in Germany) & SEM	PE, EE, SI, FC, compatibility (COM)*, trialability, trust*, and sharing* on behavioral intention (BI)
[45]	TAM and DOI	Autonomous vehicles	Survey (274 respondents in China) & SEM	PU*, PEOU*, relative advantage (RA)+, image+, COM+, result demonstrability+, visibility+, and trialability+ on BI
[13]	Behavioral reasoning theory	Autonomous vehicles	Online survey (849 individuals in China) & SEM	Attitude (ATU)*, reasons for (against) adopting Avs*, face consciousness+, need for uniqueness*, and risk aversion* on adoption intention
[15]	Cognitive appraisal theory	Autonomous vehicles	Online survey (362 responses in the United States) & CB-SEM	SI+, hedonic motivation (HM)+, trust+, PE*, perceived risk (PR)*, and emotions* on intentions
[46]	UTAUT2	Autonomous delivery vehicles	Online survey (501 German) & SEM	Trust in technology*, price sensitivity (PS)*, innovativeness*, PE*, HM*, SI*, and PR* on BI
[27]	TAM	Autonomous vehicles	Survey (340 participants in China) & SEM	ATU*, PU*, PEOU, knowledge of Avs*, PR*, public engagement, and face consciousness* on parents' intentions.
[32]	TAM	Automated vehicle	Web-based survey (647 drivers in China) & PLS-SEM	PU*, PEOU*, initial trust (IR)*, SI*, personality traits*, and sensation seeking* on BI.
[47]	Identity threat theory and identity control theory	Autonomous vehicles	Survey (353 consumers in Singapore) & SEM	Technology anxiety*, self-identity on perceived observability*, willingness to try, and intention to accept AV technology
[12]	TPB	Automated vehicle	Survey (505 Australian drivers) & linear regressions	ATU*, subjective norms (SN)*, and perceived behavioral control (PBC) on intentions to use
[48]	UTAUT2 and TPB	Shared autonomous vehicles	Survey (268 participants in Vietnam) & SEM	PE+, EE+, habit+, price value+, HM+, FC+, ATU*, subject norm*, and PBC* on BI
[28]	No specific theory	Autonomous vehicle	Survey (992 respondents in Hungary) & SEM	HM*, utilitarian motivation*, technological anxiety*, and data privacy concern on BI
[7]	No specific theory	Autonomous vehicles	Survey (1194 respondents in Florida) & SEM	Pro-technology service quality*, self-driving features*, driving assistance features*, pro-driving, trip privacy*, data privacy*, and technology concerns on AV adoption and WTP
[31]	TAM	Connected vehicles	Online survey (116 participants t) & PLS-SEM	PU, PEOU, ATU*, privacy concerns, privacy risk, trust in provider, information control, and social norm* on BI

Table 1: Continued.

Author	Theory	Types of CAVs	Methodology (collection & analysis)	Antecedent conditions (IV) and outcome (DV)
[49]	TAM	Autonomous vehicles	Online survey (316 participants)	PU, PEOU, ATU*, SI*, system characteristics*, and individual factors* on usage intention
[50]	TAM	Autonomous vehicles (AVs)	Survey (391 participants in Turkey) & SEM	PU*, PEOU*, trust*, and sustainability concerns* on BI
[6]	Social cognitive theory, TPB, prospect theory, and value perception theory	Fully autonomous vehicles	Survey (355 samples in Beijing, China) & PLS-SEM	Mass media+, social media+, self-efficacy*, SN*, PU*, and PR* on adoption intention
[51]	DOI and perceived value theory	Autonomous vehicles	Survey (526 residents in Seoul, Korea) & SEM	RA+, COM+, complexity+, trialability+, observability+, perceived value*, and trust* on public acceptance
[11]	TPB	Autonomous vehicles	Survey (526 residents in Seoul, Korea) & SEM	ATU*, SN*, behavioral control*, cognitive and emotive factors (RA+, COM+, complexity+, and HM+) on acceptance
[52]	UTAUT2	Autonomous delivery vehicles	Online survey (501 German) & SEM	PE*, EE, SI*, FC*, HM*, PS*, and PR* on BI
[8]	TAM	Autonomous shuttle services	Survey (700 respondents in Taiwan) & SEM	PU*, PEOU+, ATU*, trust*, and perceived enjoyment + on use intention
[25]	TAM and initial trust theory	Automated vehicles	Face-to-face survey (216 drivers in Shenzhen, China) & SEM	PU*, PEOU+, perceived safety risk+, perceived privacy risk, IR+, and ATU* on BI
[5]	No specific theory	Autonomous vehicles	Survey (742 Korean respondents) & PLS-SEM	Trust in technology*, perceived benefit (PB)*, PR*, BI, and willingness to pay
[53]	TAM	Autonomous vehicles	Online survey (313 Korean respondents) & PLS-SEM	RA+, psychological ownership*, self-efficiency*, PR*, PU*, and PEOU on intention to use
[54]	TAM	Autonomous electric bus	Online survey (268 passengers in Germany) & SEM	PU, PEOU, ATU*, individual differences (trust*, desire to exert control, privacy concerns, ecological awareness, and personal innovativeness), social impacts (image and SN*), and systems characteristics (perceived enjoyment, RA, and price evaluation* on intention to use
[17]	TAM and the life-oriented approach	Self-driving public bus	Online survey (268 passengers in Germany) & PLS-SEM	PU*, PEOU, ATU*, life choices, subjective well-being, factors of travel quality, and life domains on intention to use.
[55]	TAM	Automated vehicles	Online survey (1177 participants in Europe, China, and North America) & PLS-SEM	Environmental protection, innovativeness+, perceived enjoyment+, objective usability+, PU+, PEOU+, and ATU* on BI.
[2]	TAM	Autonomous electric vehicles	Online survey (470 respondents in China) & SEM	Environmental concern*, green PU*, and PEOU* on BI
[24]	Trust and TAM	Autonomous vehicles	Online survey (369 German participants) & SEM	Trust*, concern of giving up control*, PU*, PEOU, driving enjoyment*, and personal innovativeness* on the adoption intention of AVs

Table 1: Continued.

Author	Theory	Types of CAVs	Methodology (collection & analysis)	Antecedent conditions (IV) and outcome (DV)
[30]	No specific theory	Self-driving vehicles	Survey (1355 participants in Tianjin and Xi'an, China) & PLS-SEM	Demographic (familiarity, age*, gender, education*, and income*) and psychological factors (PB*, PR*, perceived dread*, and trust*) on WTP
[9]	DOI theory and agent-based simulation modeling	Connected autonomous vehicles	Survey (327 employees of the University of Memphis) & simulation modeling	Price reduction*, mass communication (marketing), and peer-to-peer communication (word-of-mouth)* on CAV market share
[4]	UTAUT	Autonomous car	Online survey (241 respondents in France) & SEM and multi-group analysis	PE*, EE*, SI*, and consumer innovativeness* on purchase intention
[21]	TAM	Autonomous driving	Web-based survey (483 respondents in Greece) & multiple regression analyses	PU*, PEOU*, perceived trust*, and SI* on BI to have AVs
[14]	UTAUT	Automated road transport systems	Survey (315 participants in Greece) & Hierarchical multiple regression	PE*, EE, SI*, FC*, and HM* on BI
[56]	No specific theory	Fully autonomous vehicles	Online survey (482 participants from the United States) & principal component analysis and SEM	ATU, social norm, trust, COM*, and system effectiveness* on BI
[57]	UTAUT	Automated road transport systems	Survey (349 respondents from France and Switzerland) & Hierarchical multiple regression	PE*, EE*, SI*, and BI
[16]	Innovation resistance model	In-vehicle infotainment (IVI) systems	Online survey (1070 samples in Korea) & PLS-SEM	Technographics, SN, prior similar experience, PU*, perceived complexity*, PR*, and resistance on intention
[58]	UTAUT	Self-driving vehicles	Online survey (556 residents of Austin, Texas) & Ordinal regression model	PE*, EE*, SI, perceived safety*, anxiety*, ATU about technology, desire for control*, technology use*, and technology acceptance on intent to use
[29]	TAM and trust theory	Autonomous vehicle	Online survey (552 drivers) & PLS-SEM	System transparency+, technical competence+, situation management+, trust*, PU*, PEOU*, PR, external locus of control*, and sensation seeking on BI

Note. Perceived usefulness, PU; perceived ease of use, PEOU, perceived benefit, PB; perceived risks, PR; performance expectancy, PE; effort expectancy, EE; social influence, SI; facilitating conditions, FC; hedonic motivation, HM; price sensitivity, PS; relative advantage, RA; compatibility, COM; attitude, ATU; subjective norms, SN; perceived behavioral control, PBC; behavioral intention, BI; "+" shows a significant impact on other variables that are not the dependent variables, and "∗" indicates a significant impact on BI.

autonomous delivery vehicles [46], automated shuttles [26], autonomous electric buses [54], shared autonomous vehicles [48, 68, 69], connected vehicles [9, 31], and Robo-taxi services [23]. Various factors have been empirically found to have a significant influence on accepting CAVs among multiple regions, e.g., Korea [5, 11, 16, 22, 53], Germany [17, 24, 26, 31, 46, 49, 54], China [2, 6, 13, 25, 27, 30, 32, 45], United States [7, 9, 15], Singapore [47], Australia [12], Vietnam [48], Hungary [28], Turkey [50], Taiwan [8], France [4], and Greece [14, 21].

In terms of theoretical models, previous studies have mainly attempted to identify the antecedents of CAVs' acceptance using a single theory, e.g., TAM [2, 8, 21–23, 27, 31, 32, 49, 50, 53, 54], TPB [11, 12], UTAUT [4, 14, 46], DOI, behavioral reasoning theory [13], cognitive appraisal theory [15], and innovation resistance model [16]. In order to offer a comprehensive comprehension of CAVs acceptance [17], some scholars have suggested that the combination of various theories or additional factors according to certain contexts [8, 17] can increase the model's explanatory power.

Increasingly scholars have used the integrated model to study CAVs' acceptance, e.g., the combination of TAM and initial trust theory [8, 24, 25, 29], or DOI [45], as well as the combination of UTAUT and DOI [26], or TPB [48]. Hence, in this study, we answer this call and integrate several primary theories and constructs to identify the comprehensive factors and improve the model's explanatory power.

In terms of methodology, most empirical studies have employed the conventional symmetric-based methods, e.g., the regression model [14, 21–23], structural equation modeling (SEM) [2, 4, 8, 11, 13, 15, 24–28, 45–48, 50, 54], and partial least square (PLS) [6, 16, 17, 29–32, 49, 53], to investigate the effects between variables. However, these methods emphasize the net effect, while ignoring possible asymmetric relations between variables in complex contexts, resulting in the correlation coefficients and significance that may differ in various models. The outcome is often produced from various combinations of antecedents in a real-life context, rather than individual ones [70].

2.2. Technology Adoption Model. In the information system (IS) domain, technology adoption has been one of the extensively enduring topics [71]. Various theories and models, e.g., theory of reasoned action (TRA) [72], technology acceptance model (TAM) [73], theory of planned behavior (TPB) [74], unified theory of acceptance and usage of technology (UTAUT) [75], initial trust model (ITM) [76], diffusion of innovation theory (DOI) [77], have aimed to identify, predict, and explain variables influencing adoption behavior at the individual level to accept and use technological innovations [78].

Among these theoretical models, TAM, initially proposed by Davis, may be possibly one of the most extensively applied models in understanding IT adoption and usage processes. Two specific constructs in TAM, i.e., perceived ease of use and usefulness, are significant for behavior intention to accept innovation. Most of the studies have found TAM as the valid, robust, and most dominant model to explain technology adoption at the individual level. Therefore, TAM is broadly applicable for exploring the users' acceptance across a wide variety of contexts, e.g., parking Apps [79], telemedicine services [80], mobile library Apps [81], and virtual reality [82, 83]. As seen in Table 1, TAM serves as an adequate theoretical foundation, which is the most used in CAVs adoption studies [8, 22, 27, 31, 32, 45, 50].

2.3. Conceptual Model and Hypotheses. Although TAM is a robust and concise model for effectively explaining the adoption of various technologies [31, 84], it may be too general to be applied to every technology in its original form (i.e., it cannot adequately reflect the specific effects of technical and contextual factors that may change user adoption) [21, 80–82]. Therefore, consideration should be given to integrating other adoption models or additional factors [8, 80]. To fully explain consumer's acceptance of CAVs, we propose an integrated framework of CAVs acceptance that combines TAM with other constructs from perceived risk theory, initial trust theory, and UTAUT, and with constructs specific to technological and environmental characteristics, as a single perspective is not sufficient for a comprehensive study [17]. Specifically, based on TAM selected as the basic theoretical framework for adaptability in the CAVs' context [16, 25], we merged initial trust and perceived risk in the initial TAM, on account of the new technology's uncertainties. Furthermore, both the two-sidedness of the innovative technologies, that is, the technological advantages (unique characteristics such as automation and ubiquitous connectivity) and the possible negative consequences (technology uncertainty) of CAVs should be investigated in parallel. In addition, environmental characteristics such as social influence, structural assurance, and corporation reputation are also considered. Moreover, CAV requires drivers to possess certain capabilities and resources. Therefore, facilitating conditions need to be considered. Considering the specific attributes of CAVs, the integrated model is shown in Figure 1.

2.4. Perceived Usefulness and Ease of Use of TAM. Perceived usefulness believes that employing a particular innovation could improve one's work performance [73]. Perceived usefulness resembles the notion of performance expectancy in UTAUT [85], and relative advantage in DOI. It has been extensively examined in previous studies on IT adoption, e.g., virtual reality [82, 83], parking App [79], telemedicine services [80], mobile library App [81], autonomous vehicles [2, 6, 21, 22, 24, 25, 27, 32, 45, 50, 53], Robotaxi [23], and autonomous shuttle [8]. Consumers are inclined to be willing to purchase innovative technologies if they possess uniqueness over existing technologies. Compared with traditional cars, CAVs can replace the driver for some or all of the driving tasks, and have several potentially helpful functions and benefits. For example, in 2016, motor vehicle–related crashes on U.S. highways claimed 37,461 lives. US Dept of Transportation research shows that 94% of serious crashes are due to human errors. CAVs can monitor the environment continuously and let drivers stay fully informed of the current driving situation, making up for lapses in driver attention, which can improve vehicle safety and eliminate crashes by the combination of passive safety, active safety, and automated driver assistance. What is more, CAVs can plan the best route, forecast traffic jams, and predict the best time to schedule a trip, which allows drivers to get an energy-optimized driving experience, and reduce fuel consumption. In addition, the drivers of CAVs will be able to solve their personal and business tasks right on the go, without the need to be distracted from driving. CAVs with a self-parking feature will also allow the drivers to save a lot of time since, after the arrival, the car will part itself independently, and the owner will be able to proceed with tasks [6, 54]. Many researchers have suggested technological usefulness could improve consumers' satisfaction and willingness to pay [53].

H1: Perceived usefulness positively influences users' willingness to adopt CAVs.

Perceived ease of use means one believes that employing a particular innovation would be effortless [73]. It is analogous to effort expectancy from UTAUT [85], and complexity from

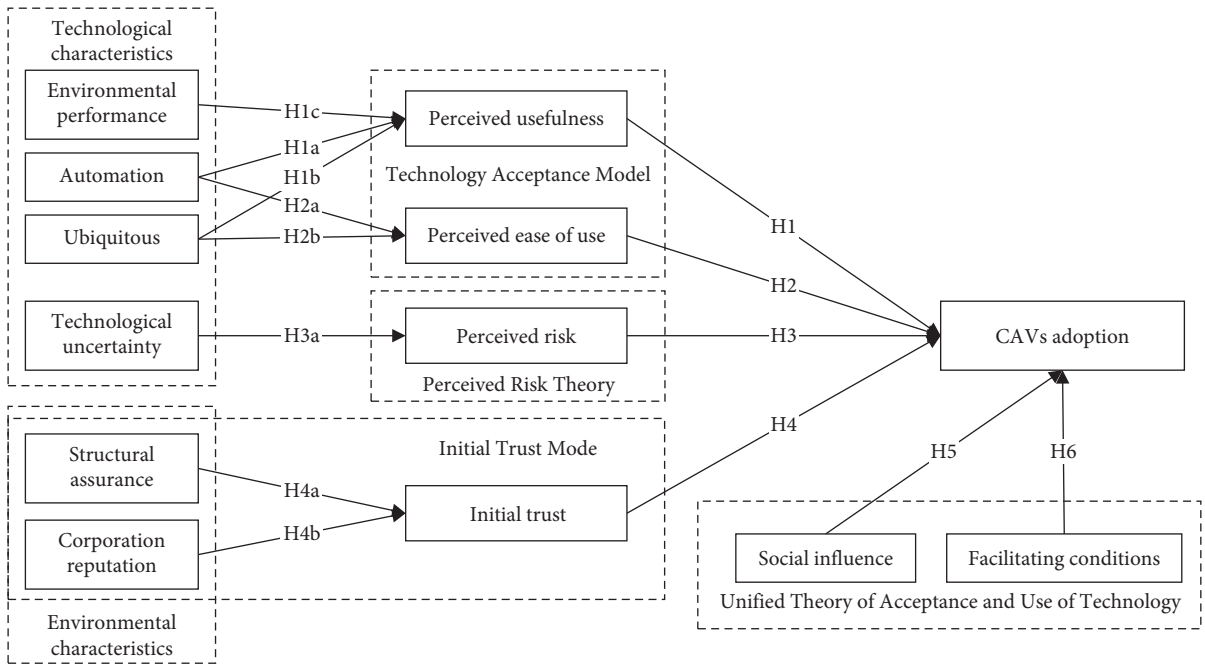


Figure 1: The research model.

DOI, which has been considered as a key precursor to the adoption of various technologies, e.g., virtual reality [82, 83], parking Apps [79], telemedicine services [80], mobile library App [81], Robo-taxi [23], and autonomous vehicles [2, 8, 21, 29, 32, 45, 50, 86]. CAVs benefit from the human-machine interface and artificial intelligence (AI), which may reduce the complexity of CAVs, and make it easier for people to operate CAVs. For example, several carmakers offer automated driving systems that allow for hands-free driving, including Tesla, GM, Ford, and Mercedes-Benz. CAVs are expected to become an option for people, particularly for impairment and disabilities. Managed with voice commands (e.g., Apps and settings, car controls, climate controls, navigation, phone, and media), they achieve better mobility with the car making all the necessary decisions. The easier it is to handle the key features and services of CAVs, the more possibility of adopting and purchasing CAVs [11]. Thus, the following hypotheses are formulated.

H2: Perceived ease of use positively affects users' willingness to adopt CAVs.

Automation refers to the performance by machine agents of functions previously performed by humans [87]. The widespread popularity of automation technology in intelligent vehicles has been accepted as wider connectivity improves the economy and simplicity of the vehicle. For CAVs, the automation defined by NHTSA is a vehicle capable of operating at least some mission-critical control without the need for human intervention [3]. Compared with traditional, fully manually controlled vehicles, automation characteristics possess several benefits for overcoming the inconvenience of controlling, managing, and monitoring vehicles [25, 68]. For example, CAV has enormous potential in decreasing accidents caused by human error and improving vehicle safety with the help of collision avoidance applications [3, 11, 43, 54]. In addition, highly automated driving and adaptive cruise control reduce driving workload and improve comfort on longer trips. Riding in CAVs with full automation takes drivers off driving tasks and allows them to choose their preferred leisure or pick-up activities unrelated to driving [88]. Moreover, CAVs offer new travel options for special groups, e.g., the older people and the disabled [68, 89], providing them with a greater extent of mobility [9]. Numerous studies have found a significant relationship related to automation with consumers' attitudes to CAVs, e.g., smart homes [87] and CAVs [9].

H1a: Automation positively affects the perceived usefulness of CAVs.

H2a: Automation positively affects the perceived ease of use of CAVs.

Ubiquitous connectivity means continuous Internet connection, which is increasingly needed by users [90]. As sensors, computing, AI, and bidirectional communication technologies are embedded in CAVs, IoT-based services in CAVs become ubiquitous. Ubiquitous connectivity plays a significant role in the safety of vehicles and passengers. CAVs could synchronize information with intelligent transportation infrastructures through wireless communication technology by capturing traffic and environmental conditions to avoid traffic jams [9]. Moreover, ubiquitous connectivity makes users to be "always connected" or "always-on," offering free access to information services, e.g., in-vehicle infotainment [16]. It is validated as a predictor of willingness to use technology applications, e.g., smartphone-based SNS [90].

H1b: Ubiquitous connectivity positively affects the perceived usefulness of CAVs.

H2b: Ubiquitous connectivity positively affects the perceived ease of use of CAVs.

Environmental performance, a nonfunctional characteristic of the product, means that the product will contribute to environmental sustainability and help reduce environmental pollution [91]. With the idea of "green and low-carbon life" embedded deeply in human nature, consumers with environmental concern [92, 93] and ecological awareness [94], are more likely to deviate from their present behavior to a more environment-friendly behavior [92] such as the purchase behavior of environmentally sound products (e.g., plugin hybrid electric vehicles, and electric vehicles [93]). CAVs could utilize real-time traffic data to efficiently navigate to a prearranged destination by better routing planning and more efficient vehicle operation (optimized braking and acceleration) [54], and thus avoid traffic congestion, reduce additional fuel consumption, and generate fewer pollutants [9, 25, 43]. This characteristic may be considered a significant factor in the willingness to buy for consumers with environmental awareness.

H1c: Environmental performance positively affects the perceived usefulness of CAVs.

2.5. Perceived Risk. Perceived risk is the extent to which users are uncertain and problematic about the possible negative consequences of using innovation, according to the perceived risk theory [95]. It is negatively related to IT adoption across multiple contexts, e.g., mobile payment, and autonomous vehicles [5, 6, 13, 15, 25, 27, 46, 53]. Despite their potential, CAVs are surrounded by risks and concerns, e.g., users' safety and data safety. For users' safety, with the progress of CAVs technology, consumers have become more familiar with the potential safety because CAVs would not get distracted or drive when tired; however, they still have expressed more concerns related to safety when systems fail or drive in complex road conditions. No carmaker offers fully autonomous vehicles, though, but only capable of navigating marked highways in ideal, safe situations. For example, more than 35% of all Autopilot crashes occur when the Tesla vehicle is rear-ended by another vehicle. For data safety, by adding network connectivity, there can be risks with security, privacy, and data analytics and aggregation due to the large volume of data being accessed and shared when systems connect to the Internet. For example, Tesla has collected vast amounts of data from a fleet of more than 1 million cars around the world for data analytics, crash simulations, and further tests, which may lead to privacy disclosure, e.g., drivers' in-vehicle experiences, activities, location, or history. Therefore, consumers are naturally going to be risk-averse, and these risks would be regarded as a significant obstacle to CAVs adoption [6, 9]. Stemming from the above arguments, the following hypothesis was suggested.

H3: Perceived risk negatively impacts users' willingness to adopt CAVs.

Technology uncertainty involves the unpredictability of technological development, the technical environment, and the technology's consequences [96]. When users perceive uncertainty concerning the technology, they tend to overestimate the likelihood of potential losses [97]. Uncertainty perception would make them perceive more risks and technology anxiety, then dampen their intentions to utilize innovation [47]. Therefore, the inherent uncertainty of CAVs when operating in complex conditions and unclear legal liability of traffic accidents can increase consumers' perceived risk of CAVs [68].

H3a: Technology uncertainty positively affects the perceived risk of CAVs.

2.6. Initial Trust. Initial trust means trustor trusts in an unfamiliar trustee without credible, meaningful information about each other [31]. Numerous studies have investigated initial trust plays a major role in affecting the acceptance of different technologies, e.g., parking App [79], telemedicine services [80], autonomous vehicles [5, 24, 25, 29, 30, 32, 46, 50], and autonomous shuttle [8, 17, 26, 32, 54, 86]. In the early days of the market of CAVs, enough trust of potential consumers could increase the perceived benefits of successfully using CAVs, and conquer the perceived risks and uncertainties of employing CAVs [25, 87].

H4: Initial trust positively has a positive influence related to users' willingness to adopt CAVs.

Structural assurance means the infrastructure supporting technology use [98]. It refers to individuals ensuring that adequate structures and supports, e.g., guarantees, regulations, commitments, or other procedures, are in place to ensure an innovative technology's successful use. In the context of mobile services, an individual with a high degree of network-related structural assurance will assume that legal and technical Internet protection, such as data encryption, can protect personal privacy, identity, or money [99]. Since the CAVs market is still in its early stages, many surveys report consumer concerns about the unclear legal liability of traffic accidents, and the safety of CAVs might be compromised because of online intrusion and hacking thefts [68]. For example, no carmaker is willing to accept liability for accidents that happen during self-driving sessions, with the notable exception of Mercedes-Benz, which will take full legal responsibility for its cars when its Level 3 autonomous driving system is active. Especially in the absence of direct experience, such structural assurance can help individuals develop confidence in innovative technologies, thereby promoting trust in a specific technology [5], especially safety, privacy, and security issues of CAVs [21].

H4a: Structural assurance positively affects users' trust in CAVs.

A corporation's reputation reflects consumers' perceptions of its ability to deliver products and services effectively, the organization's credibility, and business engagement's reliability [98]. It plays a critical part in forming confidence and willingness to use offered information services [99]. Using big data analysis of Chinese consumers' behavior, nationality, reputation, and brand ranking are the determinants of consumers' consideration and preferences for electric

vehicles [100]. Motor corporation reputation can increase consumers' recognition of a newly introduced product and service and help maintain confidence in future purchases, especially for novice consumers. For example, Tesla has built its reputation for quality and customer support, and has the most loyal customers. When Tesla was hit with negative publicity involving the fatal accidents, CEO Musk was quick to address the issue and salvage its reputation via social media channels. Many studies find that firm reputation is a significant antecedent in users' trust [99]. It strongly influences the willingness to use innovative technology.

H4b: Corporation reputation of motor positively affects the user's trust in CAVs.

2.7. Social Influence and Facilitating Conditions of UTAUT. Social influence is defined as an individual feels the importance that the others believe he or she should use the new technology [85]. Previous scholars have revealed the impact of social influence on IT adoption in different contexts, e.g., AI [101], mobile banking, autonomous vehicle [3, 4, 21, 22, 32, 46], Robo-taxi [23], and automated public transport. Humans are essentially social [11], so they are susceptibly influenced by social networks [17], and social norms [12,31,48], even if making decisions alone. For CAVs, that is a disruptive innovation in an emerging market, and while it will be a while before we have fully AVs on our road, many of them already exist. Therefore, increasingly consumers will follow the mainstream. Moreover, the popularity of CAVs among friends and family would result in a higher intention to accept and purchase CAVs [4].

H5: Social influence positively influences users' willingness to adopt CAVs.

Facilitating conditions means users' awareness of the resources and supports available to carry out a certain behavior [85]. Willingness to utilize innovative technologies could be influenced by facilitating conditions in varying contexts, e.g., AI [101], mobile banking [99], telemedicine services [80], autonomous vehicles [22], and automated public transport [14]. Autonomous vehicle technology as a developing technology requires R&D expenses, which means that CAVs are sold at a premium. The unaffordable price of purchasing CAVs might be one major obstacle to users' decision to utilize CAVs [23,46,54,91]. Moreover, some functions of CAVs may require users to have certain knowledge. If users have the necessary money and operational knowledge, they will accept or purchase a CAV.

H6: Facilitating conditions are positively related to users' willingness to adopt CAVs.

3. Methodology

3.1. Research Design. We employ a survey to gather empirical data from potential consumers in China. Based on the theoretical models, our quantitative questionnaire was designed and validated. The measurement model of PLS-SEM is used to examine the instruments. The structural model of PLS-SEM is used to test our model and hypotheses, and provide symmetrical "net effect" explanations [102]. Besides, considering the limitations of symmetric statistical approaches and the occurrence of multiple realities (i.e., complex causality), the same data were calibrated and analyzed by fsQCA to explore the cause-effect relations between antecedent conditions and outcome, and to provide a holistic perspective of the interrelationships that jointly influence CAVs acceptance. The sections below describe the instruments development, data collection, and the reasons why to choose PLS-SEM and fsQCA.

3.2. Instruments Development. To improve content validity [73,103], the scales are adapted from prior literature and modified to match CAVs' context. The translation and back-translation method is employed to make sure that the meaning of the original language statements was retained to realize the semantic equivalence [22]. We measured the constructs using a five-point Likert scale (see Table 2).

The instrument was refined by revising ambiguous items during a pre-test to improve the understandability, with the involvement of several postgraduate students. Afterward, a pilot sample including 30 undergraduate students was conducted to confirm the scales' reliability, validity, and translational equivalence.

3.3. Data Collection. A self-administered online survey was conducted to collect data using Wenjuanxing (http://www.wenjuan.com), which is a professional online questionnaire survey and voting platform, in China from 22 to 26 June 2020. The questionnaire includes short explanations of the definition and functions of CAVs to help respondents understand this concept. The questionnaire was distributed randomly to actual respondents by Wenjuanxing. In total, 420 participants participated in our survey. Excluding invalid questionnaires with uncompleted data or identical answers, 362 questionnaires were deemed valid for the study. The detailed demographic profile is shown in Table 3.

As shown in Table 3, the gender distribution for the sample was comparable with the population distribution for China (i.e., 51.24% of the population are male). Age is biased towards young people, and education was slightly skewed toward higher educational attainment. It is mainly because of the online survey method which is mostly used by young people and highly educated people. For CAVs services most frequently used, unconscious pay is the most frequently used (54.70%). One probable reason is that electronic toll collection (ETC) systems are operating today for the turnpike covering 29 provinces in China, which reduces congestion and improves the efficiency of revenue collection while passing the highway toll station. Another is that increasingly parking lots have been connected to third-party payment (e.g., Ali pays and WeChat pay) to allow unconscious pay, so that users can pay the parking fee through unconscious pay and enjoy a convenient and efficient smart parking experience. The next is the connectivity-based intelligent services, i.e., human-machine interaction (46.41%), intelligent navigation (30.11%), and in-vehicle infotainment (22.38%). In

Table 2: Measurement scale.

Constructs	Items	Description	References
Automation	AU1	CAVs could help the drivers voluntarily without human intervention.	[87, 104]
	AU2	CAVs offer auto-adjusted control.	
	AU3	I could control CAVs through a simple operation.	
Ubiquitous connectivity	UC1	CAVs can assist me in being well informed about my vehicle.	[90, 105, 106]
	UC2	CAVs can allow me to obtain useful information anytime.	
	UC3	CAVs can help me to get information and monitor my vehicle regardless of where I am.	
Environmental performance	EP1	CAVs will contribute to environmental sustainability.	[2, 91]
	EP2	CAVs will help to reduce environmental pollution.	
	EP3	CAVs are important to save natural resources.	
Perceived usefulness	PU1	CAVs could improve my driving productivity.	[6, 29]
	PU2	CAVs could increase my driving performance.	
	PU3	CAVs could enhance my driving effectiveness.	
	PU4	Overall, I find CAVs very useful.	
Perceived ease of use	PEOU1	I realize that learning to operate CAVs would be easy.	[25, 29]
	PEOU2	I could easily understand how to make CAVs do what I wanted to do.	
	PEOU3	I do not have difficulty in interacting with CAVs.	
Technological uncertainty	TU1	The wireless network of CAVs is unstable.	[97, 107]
	TU2	The security of CAVs is questionable.	
	TU3	The technologies related to CAVs are undeveloped.	
Perceived risk	PR1	I fear for the general safety of CAVs.	[5, 6, 25, 84]
	PR2	I am more concerned that malfunctions of CAVs will lead to an accident.	
	PR3	I 'm afraid CAVs may steal too much personal data from me privately.	
	PR4	I'm worried CAVs may misuse my personal data for commercial purposes without my authorization.	
	PR5	I'm worried CAVs will share my personal data with other entities without my authorization.	
Structural assurance	SA1	I believe that advances in intelligence, communication, encryption, and other technologies make it safe for me to use CAVs.	[99, 108]
	SA2	I am confident that legal and technological structures are sufficient to protect me from the problems associated with using CAVs.	
	SA3	Mobile communication is a robust and safe environment where to use CAVs.	
Corporation reputation	CR1	Motor corporation of CAVs is reliable.	[87, 99, 108]
	CR2	The motor corporation keeps promises and commitments.	
	CR3	The motor corporation keeps consumers' best interests in mind.	
	CR4	I feel confident in the brand of the motor corporation.	
Initial trust	IT1	CAVs are very dependable.	[25, 108]
	IT2	CAVs are very reliable.	
	IT3	In general, I could trust in CAVs.	
Social influence	SI1	People who matter to me think that I should use CAVs.	[14, 85, 109]
	SI2	People who influence my behavior think that I should use CAVs.	
	SI3	People whose opinions that I appreciate suggest that I use CAVs.	
Facilitating conditions	FC1	I have the resources (budget) necessary to use CAVs.	[14, 85, 109]
	FC2	I have the knowledge necessary to use CAVs.	
	FC3	When there are difficulties with CAVs, the assistance of specific people can be provided.	
Users' willingness to adopt CAVs	AD1	I am willing to accept and buy CAVs.	[6,14,85]
	AD2	I predict that I will accept and buy CAVs in the future.	
	AD3	I plan to accept and buy CAVs.	

making CAVs, manufacturers have introduced human-machine interaction and in-vehicle infotainment (IVI) systems into their major product lines, which provide a variety of information and entertainment services, including vehicle-specific information, navigation, and much more [16]. The proportions of the respondents that use CAVs services of the lower automation level are adaptive cruise control (20.17%), automatic parking assist (15.47%), and collision avoidance system (14.09%), respectively. These advanced driver-assistance systems are designed to increase the safety of driving a vehicle, which is great for applications like blind-spot monitoring, lane-keep assistance, and forward-collision warning. For the high automation level, only 2.49% of the respondents have used self-driving, the smallest

Table 3: Demographic profile.

Characteristics		n	%
Gender	Male	186	51.38
	Female	176	48.62
Age	18–25	121	33.43
	26–30	66	18.23
	31–40	131	36.19
	41–50	32	8.84
	More than 51	12	3.31
Education	High school and below	8	2.21
	Junior College	16	4.42
	Undergraduate	133	36.74
	Graduate (Master)	117	32.87
	Graduate (Ph.D.)	86	23.76
Occupation	Senior manager	55	15.19
	Professionals	93	25.69
	Civil servant	3	0.83
	Company employee	50	13.81
	Service worker	3	0.83
	Labor	2	0.55
	Private entrepreneurs	8	2.21
	Self-employed	10	2.76
	Student	118	32.60
	Unemployed	7	1.93
	Other	13	3.59
Monthly household income (CNY)	Less than 3001	22	6.08
	3001–5000	45	12.43
	5001–10000	110	30.66
	10001–15000	80	22.10
	15001–20000	47	12.98
	20001–30000	34	9.39
	More than 30000	23	6.35
License holder	Yes	318	87.85
	No	44	12.15
Purchase experience	Yes	193	53.31
	No	169	46.69
Number of cars owned by the household	0	54	14.92
	1	206	56.91
	2	91	25.14
	3	8	2.21
	>3	3	0.83
CAVs service most frequently used	Automatic parking assist	56	15.47
	Adaptive cruise control	73	20.17
	Collision avoidance system	51	14.09
	In-vehicle infotainment	81	22.38
	Human-machine interaction	168	46.41
	Intelligent navigation	109	30.11
	Unconscious pay	198	54.70
	Self-driving	9	2.49
	Other	79	21.82

proportion in our study. This suggests that the wide used of self-driving is not as easy as expected.

Common method bias (CMB), in the context of PLS-SEM, is a phenomenon that is caused by the measurement method used in an SEM study, and not by the network of causes and effects in the model being studied. CMB is evaluated through a full collinearity assessment approach recommended by Kock [110]. Kock notes that if all variance inflation factors (VIF) resulting from a full collinearity test are equal to or lower than 5 (3.3), the model with measurement error can be considered free of CMB. Since all VIFs are below 5, the probability of CMB could be excluded.

3.4. Data Analysis Approach. Data were analyzed by PLS-SEM for estimating the measurement and structural model [103], as well as fsQCA for revealing the configurations of the antecedent conditions on outcome [38].

3.4.1. PLS-SEM Approach. Compared to other approaches, PLS-SEM provides great flexibility in modeling (e.g., complex models and formative constructs) and data requirements (e.g., small samples and nonnormally distributed data) [103]. Therefore, PLS-SEM has been widely used in various study domains, e.g., organization management, information management [111], and transportation management [53].

Following the recommendations [103], our research aims to identify key "driver" constructs, and our proposed structural model is complex. Therefore, we select the PLS-SEM approach through the specific software SmartPLS 3.2.9.

3.4.2. QCA Approach. QCA, originally developed by Ragin [112], is a methodology of in-depth analysis of the causal contribution of different conditions to an outcome of interest, based on set theory [113]. QCA can distinguish various complex forms of causation, e.g., configurations of causal conditions (not just single causes), equifinality (multiple causal pathways are leading to the same outcome's occurrence) [114], multifinality (identical conditions could generate different outcomes), and causal asymmetry (the causes of failure may not simply be the absence of the cause of success) [38, 115].

In everyday life, the outcome often results from different combinations of antecedent conditions, rather than any one individual condition, in the context of great causal complexity [116]. FsQCA is particularly suited to analyzing causal processes, providing a configurational understanding of how causes combine to produce outcomes, and dealing with significant causal complexity [115]. Therefore, fsQCA is very appealing to many researchers and is widely used in different contexts to assess cause-effect relations, e.g., strategy and organization management [115] and information management [117].

PLS-SEM and fsQCA rely on distinct principles. PLS-SEM is a variable-oriented approach that verifies each independent variable's net effect and significance on the dependent variable by a series of regression analyses [37, 118]. It does not answer what variable is sufficient or necessary for a certain outcome. Conversely, fsQCA is a case-oriented approach that analyzes the causal contribution of different conditions. FsQCA is a complementary analysis suitable for PLS-SEM when effects due to unobserved heterogeneity are detected, as it explains how factors work together in producing an outcome [117]. Thus, many studies empirically test the proposed models, both employing a symmetrical approach with PLS-SEM and an asymmetrical approach with fsQCA to analyze the casual and outcome conditions in complex situations, e.g., technology adoption [44, 119], organizational learning [102], and mobile shopping [37].

4. Results

4.1. Results of PLS-SEM Analysis. Based on the two-step process of PLS-SEM assessment recommendation [103], the measurement model is adopted to examine the reliability and validity of our instrument, and the structural model is employed to estimate the model and hypotheses.

4.1.1. Measurement Model. A reflective measurement model assessment includes indicator and internal consistency reliability, and convergent and discriminant validity [103].

Indicator reliability, obtained from squaring outer loadings of reflective constructs, clearly describes the relationship between the latent variable and its measures [111]. Table 4 shows that the constructs' outer loadings are greater than 0.708 [103], thus suggesting acceptable indicator reliability.

Internal consistency reliability is assessed by Cronbach's alpha (CA), ρ_A, and composite reliability (CR) [103,120]. Table 5 suggests that all criteria are above 0.7, thus indicating good measurement reliability.

Convergent validity is examined according to the average variance extracted (AVE) [120]. Table 5 shows that AVE is greater than 0.5, thus indicating that the measurement shows good convergent validity [103].

Discriminant validity is evaluated by Fornelle–Larcker criterion [120]. Table 6 indicates that AVE's square root is greater than the correlations between variables, confirming adequate discriminant validity.

4.1.2. Structural Model. Standard assessment criteria of the structural model contain R^2, and the significance and relevance of the path coefficients [103,111].

Before assessing the structural model, multicollinearity must be tested through the variance inflation factor (VIF). Table 5 shows that all VIF values are between 1 and 4.11, less than 5 [103], indicating no concern about multicollinearity issues.

Through PLS bootstrapping with 5000 iterations of resampling, the result of the structural model is presented in Figure 2.

As depicted in Figure 2, the effect of automation ($\beta = 0.339$, $p < 0.001$), ubiquitous connectivity ($\beta = 0.223$, $p < 0.05$), and environmental performance ($\beta = 0.260$, $p < 0.001$) on perceived usefulness are statistically significant. Thus, the hypothesis of automation (H1a), ubiquitous connectivity (H1b), and environmental performance (H1c) as a predictor of perceived usefulness are supported, and the model explains 58.0% of the variation in perceived usefulness, showing a moderate explanatory power. The effects of automation ($\beta = 0.531$, $p < 0.001$) and ubiquitous connectivity ($\beta = 0.225$, $p < 0.05$) as predictors of perceived ease of use are statistically significant. Thus, the hypothesis of automation (H2a) and ubiquitous connectivity (H2b) are confirmed, and the model explains 53.1% of the variation in perceived ease of use, suggesting a moderate explanatory power. The effect of technological uncertainty ($\beta = 0.614$, $p < 0.001$) as an antecedent of perceived risk is statistically significant. Thus, hypothesis H3a is confirmed, and the model explains 37.70% of the variation in perceived risk, indicating a weak explanatory power. The effects of structural assurance ($\beta = 0.637$, $p < 0.001$), and corporation reputation ($\beta = 0.300$, $p \leq 0.001$) on initial trust, are statistically significant. Hence, H4a and H4b are supported. The model explains 81.7% of the variation in initial trust, indicating a substantial explanatory power. For the direct effect of each variable on CAVs adoption, the

Table 4: Loadings and cross-loadings.

	PU	PEOU	PR	IT	SI	FC	AU	UB	EP	TU	SA	CR	AD
PU1	**0.92**	0.72	−0.26	0.68	0.73	0.67	0.67	0.63	0.64	−0.20	0.66	0.63	0.69
PU2	**0.85**	0.61	−0.28	0.59	0.70	0.59	0.60	0.55	0.53	−0.18	0.59	0.61	0.57
PU3	**0.92**	0.76	−0.28	0.72	0.78	0.70	0.69	0.66	0.67	−0.19	0.70	0.70	0.73
PU4	**0.86**	0.73	−0.32	0.64	0.67	0.64	0.62	0.61	0.59	−0.18	0.63	0.54	0.63
PEOU1	0.75	**0.96**	−0.31	0.62	0.70	0.72	0.66	0.61	0.62	−0.26	0.60	0.58	0.67
PEOU2	0.77	**0.97**	−0.29	0.65	0.73	0.77	0.69	0.64	0.64	−0.26	0.63	0.62	0.69
PEOU3	0.78	**0.96**	−0.33	0.69	0.77	0.79	0.72	0.66	0.67	−0.26	0.67	0.64	0.70
PR1	−0.27	−0.26	**0.88**	−0.21	−0.31	−0.23	−0.33	−0.42	−0.24	0.50	−0.26	−0.20	−0.25
PR2	−0.29	−0.31	**0.89**	−0.26	−0.33	−0.28	−0.36	−0.45	−0.29	0.50	−0.31	−0.22	−0.27
PR3	−0.27	−0.29	**0.89**	−0.24	−0.34	−0.23	−0.32	−0.42	−0.24	0.49	−0.29	−0.23	−0.26
PR4	−0.29	−0.28	**0.86**	−0.23	−0.30	−0.26	−0.36	−0.37	−0.32	0.53	−0.25	−0.21	−0.29
PR5	−0.25	−0.24	**0.75**	−0.27	−0.31	−0.26	−0.37	−0.34	−0.35	0.58	−0.28	−0.28	−0.28
IT1	0.73	0.66	−0.29	**0.97**	0.72	0.71	0.79	0.74	0.69	−0.16	0.84	0.78	0.75
IT2	0.70	0.62	−0.24	**0.96**	0.68	0.69	0.74	0.69	0.65	−0.13	0.83	0.80	0.71
IT3	0.73	0.68	−0.30	**0.95**	0.73	0.70	0.79	0.73	0.69	−0.16	0.89	0.82	0.77
SI1	0.79	0.75	−0.33	0.73	**0.94**	0.73	0.70	0.68	0.65	−0.23	0.72	0.70	0.68
SI2	0.75	0.72	−0.36	0.70	**0.95**	0.73	0.70	0.66	0.65	−0.26	0.71	0.71	0.68
SI3	0.75	0.69	−0.36	0.66	**0.93**	0.74	0.69	0.67	0.65	−0.25	0.68	0.70	0.68
FC1	0.58	0.61	−0.25	0.63	0.67	**0.87**	0.64	0.56	0.60	−0.21	0.58	0.60	0.57
FC2	0.68	0.74	−0.25	0.61	0.67	**0.89**	0.61	0.56	0.59	−0.23	0.57	0.56	0.62
FC4	0.67	0.72	−0.27	0.68	0.71	**0.86**	0.74	0.63	0.64	−0.26	0.63	0.65	0.65
AU1	0.58	0.57	−0.35	0.68	0.61	0.63	**0.86**	0.69	0.64	−0.22	0.65	0.71	0.65
AU2	0.67	0.65	−0.39	0.74	0.67	0.67	**0.93**	0.78	0.74	−0.22	0.73	0.75	0.74
AU3	0.70	0.71	−0.37	0.75	0.72	0.73	**0.90**	0.75	0.74	−0.26	0.74	0.69	0.80
UB1	0.65	0.63	−0.46	0.71	0.66	0.62	0.76	**0.94**	0.68	−0.29	0.70	0.70	0.70
UB2	0.61	0.59	−0.45	0.65	0.63	0.61	0.74	**0.94**	0.66	−0.26	0.66	0.65	0.63
UB3	0.69	0.66	−0.41	0.74	0.72	0.66	0.82	**0.93**	0.72	−0.23	0.75	0.75	0.76
EP1	0.63	0.62	−0.31	0.66	0.66	0.67	0.73	0.67	**0.94**	−0.34	0.65	0.67	0.75
EP2	0.63	0.61	−0.35	0.64	0.64	0.66	0.74	0.68	**0.94**	−0.29	0.65	0.67	0.74
EP3	0.69	0.66	−0.31	0.68	0.65	0.64	0.76	0.71	**0.95**	−0.26	0.70	0.67	0.81
TU1	−0.22	−0.26	0.54	−0.16	−0.27	−0.28	−0.26	−0.27	−0.30	**0.91**	−0.17	−0.19	−0.25
TU2	−0.15	−0.21	0.57	−0.13	−0.22	−0.22	−0.22	−0.22	−0.28	**0.93**	−0.13	−0.14	−0.17
TU3	−0.21	−0.28	0.58	−0.14	−0.24	−0.25	−0.24	−0.27	−0.30	**0.92**	−0.15	−0.13	−0.24
SA1	0.71	0.65	−0.31	0.84	0.73	0.64	0.74	0.74	0.69	−0.15	**0.95**	0.78	0.76
SA2	0.66	0.59	−0.28	0.82	0.70	0.64	0.73	0.69	0.63	−0.15	**0.93**	0.80	0.68
SA3	0.70	0.63	−0.33	0.85	0.69	0.65	0.76	0.71	0.69	−0.17	**0.95**	0.79	0.75
CR1	0.64	0.60	−0.28	0.81	0.68	0.64	0.74	0.70	0.64	−0.15	0.80	**0.93**	0.69
CR2	0.65	0.56	−0.26	0.77	0.68	0.62	0.72	0.70	0.65	−0.15	0.80	**0.94**	0.68
CR3	0.67	0.60	−0.22	0.75	0.71	0.66	0.73	0.68	0.67	−0.17	0.75	**0.94**	0.66
CR4	0.68	0.62	−0.25	0.80	0.73	0.67	0.79	0.72	0.70	−0.15	0.80	**0.94**	0.71
AD1	0.69	0.64	−0.30	0.74	0.69	0.66	0.79	0.71	0.79	−0.25	0.73	0.71	**0.93**
AD2	0.71	0.68	−0.30	0.72	0.67	0.66	0.75	0.69	0.75	−0.21	0.73	0.66	**0.95**
AD3	0.71	0.70	−0.30	0.74	0.69	0.67	0.78	0.71	0.78	−0.22	0.74	0.70	**0.96**

Bold values are the factor loadings of the constructs.

arrows in Figure 2 indicate that perceived usefulness ($\beta = 0.184$, $p < 0.05$), perceived ease of use ($\beta = 0.163$, $p < 0.05$), and initial trust ($\beta = 0.410$, $p < 0.001$) are statistically significant to explain CAVs' adoption. Therefore, H1, H2, and H4 are confirmed. In contrast, perceived risk ($\beta = -0.042$, $p > 0.05$), social influence ($\beta = 0.073$, $p > 0.05$), and facilitating conditions ($\beta = 0.070$, $p > 0.05$) have insignificant direct relationships with CAVs adoption. Accordingly, H3, H5, and H6 are not confirmed.

For the indirect effect of each factor on CAVs' adoption, the results suggest that the total indirect effects of automation ($\beta = 0.149$, $p < 0.001$), ubiquitous connectivity ($\beta = 0.078$, $p < 0.05$), structural assurance ($\beta = 0.261$, $p < 0.001$), and corporation reputation ($\beta = 0.123$, $p < 0.001$) are statistically significant on the intention to adopt CAVs. In contrast, environmental performance ($\beta = 0.048$, $p > 0.05$) and technological uncertainty ($\beta = -0.026$, $p > 0.05$) have no significant indirect influence on CAVs adoption. To sum up, the model explains the variance of 68.7% (R^2) in CAVs acceptance, indicating a moderate explanatory power [103]. Table 7 summarizes the results for the fourteen hypotheses.

4.2. Results of Qualitative Comparative Analysis. The key procedures of fsQCA empirical research include model development, sampling, data calibration, analysis of necessary conditions, analysis of sufficient conditions, and findings' interpretation.

Table 5: CA, ρ_A, CR, AVE, and VIF.

	CA	ρ_A	CR	AVE	VIF (AD)	VIF (PU)	VIF (PEOU)	VIF (PR)	VIF (IT)
PU	0.91	0.91	0.94	0.78	4.10				
PEOU	0.96	0.96	0.97	0.92	3.71				
PR	0.91	0.91	0.93	0.73	1.17				
IT	0.96	0.96	0.97	0.93	2.80				
SI	0.93	0.93	0.96	0.88	4.11				
FC	0.85	0.85	0.91	0.77	3.58				
AU	0.88	0.89	0.93	0.81		4.10	3.13		
UB	0.93	0.94	0.96	0.88		3.31	3.13		
EP	0.94	0.94	0.96	0.89		2.81			
TU	0.91	0.91	0.94	0.84				1.00	
SA	0.94	0.94	0.96	0.89					3.35
CR	0.95	0.96	0.97	0.88					3.35
AD	0.94	0.94	0.96	0.89					

Note. perceived usefulness, PU; perceived ease of use, PUOE; perceived risk, PR; initial trust, IT; social influence, SI; facilitating conditions, FC; automation, AU; ubiquitous connectivity, UC; environmental performance, EP; technological uncertainty, TU; structural assurance, SA; corporation reputation, CR; users' willingness to adopt CAVs, AD.

Table 6: Discriminant validity of the constructs.

	PU	PEOU	PR	IT	SI	FC	AU	UB	EP	TU	SA	CR	AD
PU	**0.89**												
PEOU	0.80	**0.96**											
PR	−0.32	−0.32	**0.86**										
IT	0.75	0.68	−0.29	**0.96**									
SI	0.81	0.77	−0.37	0.74	**0.94**								
FC	0.74	0.79	−0.29	0.73	0.78	**0.88**							
AU	0.73	0.72	−0.41	0.81	0.74	0.76	**0.90**						
UB	0.69	0.66	−0.47	0.75	0.72	0.67	0.83	**0.94**					
EP	0.69	0.67	−0.34	0.70	0.69	0.70	0.79	0.73	**0.94**				
TU	−0.21	−0.27	0.61	−0.16	−0.26	−0.27	−0.26	−0.28	−0.32	**0.92**			
SA	0.73	0.66	−0.33	0.89	0.75	0.68	0.79	0.75	0.71	−0.17	**0.94**		
CR	0.70	0.63	−0.27	0.83	0.75	0.69	0.80	0.75	0.71	−0.16	0.84	**0.94**	
AD	0.74	0.71	−0.32	0.78	0.72	0.70	0.82	0.74	0.82	−0.24	0.78	0.73	**0.95**

Note. The bold diagonal is AVE's square root.

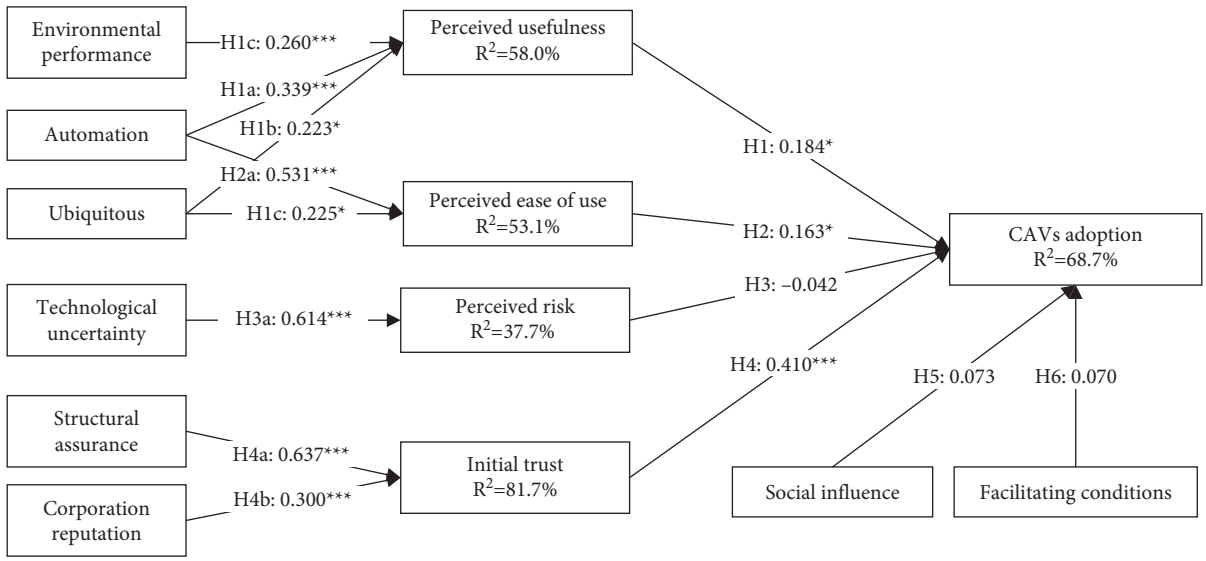


Figure 2: The results of the structural model of CAVs adoption. *Note.* $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 7: Summary of the hypothesis test.

Hypothesis	Structural path	β	SE	T value	p value	Supported
H1a	Automation ⟶Perceived usefulness	0.339***	0.084	4.052	0.001	Yes
H1b	Ubiquitous ⟶ Perceived usefulness	0.223*	0.091	2.436	0.015	Yes
H1c	Environmental performance ⟶ Perceived usefulness	0.260***	0.069	3.765	0.001	Yes
H1	Perceived usefulness⟶ Adoption	0.184*	0.075	2.464	0.014	Yes
H2a	Automation ⟶ Perceived ease of use	0.531***	0.092	5.775	0.001	Yes
H2b	Ubiquitous ⟶ Perceived ease of use	0.225*	0.107	2.109	0.035	Yes
H2	Perceived ease of use⟶ Adoption	0.163*	0.071	2.286	0.022	Yes
H3a	Technological uncertainty⟶ Perceived risk	0.614***	0.049	12.417	0.001	Yes
H3	Perceived risk ⟶ Adoption	−0.042	0.033	1.285	0.199	No
H4a	Structural assurance ⟶ Initial trust	0.637***	0.053	11.987	0.001	Yes
H4b	Corporation reputation ⟶ Initial trust	0.300***	0.055	5.470	0.001	Yes
H4	Initial trust ⟶ Adoption	0.410***	0.068	5.988	0.001	Yes
H5	Social influence ⟶ Adoption	0.073	0.080	0.917	0.359	No
H6	Facilitating conditions ⟶ Adoption	0.070	0.075	0.926	0.354	No

Note. $^{*}p \le 0.05$; $^{**}p \le 0.01$; $^{***}p \le 0.001$.

Table 8: Data calibration.

	Full nonmembership	Crossover point	Full membership
PU	1	3.9	5
PEOU	1	3.9	5
PR	1	2.3	5
IT	1	3.9	5
SI	1	3.9	5
FC	1	3.9	5
AD	1	4.1	5

4.2.1. Calibration. The data used in PLS-SEM should be calibrated into fuzzy sets for fsQCA analysis. The fuzzy set ranges from 0 to 1 on a continuous scale, where 0 represents the full nonset membership, and 1 shows the full set membership. As our actual data are not a normal distribution, consequently, the mean value of each condition is selected as a crossover point [37, 115]. By the procedure of fsQCA 3.0, data calibration is automatically calculated (see Table 8).

4.2.2. Analysis of Necessary Conditions. An analysis of necessary conditions tests whether there are any causal conditions that can be considered necessary to produce an outcome, which is CAVs' adoption in our study. According to the recommendations by previous studies [38], a condition is necessary when its consistency must be more than 0.9 [37]. As seen in Table 9, except for the perceived usefulness and ease of use, any condition alone is not necessary for CAVs adoption ("AD"). In addition, no condition alone is necessary for the negation of CAVs' adoption ("~AD"). The results suggest that no single condition on its own can result in output "~AD."

4.2.3. Analysis of Sufficient Conditions for CAVs' Acceptance. An analysis of sufficient conditions determines all conditions that are sufficient for an outcome. This study sets the frequency threshold (five) and the raw consistency threshold (0.8) to avoid distractions of less important configurations. According to the threshold setting, the remainder configurations capture 82% of the cases, meeting the recommendation for at least 75–80%.

Table 9: Analysis of necessary conditions.

Conditions	Outcome			
	AD (CAVs adoption)		~AD (negation of CAVs adoption)	
	Consistency	Coverage	Consistency	Coverage
PU	**0.932**	0.841	0.734	0.494
~PU	0.439	0.689	0.764	0.893
PEOU	**0.931**	0.827	0.745	0.493
~PEOU	0.430	0.693	0.739	0.888
PR	0.552	0.752	0.747	0.759
~PR	0.823	0.813	0.756	0.557
IT	0.893	0.893	0.693	0.516
~IT	0.516	0.693	0.856	0.857
SI	0.881	0.887	0.677	0.508
~SI	0.512	0.680	0.849	0.842
FC	0.854	0.891	0.670	0.521
~FC	0.541	0.688	0.860	0.814

Bold values show that the consistency value of each condition is greater than 0.9.

Based on the "Standard Analyses" procedure of fsQCA 3.0, the complex, intermediate, and parsimonious solutions are automatically given [38]. As the intermediate solution is most suitable for theoretical interpretation [115], it was used for our analyses. Table 10 reports the combinations of conditions that result in the high-level CAVs adoption. Any isolated condition is not sufficient for "AD," and five equifinal configurations (divided into four types by the core conditions) with a consistency of more than 0.8, indicating that all these configurations are sufficient. Moreover, every configuration's coverage (a similar approach of R^2 of regression model) is more than 0, suggesting they are empirically relevant [38]. Solution consistency (0.908) and solution coverage (0.869) should be more than 0.75 and 0.25, respectively [38]. In addition, solution coverage presents that total solutions account for 86.9% of samples related to high CAVs adoption.

As shown in Table 10, the presence of PU and PEOU and the absence of PR as core conditions are, respectively, listed in the different configurations. This result means that consumers' perception of technological characteristics

Table 10: Sufficient configurations for high CAVs adoption.

Causal condition	Configurations				
	1a	1b	2	3	4
PU	●	●		●	○
PEOU	●	●	●	●	○
PR		•		◯	◯
IT	•		•		
SI		○	○	•	○
FC		○	○		○
Raw coverage	0.840	0.377	0.423	0.734	0.381
Unique coverage	0.104	0.004	0.001	0.025	0.023
Consistency	0.923	0.953	0.946	0.941	0.861
Solution consistency			0.908		
Solution coverage			0.869		

Note. Black ("●") and hollow circles ("○") show the presence and absence of a condition, respectively. Moreover, large and small circles show core and peripheral conditions, respectively. Blank cells show a "do not care" situation.

Table 11: Sufficient configurations for the negation of CAVs' acceptance.

Causal condition	Configurations						
	1	2a	2b	2c	3	4	5
PU	○	○	●		●	●	●
PEOU	○		●	●	●	●	●
PR		○	●		○		
IT	◯			●	◯	●	●
SI	◯	◯	◯	◯	•	◯	
FC	◯	◯	◯	◯			◯
Raw coverage	0.641	0.555	0.490	0.545	0.527	0.548	0.559
Unique coverage	0.125	0.013	0.005	0.003	0.022	0.0130	0.013
Consistency	0.950	0.934	0.924	0.907	0.875	0.876	0.840
Solution consistency				0.822			
Solution coverage				0.844			

(usefulness, ease of use, and risk) of CAVs is the most important significant condition for realizing consumers' high willingness to adopt CAVs.

Solution 1a shows that the presence of PU and PEOU as core conditions combines with the presence of IT as peripheral conditions are sufficient for high CAVs adoption, which is in line with PLS-SEM results. Furthermore, solution 1a is the most important solution of the five causal paths, as it is the most empirically relevant (unique coverage = 0.104). Solution 1b demonstrates that the presence of PU and PEOU as core conditions, in case of the presence of PR and the absence of SI and FC, can fit together to produce the outcome. Solution 2 indicates that the configuration of the presence of PEOU as core conditions with the presence of IT as peripheral conditions, and the absence of SI and FC as peripheral conditions, is sufficient for high willingness to adopt CAVs.

Solution 3 demonstrates that the presence of PU and PEOU as core conditions and the absence of PR as core conditions, with the presence of SI as peripheral conditions, is sufficient for high intention to accept CAVs. Solution 4 indicates an important path to high CAVs adoption, combining the absence of PR as core conditions as well as PU, SI, and FC as peripheral conditions. Based on the two-factor theory [121], although PR, as the inhibitor of IT adoption, can hinder adoption in case of the absence of enablers (PU, SI, and FC), the absence of inhibitor could also enable the adoption. Besides, the absence of PR is a core condition in solution 3 and solution 4 in fsQCA, while PR is not statistically significant on CAVs acceptance in PLS-SEM, suggesting fsQCA complements the net effect perspective.

4.2.4. Analysis of Sufficient Conditions for the Negation of CAVs' Acceptance. Unlike conventional approaches, e.g., SEM and regression model, fsQCA is good at dealing with causal asymmetry [38]. Therefore, this study also explores which conditions work together on the negation of the outcome (~AD) by applying the same thresholds setting. The results of the negation of CAVs acceptance are shown in Table 11. Table 11 indicates that seven identified configurations are sufficient and empirically relevant, because each configuration's consistency and coverage are more than 0.8 and 0, respectively. The causal conditions account for 84.4% of samples (negation of CAVs adoption). These findings also indicate causal asymmetry.

As seen in Table 11, seven configurations (divided into five types by the core conditions) were found for the negation of CAVs' adoption.

5. Discussion and Conclusion

CAVs are emerging as a significant development in the automobile industry, potentially increasing safety, convenience, and efficiency in driving. Using a mixed-method within PLS-SEM and fsQCA, we sought to deeply understand CAVs' adoption by the net effect and combined effect of conditions.

It is found from PLS-SEM results that initial trust ($\beta = 0.41$) is one of the most significant factors that influence users' willingness to accept CAVs. The significantly net effect of initial trust in shaping adoption intention corroborates previous studies, e.g., mobile banking [108], autonomous vehicles [5, 24, 29, 30], and autonomous shuttle/bus [8, 17, 54]. This finding indicates that trust, as "a tool for the reduction of cognitive complexity," may help reduce cognitive complexity and simplify decision-making processes, particularly in situations involving risks or uncertainty [25]. As a result, initial trust is of particular importance for CAVs, as it indicates that the potential consumer must have confidence in CAVs in order to accept it. In addition, the results of fsQCA shows that "the presence of initial trust" is just a peripheral condition of configuration (solution 1a and 2) for CAVs adoption, and more importantly, "the absence of initial trust" is indeed a core condition in two of the seven configurations (solution 1 and 3) for the negation of CAVs adoption (i.e., the presence of AD), which indicates that the lack of initial trust in CAVs is likely to reduce the willingness to adopt CAVs. This finding of fsQCA is in accordance with previous studies that a major psychological barrier to the widespread adoption of CAVs is an inadequate sense of trust [24, 30]. Structural assurance ($\beta = 0.637$) and corporation reputation ($\beta = 0.300$) directly influence initial trust and indirectly influence CAVs adoption through initial trust, which concur with previous studies suggesting they are the important antecedents of technology adoption [99, 108]. In order to gain users' trust and promote the adoption of CAVs, providing structural assurance and building corporation reputation are necessary.

The results of PLS-SEM indicate that either perceived usefulness (PU) ($\beta = 0.184$), or perceived ease of use (PEOU) ($\beta = 0.163$) significantly affects users' willingness to adopt CAVs. The significantly net effect of PU is consistent with previous studies on various technologies, e.g., mobile payment [122], social commerce [123], IoT [124], autonomous vehicles [2, 6, 21, 24, 25, 53], and autonomous shuttle [8], which highlight the importance of the users' perception of usefulness to achieve the behavioral intention to use innovation. The relevance of PEOU is in line with similar studies of IoT [124], mobile payment [122], social commerce [123], and automated vehicles [2, 21, 29]. The findings from fsQCA reconfirm that PU (Consistency = 0.932) and PEOU (Consistency = 0.931) are necessary for CAVs' adoption. Moreover, "the presence of PU" and "the presence of PEOU" are core conditions in three (solution 1a, 1b, and 3) and four (solution 1a, 1b, 2, and 3) of the five configurations leading to the presence of CAVs adoption, respectively, providing additional support for the related hypotheses of PLS-SEM. Findings from these two distinct methods demonstrate that higher PU and PEOU could increase the intention to adopt CAVs [24, 117]. In addition, it is found that automation and ubiquitous connectivity directly impact PU and PEOU, and indirectly impact CAVs' adoption. Environmental performance has a direct impact on PU, while no indirect impact on CAVs' adoption.

Despite perceived risk (PR) has been frequently cited as one major concern in adopting CAVs in surveys [125], PLS-SEM failed to identify its significant net effect on behavioral intention to accept CAVs. This result is aligned with the previous studies that find a nonsignificant direct effect of PR on the behavioral intention to use social commerce [123], autonomous vehicles [29, 30], and connected vehicles [31]. Interestingly, fsQCA results show that the absence of PR is indeed a core condition in two of the five configurations (solution 3 and 4) for user's willingness to adopt CAVs, which corroborates previous studies that observe a negative direct effect of PR on the behavioral intention of innovative products or services such as mobile payment [122], social media purchase [126], and autonomous vehicle [5, 6, 52, 53]. The empirical findings of previous studies are rather mixed, resulting in the argument that PR cannot be seen as a steady predictor of CAVs' adoption [5, 52], indicating the lower the risk perception by potential users, the higher the acceptance of CAVs. In addition, technological uncertainty ($\beta = 0.614$) as an antecedent of perceived risk is statistically significant, while it has no significant indirect influence on CAVs' adoption, which is consistent with previous studies [97]. As the CAV is a technology-intensive product, the more consumers feel uncertain about their technology, the more they will associate risk with CAVs.

Different from the significant net effect of social influence (SI) and facilitating conditions (FC) in previous studies [14, 52], PLS-SEM results show that SI and FC have no significant net effect on user's willingness to adopt CAVs, which is aligned with previous studies that find a nonsignificant direct effect of SI on user's intention [127] and nonsignificant direct effect of FC on user's intention [128]. However, SI, as peripheral conditions, is present in one configuration (solution 3) for AD in the fsQCA result, suggesting that the relevance of SI on AD is consistent with numerous studies in IoT [124] and autonomous vehicles [3, 4, 21]. In addition, fsQCA shows that the absence of SI (solution 1, 2a, 2b, 2c, and 4) and the absence of FC (solution 1, 2a, 2b, 2c, and 5) are core conditions in five of the seven configurations for the negation of behavioral intention to accept CAVs, showing the existence of causal asymmetry in a complex context [35, 41].

In short, results from PLS-SEM not only indicate that perceived usefulness, perceived ease of use, and initial trust increase user's willingness but also identify the predictors of perceived usefulness (automation, ubiquitous connectivity, and environmental performance), perceived risk (technological uncertainty), and initial trust (structural assurance and corporation reputation). Moreover, results from fsQCA which are used as supplementary analysis technique [129] reinforce the symmetric findings of PLS-SEM and offer additional novel, interesting, and more nuanced insights that indicate a combination of the conditions needs to be

taken into account to explain the outcome. Five configurations for "AD" and seven configurations for "~AD" are identified. This synergetic effect could not be captured by PLS-SEM since it examines the condition in isolation from the other conditions [130, 131].

5.1. Theoretical Implications. Firstly, our study uses a holistic and integrative approach to exploring CAVs' acceptance, making important contributions to the literature on CAVs' acceptance, and responding to the call for comprehensive approaches that integrate multiple theories to understand the adoption of advanced technologies [8, 17–20, 37, 101], due to single perspectives not being enough for comprehensive investigations [17]. Although previous literature has addressed CAVs' adoption, the advantage of our study is that it combines the perceived, technological, and environmental factors from several technology adoption theories, e.g., TAM, perceived risk theory, initial trust model, and UTAUT. Empirical studies show that our integrated model overcomes the shortcoming of a single model, and explains 68.7% variance (R^2) of CAVs' adoption, indicating a high explanatory power.

Secondly, from a methodological perspective, our study contributes to illustrating the complementarity of PLS-SEM and fsQCA in the context of CAVs' acceptance [44]. Numerous studies on CAVs' acceptance generally use the multiple regression model, SEM, and PLS-SEM techniques, and test the net effect of each isolated antecedent on the consumer's willingness to accept CAVs. Hence, we responded to the call for applying the QCA methodology to complex behavioral studies [117]. Specifically, the PLS-SEM methodology is appropriate for identifying key "driver" constructs for CAVs' acceptance. In contrast, fsQCA provides a deeper understanding of a configuration of the conditions that need to be considered to explain the complex, nonlinear, and asymmetric influences of causal conditions for the outcome (CAVs' adoption).

5.2. Managerial Implications. Our study provides automotive manufacturers with several managerial implications for pushing CAVs into the wider mass market.

Firstly, the perceived usefulness and perceived ease of use of CAVs are very crucial to potential users. Hence, automotive manufacturers should focus on the core technology development of CAVs (e.g., higher automation, higher ubiquitous connectivity, and higher environmental performance) for improving the usefulness and ease of use of CAVs.

Second, lower perceived risk is also found to be crucial for CAVs' acceptance. Therefore, automotive manufacturers should pay attention to reducing technological uncertainty and market uncertainty, by R&D of the hardware and software of autonomous driving technology (e.g., sensing and data input, computation, and decision-making).

Third, initial trust is of great significance in predicting CAVs' acceptance, and the lack of initial trust is likely to result in lower willingness. Hence, consumers' trust-building has become a top priority for automotive manufacturers [5]. Based on our research on the antecedents of consumers' initial trust, structural assurance and motor corporation should be strengthened for building initial trust in CAVs.

Finally, a lack of social influence and facilitating conditions would result in a lower willingness to buy CAVs. Hence, marketers should display CAVs' features and advantages through auto shows and advertising to form positive attitudes and a good reputation.

5.3. Limitations and Future Directions. First, our study only identifies and evaluates several key antecedents of CAVs' acceptance, based on IT adoption theories and previous literature on CAVs' acceptance. However, there may be several other factors influencing consumers' willingness to buy CAVs. Therefore, future research should consider more comprehensive factors for CAVs' acceptance. Second, this study relies only on a group of respondents in a single country, China. However, the consumer in other regions or countries may have different attitudes toward CAVs. Hence, future studies should conduct a comparative analysis across multiple regions, to enhance its generalizability.

Data Availability

The sample data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Shandong Provincial Natural Science Foundation of China (ZR2020MG075).

References

[1] M. Jaller, C. Otero-Palencia, and A. Pahwa, "Automation, electrification, and shared mobility in urban freight: opportunities and challenges," *Transportation Research Procedia*, vol. 46, pp. 13–20, 2020.

[2] J. Wu, H. Liao, J.-W. Wang, and T. Chen, "The role of environmental concern in the public acceptance of autonomous electric vehicles: a survey from China," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 60, pp. 37–46, 2019.

[3] N. Adnan, S. Md Nordin, M. A. bin Bahruddin, and M. Ali, "How trust can drive forward the user acceptance to the technology? In-vehicle technology for autonomous vehicle," *Transportation Research Part A: Policy and Practice*, vol. 118, pp. 819–836, 2018.

[4] T. Leicht, A. Chtourou, and K. Ben Youssef, "Consumer innovativeness and intentioned autonomous car adoption," *The Journal of High Technology Management Research*, vol. 29, no. 1, pp. 1–11, 2018.

[5] H. Liu, R. Yang, L. Wang, and P. Liu, "Evaluating initial public acceptance of highly and fully autonomous vehicles," *International Journal of Human-Computer Interaction*, vol. 35, no. 11, pp. 919–931, 2019.

[6] G. Zhu, Y. Chen, and J. Zheng, "Modelling the acceptance of fully autonomous vehicles: a media-based perception and adoption model," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 73, pp. 80–91, 2020.

[7] A. Rahimi, G. Azimi, H. Asgari, and X. Jin, "Adoption and willingness to pay for autonomous vehicles: attitudes and latent classes," *Transportation Research Part D: Transport and Environment*, vol. 89, Article ID 102611, 2020.

[8] C.-F. Chen, "Factors affecting the decision to use autonomous shuttle services: evidence from a scooter-dominant urban context," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 67, pp. 195–204, 2019.

[9] A. Talebian and S. Mishra, "Predicting the adoption of connected autonomous vehicles: a new approach based on the theory of diffusion of innovations," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 363–380, 2018.

[10] P. Bansal and K. M. Kockelman, "Are we ready to embrace connected and self-driving vehicles? A case study of Texans," *Transportation*, vol. 45, no. 2, pp. 641–675, 2018.

[11] K. F. Yuen, G. Chua, X. Wang, F. Ma, and K. X. Li, "Understanding public acceptance of autonomous vehicles using the theory of planned behaviour," *International Journal of Environmental Research and Public Health*, vol. 17, no. 12, p. 4419, 2020.

[12] S.-A. Kaye, I. Lewis, L. Buckley, and A. Rakotonirainy, "Assessing the feasibility of the theory of planned behaviour in predicting drivers' intentions to operate conditional and full automated vehicles," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 74, pp. 173–183, 2020.

[13] Y. Huang and L. Qian, "Understanding the potential adoption of autonomous vehicles in China: the perspective of behavioral reasoning theory," *Psychology and Marketing*, vol. 38, no. 4, pp. 669–690, 2021.

[14] R. Madigan, T. Louw, M. Wilbrink, A. Schieben, and N. Merat, "What influences the decision to use automated public transport? Using UTAUT to understand public acceptance of automated road transport systems," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 50, pp. 55–64, 2017.

[15] M. A. Ribeiro, D. Gursoy, and O. H. Chi, "Customer acceptance of autonomous vehicles in travel and tourism," *Journal of Travel Research*, vol. 61, no. 3, pp. 620–636, 2022.

[16] J. Kim, S. Kim, and C. Nam, "User resistance to acceptance of In-Vehicle Infotainment (IVI) systems," *Telecommunications Policy*, vol. 40, no. 9, pp. 919–930, 2016.

[17] B. Herrenkind, I. Nastjuk, A. B. Brendel, S. Trang, and L. M. Kolbe, "Young people's travel behavior - using the life-oriented approach to understand the acceptance of autonomous driving," *Transportation Research Part D: Transport and Environment*, vol. 74, pp. 214–233, 2019.

[18] P. R. Palos-Sanchez, F. J. Arenas-Marquez, and M. Aguayo-Camacho, "Cloud computing (SaaS) adoption as a strategic technology: results of an empirical study," *Mobile Information Systems*, vol. 2017, Article ID 2536040, 2017.

[19] M. J. Khan and S. Mahmood, "Assessing the determinants of adopting component-based development in a global context: a client-vendor analysis," *IEEE Access*, vol. 6, pp. 79060–79073, 2018.

[20] O. Ali, A. Shrestha, V. Osmanaj, and S. Muhammed, "Cloud Computing Technology Adoption: An Evaluation of Key Factors in Local Governments," *Information Technology & People*, vol. 34, no. 2, 2020.

[21] I. Panagiotopoulos and G. Dimitrakopoulos, "An empirical investigation on consumers' intentions towards autonomous driving," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 773–784, 2018.

[22] J. Park, E. Hong, and H. T. Le, "Adopting autonomous vehicles: the moderating effects of demographic variables," *Journal of Retailing and Consumer Services*, vol. 63, Article ID 102687, 2021.

[23] M. Liu, J. Wu, C. Zhu, and K. Hu, "A study on public adoption of robo-taxis in China," *Journal of Advanced Transportation*, vol. 2020, Article ID 8877499, 2020.

[24] S. M. Hegner, A. D. Beldad, and G. J. Brunswick, "In automatic we trust: investigating the impact of trust, control, personality characteristics, and extrinsic and intrinsic motivations on the acceptance of autonomous vehicles," *International Journal of Human-Computer Interaction*, vol. 35, no. 19, pp. 1769–1780, 2019.

[25] T. Zhang, D. Tao, X. Qu, X. Zhang, R. Lin, and W. Zhang, "The roles of initial trust and perceived risk in public's acceptance of automated vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 98, pp. 207–220, 2019.

[26] S. Nordhoff, V. Malmsten, B. van Arem, P. Liu, and R. Happee, "A structural equation modeling approach for the acceptance of driverless automated shuttles based on constructs from the Unified Theory of Acceptance and Use of Technology and the Diffusion of Innovation Theory," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 78, pp. 58–73, 2021.

[27] P. Jing, L. Du, Y. Chen, Y. Shi, F. Zhan, and J. Xie, "Factors that influence parents' intentions of using autonomous vehicles to transport children to and from school," *Accident Analysis & Prevention*, vol. 152, Article ID 105991, 2021.

[28] T. Keszey, "Behavioural intention to use autonomous vehicles: systematic review and empirical extension," *Transportation Research Part C: Emerging Technologies*, vol. 119, Article ID 102732, 2020.

[29] J. K. Choi and Y. G. Ji, "Investigating the importance of trust on adopting an autonomous vehicle," *International Journal of Human-Computer Interaction*, vol. 31, no. 10, pp. 692–702, 2015.

[30] P. Liu, Q. Guo, F. Ren, L. Wang, and Z. Xu, "Willingness to pay for self-driving vehicles: influences of demographic and psychological factors," *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 306–317, 2019.

[31] J. Walter and B. Abendroth, "On the role of informational privacy in connected vehicles: a privacy-aware acceptance modelling approach for connected vehicular services," *Telematics and Informatics*, vol. 49, Article ID 101361, 2020.

[32] T. Zhang, D. Tao, X. Qu et al., "Automated vehicle acceptance in China: social influence and initial trust are key determinants," *Transportation Research Part C: Emerging Technologies*, vol. 112, pp. 220–233, 2020.

[33] D. Gligor and S. Bozkurt, "FsQCA versus regression: the context of customer engagement," *Journal of Retailing and Consumer Services*, vol. 52, Article ID 101929, 2020.

[34] T. Greckhamer, S. Furnari, P. C. Fiss, and R. V. Aguilera, "Studying configurations with qualitative comparative analysis: best practices in strategy and organization research," *Strategic Organization*, vol. 16, no. 4, pp. 482–495, 2018.

[35] P. C. Fiss, "Building better causal theories: a fuzzy set approach to typologies in organization research," *Academy of Management Journal*, vol. 54, no. 2, pp. 393–420, 2011.

[36] P. E. Kourouthanassis, P. Mikalef, I. O. Pappas, and P. Kostagiolas, "Explaining travellers online information satisfaction: a complexity theory approach on information needs, barriers, sources and personal characteristics," *Information & Management*, vol. 54, no. 6, pp. 814–824, 2017.

[37] L. Wang, Z. Wang, X. Wang, and Y. Zhao, "Explaining consumer implementation intentions in mobile shopping with SEM and fsQCA: roles of visual and technical perceptions," *Electronic Commerce Research and Applications*, vol. 49, p. 101080, 2021.

[38] C. C. Ragin, *Redesigning Social Inquiry: Fuzzy Sets and beyond*, University of Chicago Press, Oxford, 2008.

[39] A. G. Woodside, "Moving beyond multiple regression analysis to algorithms: calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory," *Journal of Business Research*, vol. 66, no. 4, pp. 463–472, 2013.

[40] M. Ali, K. A. Seny Kan, and M. Sarstedt, "Direct and configurational paths of absorptive capacity and organizational innovation to successful organizational performance," *Journal of Business Research*, vol. 69, no. 11, pp. 5317–5323, 2016.

[41] Y. Liu, J. Mezei, V. Kostakos, and H. Li, "Applying configurational analysis to IS behavioural research: a methodological alternative for modelling combinatorial complexities," *Information Systems Journal*, vol. 27, no. 1, pp. 59–89, 2017.

[42] I. O. Pappas, P. E. Kourouthanassis, P. Mikalef, and M. N. Giannakos, "Combining system success factors with trust to explain e-government adoption using fsQCA," in *Proceedings of the Americas Conference on Information Systems*, New Orleans LA.USA, August 2018.

[43] D. J. Fagnant and K. Kockelman, "Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations," *Transportation Research Part A: Policy and Practice*, vol. 77, pp. 167–181, 2015.

[44] F. McLeay, H. Olya, H. Liu, C. Jayawardhena, and C. Dennis, "A multi-analytical approach to studying customers motivations to use innovative totally autonomous vehicles," *Technological Forecasting and Social Change*, vol. 174, Article ID 121252, 2022.

[45] K. F. Yuen, L. Cai, G. Qi, and X. Wang, "Factors influencing autonomous vehicle adoption: an application of the technology acceptance model and innovation diffusion theory," *Technology Analysis & Strategic Management*, vol. 33, no. 5, pp. 505–519, 2021.

[46] S. Kapser, M. Abdelrahman, and T. Bernecker, "Autonomous delivery vehicles to fight the spread of Covid-19 - how do men and women differ in their acceptance?" *Transportation Research Part A: Policy and Practice*, vol. 148, pp. 183–198, 2021.

[47] X. Wang, Y. D. Wong, K. X. Li, and K. F. Yuen, "This is not me! Technology-identity concerns in consumers' acceptance of autonomous vehicle technology," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 74, pp. 345–360, 2020.

[48] K. F. Yuen, D. T. K. Huyen, X. Wang, and G. Qi, "Factors influencing the adoption of shared autonomous vehicles," *International Journal of Environmental Research and Public Health*, vol. 17, no. 13, p. 4868, 2020.

[49] I. Nastjuk, B. Herrenkind, M. Marrone, A. B. Brendel, and L. M. Kolbe, "What drives the acceptance of autonomous driving? An investigation of acceptance factors from an end-user's perspective," *Technological Forecasting and Social Change*, vol. 161, Article ID 120319, 2020.

[50] T. Dirsehan and C. Can, "Examination of trust and sustainability concerns in autonomous vehicle adoption," *Technology in Society*, vol. 63, p. 101361, 2020.

[51] K. F. Yuen, Y. D. Wong, F. Ma, and X. Wang, "The determinants of public acceptance of autonomous vehicles: an innovation diffusion perspective," *Journal of Cleaner Production*, vol. 270, Article ID 121904, 2020.

[52] S. Kapser and M. Abdelrahman, "Acceptance of autonomous delivery vehicles for last-mile delivery in Germany - e," *Transportation Research Part C: Emerging Technologies*, vol. 111, pp. 210–225, 2020.

[53] J. Lee, D. Lee, Y. Park, S. Lee, and T. Ha, "Autonomous vehicles can be shared, but a feeling of ownership is important: examination of the influential factors for intention to use autonomous vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 107, pp. 411–422, 2019.

[54] B. Herrenkind, A. B. Brendel, I. Nastjuk, M. Greve, and L. M. Kolbe, "Investigating end-user acceptance of autonomous electric buses to accelerate diffusion," *Transportation Research Part D: Transport and Environment*, vol. 74, pp. 255–276, 2019.

[55] J. M. Müller, "Comparing technology acceptance for autonomous vehicles, battery electric vehicles, and car sharing-A study across Europe, China, and North America," *Sustainability*, vol. 11, no. 16, p. 4333, 2019.

[56] S. Deb, L. Strawderman, D. W. Carruth, J. DuBien, B. Smith, and T. M. Garrison, "Development and validation of a questionnaire to assess pedestrian receptivity toward fully autonomous vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 84, pp. 178–195, 2017.

[57] R. Madigan, T. Louw, M. Dziennus et al., "Acceptance of automated road transport systems (ARTS): an adaptation of the UTAUT model," *Transportation Research Procedia*, vol. 14, pp. 2217–2226, 2016.

[58] J. Zmud, I. N. Sener, and J. Wagner, "Self-driving vehicles: determinants of adoption and conditions of usage," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2565, no. 1, pp. 57–64, 2016.

[59] M. Ghasri, A. Ardeshiri, and T. Rashidi, "Perception towards electric vehicles and the impact on consumers' preference," *Transportation Research Part D: Transport and Environment*, vol. 77, pp. 271–291, 2019.

[60] J. R. Simpson and S. Mishra, "Developing a methodology to predict the adoption rate of Connected Autonomous Trucks in transportation organizations using peer effects," *Research in Transportation Economics*, vol. 90, Article ID 100866, 2020.

[61] S. Wang and J. Zhao, "Risk preference and adoption of autonomous vehicles," *Transportation Research Part A: Policy and Practice*, vol. 126, pp. 215–229, 2019.

[62] R. Krueger, T. H. Rashidi, and J. M. Rose, "Preferences for shared autonomous vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 69, pp. 343–355, 2016.

[63] T. U. Saeed, M. W. Burris, S. Labi, and K. C. Sinha, "An empirical discourse on forecasting the use of autonomous vehicles using consumers' preferences," *Technological Forecasting and Social Change*, vol. 158, Article ID 120130, 2020.

[64] K. Maeng, W. Kim, and Y. Cho, "Consumers' attitudes toward information security threats against connected and autonomous vehicles," *Telematics and Informatics*, vol. 63, Article ID 101646, 2021.

[65] C. J. Haboucha, R. Ishaq, and Y. Shiftan, "User preferences regarding autonomous vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 78, pp. 37–49, 2017.

[66] P. Bansal and K. M. Kockelman, "Forecasting Americans' long-term adoption of connected and autonomous vehicle technologies," *Transportation Research Part A: Policy and Practice*, vol. 95, pp. 49–63, 2017.

[67] M. Lavasani, X. Jin, and Y. Du, "Market penetration model for autonomous vehicles on the basis of earlier technology adoption experience," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2597, no. 1, pp. 67–74, 2016.

[68] K. Maeng and Y. Cho, "Who will want to use shared autonomous vehicle service and how much? A consumer experiment in South Korea," *Travel Behaviour and Society*, vol. 26, pp. 9–17, 2022.

[69] M. S. Hossain and M. R. Fatmi, "Modelling the adoption of autonomous vehicle: how historical experience inform the future preference," *Travel Behaviour and Society*, vol. 26, pp. 57–66, 2022.

[70] S. H.-W. Chuah, M.-L. Tseng, K.-J. Wu, and C.-F. Cheng, "Factors influencing the adoption of sharing economy in B2B context in China: findings from PLS-SEM and fsQCA," *Resources, Conservation and Recycling*, vol. 175, Article ID 105892, 2021.

[71] H. Gangwar, H. Date, and A. D. Raoot, "Review on IT adoption: insights from recent technologies," *Journal of Enterprise Information Management*, vol. 27, no. 4, pp. 488–502, 2014.

[72] M. Fishbein and I. Ajzen, *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*, Addison-Wesley Publishing Company, California, Riverside, 1975.

[73] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, 1989.

[74] I. Ajzen, "The theory of planned behavior," *Organizational Behavior and Human Decision Processes*, vol. 50, no. 2, pp. 179–211, 1991.

[75] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: toward a unified view," *MIS Quarterly*, vol. 27, no. 3, pp. 425–478, 2003.

[76] D. H. McKnight and N. L. Chervany, "What trust means in E-commerce customer relationships: an interdisciplinary conceptual typology," *International Journal of Electronic Commerce*, vol. 6, no. 2, pp. 35–59, 2001.

[77] E. M. Rogers, *Diffusion of Innovations*, Simon & Schuster, Manhattan, New York City, 2020.

[78] Y. Liang, G. Qi, K. Wei, and J. Chen, "Exploring the determinant and influence mechanism of e-Government cloud adoption in government agencies in China," *Government Information Quarterly*, vol. 34, no. 3, pp. 481–495, 2017.

[79] C. Yang, X. Ye, J. Xie et al., "Analyzing drivers' intention to accept parking App by structural equation model," *Journal of Advanced Transportation*, vol. 2020, Article ID 3051283, 2020.

[80] S. A. Kamal, M. Shafiq, and P. Kakria, "Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM)," *Technology in Society*, vol. 60, Article ID 101212, 2020.

[81] H. Rafique, A. O. Almagrabi, A. Shamim, F. Anwar, and A. K. Bashir, "Investigating the acceptance of mobile library applications with an extended technology acceptance model (TAM)," *Computers & Education*, vol. 145, Article ID 103732, 2020.

[82] K. T. Manis and D. Choi, "The virtual reality hardware acceptance model (VR-HAM): extending and individuating the technology acceptance model (TAM) for virtual reality hardware," *Journal of Business Research*, vol. 100, pp. 503–513, 2019.

[83] C. Sagnier, E. Loup-Escande, D. Lourdeaux, I. Thouvenin, and G. Valléry, "User acceptance of virtual reality: an extended technology acceptance model," *International Journal of Human-Computer Interaction*, vol. 36, no. 11, pp. 993–1007, 2020.

[84] A. Shuhaiber and I. Mashal, "Understanding users' acceptance of smart homes," *Technology in Society*, vol. 58, Article ID 101110, 2019.

[85] V. Venkatesh, J. Y. L. Thong, and X. Xu, "Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology," *MIS Quarterly*, vol. 36, no. 1, pp. 157–178, 2012.

[86] S. Nordhoff, J. de Winter, M. Kyriakidis, B. van Arem, and R. Happee, "Acceptance of driverless vehicles: results from a large cross-national questionnaire study," *Journal of Advanced Transportation*, vol. 2018, Article ID 5382192, 2018.

[87] H. Yang, H. Lee, and H. Zo, "User acceptance of smart home services: an extension of the theory of planned behavior," *Industrial Management & Data Systems*, vol. 117, no. 1, pp. 68–89, 2017.

[88] J. C. F. de Winter, R. Happee, M. H. Martens, and N. A. Stanton, "Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 27, pp. 196–217, 2014.

[89] F. Hartwich, M. Beggiato, and J. F. Krems, "Driving comfort, enjoyment and acceptance of automated driving - effects of drivers' age and driving style familiarity," *Ergonomics*, vol. 61, no. 8, pp. 1017–1032, 2018.

[90] W. Gao, Z. Liu, Q. Guo, and X. Li, "The dark side of ubiquitous connectivity in smartphone-based SNS: an integrated model from information perspective," *Computers in Human Behavior*, vol. 84, pp. 185–193, 2018.

[91] K. Degirmenci and M. H. Breitner, "Consumer purchase intentions for electric vehicles: is green more important than price and range?" *Transportation Research Part D: Transport and Environment*, vol. 51, pp. 250–260, 2017.

[92] S. Wang, J. Li, and D. Zhao, "The impact of policy measures on consumer intention to adopt electric vehicles: evidence from China," *Transportation Research Part A: Policy and Practice*, vol. 105, pp. 14–26, 2017.

[93] M. Mohamed, C. D. Higgins, M. Ferguson, and W. J. Réquia, "The influence of vehicle body type in shaping behavioural intention to acquire electric vehicles: a multi-group structural equation approach," *Transportation Research Part A: Policy and Practice*, vol. 116, pp. 54–72, 2018.

[94] T. Okada, T. Tamaki, and S. Managi, "Effect of environmental awareness on purchase intention and satisfaction pertaining to electric vehicles in Japan," *Transportation Research Part D: Transport and Environment*, vol. 67, pp. 503–513, 2019.

[95] V. W. Mitchell, "Understanding consumers' behaviour: can perceived risk theory help?" *Management Decision*, vol. 30, no. 3, 1992.

[96] W. M. Al-Rahmi, N. Yahaya, A. A. Aldraiweesh et al., "Big data adoption and knowledge management sharing: an

empirical investigation on their adoption and sustainability as a purpose of education," *IEEE Access*, vol. 7, pp. 47245–47258, 2019.

[97] A. Hong, C. Nam, and S. Kim, "What will be the possible barriers to consumers' adoption of smart home services?" *Telecommunications Policy*, vol. 44, no. 2, Article ID 101867, 2020.

[98] D. H. McKnight, L. L. Cummings, and N. L. Chervany, "Initial trust formation in new organizational relationships," *Academy of Management Review*, vol. 23, no. 3, pp. 473–490, 1998.

[99] G. Kim, B. Shin, and H. G. Lee, "Understanding dynamics between initial trust and usage intentions of mobile banking," *Information Systems Journal*, vol. 19, no. 3, pp. 283–311, 2009.

[100] S.-C. Ma, Y. Fan, J.-F. Guo, J.-H. Xu, and J. Zhu, "Analysing online behaviour to determine Chinese consumers' preferences for electric vehicles," *Journal of Cleaner Production*, vol. 229, pp. 244–255, 2019.

[101] G. Cao, Y. Duan, J. S. Edwards, and Y. K. Dwivedi, "Understanding managers' attitudes and behavioral intentions towards using artificial intelligence for organizational decision-making," *Technovation*, vol. 106, Article ID 102312, 2021.

[102] N. Valaei, S. Rezaei, and W. K. W. Ismail, "Examining learning strategies, creativity, and innovation at SMEs using fuzzy set Qualitative Comparative Analysis and PLS path modeling," *Journal of Business Research*, vol. 70, pp. 224–233, 2017.

[103] J. F. Hair, J. J. Risher, M. Sarstedt, and C. M. Ringle, "When to use and how to report the results of PLS-SEM," *European Business Review*, vol. 31, no. 1, pp. 2–24, 2019.

[104] T. Luor, H.-P. Lu, H. Yu, and Y. Lu, "Exploring the critical quality attributes and models of smart homes," *Maturitas*, vol. 82, no. 4, pp. 377–386, 2015.

[105] R. P. Q. Falcao, J. B. Ferreira, and M. Carrazedo Marques da Costa Filho, "The influence of ubiquitous connectivity, trust, personality and generational effects on mobile tourism purchases," *Information Technology & Tourism*, vol. 21, no. 4, pp. 483–514, 2019.

[106] H. Kim, I. Lee, and J. Kim, "Maintaining continuers vs. converting discontinuers: relative importance of post-adoption factors for mobile data services," *International Journal of Mobile Communications*, vol. 6, no. 1, pp. 108–132, 2008.

[107] Y. Yang, Y. Liu, H. Li, and B. Yu, "Understanding perceived risks in mobile payment acceptance," *Industrial Management & Data Systems*, vol. 115, no. 2, pp. 253–269, 2015.

[108] T. Oliveira, M. Faria, M. A. Thomas, and A. Popovič, "Extending the understanding of mobile banking adoption: when UTAUT meets TTF and ITM," *International Journal of Information Management*, vol. 34, no. 5, pp. 689–703, 2014.

[109] S. Nordhoff, T. Louw, S. Innamaa et al., "Using the UTAUT2 model to explain public acceptance of conditionally automated (L3) cars: a questionnaire study among 9,118 car drivers from eight European countries," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 74, pp. 280–297, 2020.

[110] N. Kock, "Common method bias in PLS-SEM," *International Journal of E-Collaboration*, vol. 11, no. 4, pp. 1–10, 2015.

[111] Y. Liang, W. Wang, K. Dong, G. Zhang, and G. Qi, "Adoption of mobile government cloud from the perspective of public sector," *Mobile Information Systems*, vol. 2021, Article ID 8884594, 2021.

[112] C. C. Ragin, *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*, Univ of California Press, Oakland, California, 2014.

[113] C. C. Ragin, "Set relations in social research: evaluating their consistency and coverage," *Political Analysis*, vol. 14, no. 3, pp. 291–310, 2006.

[114] A. Ortiz de Guinea and L. Raymond, "Enabling innovation in the face of uncertainty through IT ambidexterity: a fuzzy set qualitative comparative analysis of industrial service SMEs," *International Journal of Information Management*, vol. 50, pp. 244–260, 2020.

[115] P. C. Fiss, "A set-theoretic approach to organizational configurations," *Academy of Management Review*, vol. 32, no. 4, pp. 1180–1198, 2007.

[116] C. C. Ragin, "Using qualitative comparative analysis to study causal complexity," *Health Services Research*, vol. 34, no. 5, pp. 1225–1239, 1999.

[117] I. O. Pappas and A. G. Woodside, "Fuzzy-set qualitative comparative analysis (fsQCA): guidelines for research practice in information systems and marketing," *International Journal of Information Management*, vol. 58, Article ID 102310, 2021.

[118] B. Kaya, A. M. Abubakar, E. Behravesh, H. Yildiz, and I. S. Mert, "Antecedents of innovative performance: findings from PLS-SEM and fuzzy sets (fsQCA)," *Journal of Business Research*, vol. 114, pp. 278–289, 2020.

[119] Y. Liang, G. Zhang, F. Xu, and W. Wang, "User acceptance of Internet of Vehicles services: empirical findings of partial least square structural equation modeling (PLS-SEM) and fuzzy sets qualitative comparative analysis (fsQCA)," *Mobile Information Systems*, vol. 2020, Article ID 6630906, 2020.

[120] C. Fornell and D. F. Larcker, "Evaluating structural equation models with unobservable variables and measurement Error," *Journal of Marketing Research*, vol. 18, no. 1, pp. 39–50, 1981.

[121] R. T. Cenfetelli and A. Schwarz, "Identifying and testing the inhibitors of technology usage intentions," *Information Systems Research*, vol. 22, no. 4, pp. 808–823, 2011.

[122] J. Li, J. Wang, S. Wangh, and Y. Zhou, "Mobile payment with alipay: an application of extended technology acceptance model," *IEEE Access*, vol. 7, pp. 50380–50387, 2019.

[123] S. T. Biucky, N. Abdolvand, and S. Rajaee Harandi, "The effects of perceived risk on social commerce adoption based on the tam model," *International Journal of Electronic Commerce Studies*, vol. 8, no. 2, pp. 173–196, 2017.

[124] L. Gao and X. Bai, "A unified perspective on the factors influencing consumer acceptance of internet of things technology," *Asia Pacific Journal of Marketing & Logistics*, vol. 26, no. 2, pp. 211–231, 2014.

[125] J. Wu, H. Liao, and J.-W. Wang, "Analysis of consumer attitudes towards autonomous, connected, and electric vehicles: a survey in China," *Research in Transportation Economics*, vol. 80, Article ID 100828, 2020.

[126] D. D. Gunawan and K.-H. Huarng, "Viral effects of social network and media on consumers' purchase intention," *Journal of Business Research*, vol. 68, no. 11, pp. 2237–2241, 2015.

[127] M. A. Almaiah, M. M. Alamri, and W. Al-Rahmi, "Applying the UTAUT model to explain the students' acceptance of

Mobile learning system in higher education," *IEEE Access*, vol. 7, pp. 174673–174686, 2019.

[128] C. Martins, T. Oliveira, and A. Popovič, "Understanding the Internet banking adoption: a unified theory of acceptance and use of technology and perceived risk application," *International Journal of Information Management*, vol. 34, no. 1, pp. 1–13, 2014.

[129] M. A. Stanko, "Toward a theory of remixing in online innovation communities," *Information Systems Research*, vol. 27, no. 4, pp. 773–791, 2016.

[130] P. Mikalef and A. Pateli, "Information technology-enabled dynamic capabilities and their indirect effect on competitive performance: findings from PLS-SEM and fsQCA," *Journal of Business Research*, vol. 70, pp. 1–16, 2017.

[131] S. F. Jahanmir, G. M. Silva, P. J. Gomes, and H. M. Gonçalves, "Determinants of users' continuance intention toward digital innovations: are late adopters different?" *Journal of Business Research*, vol. 115, pp. 225–233, 2020.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 5218254, 16 pages
https://doi.org/10.1155/2022/5218254

Research Article

Hotspots Identification and Classification of Dockless Bicycle Sharing Service under Electric Fence Circumstances

Ying Hui, Yingkun Xie, Qing Yu, Xiaolei Liu, and Xinyuan Wang

The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, College of Transportation Engineering, 4800 Cao'an Road, Shanghai 201804, China

Correspondence should be addressed to Qing Yu; 1810514@tongji.edu.cn

Received 18 December 2021; Revised 27 February 2022; Accepted 11 March 2022; Published 31 March 2022

Academic Editor: Linchuan Yang

Dockless bicycle sharing is one of the low-carbon transportation modes towards sustainable mobilities. Electric fences, as an effective solution for parking management, may have a high potential in guiding the usage of dockless bicycles at a low operation cost. However, new issues arise with the implementation of electric fences. The location of electric fences in hotspots fails to match the parking demand, leading the parking congestion in urban central areas. In this paper, a novel methodology of bicycle hotspots identification and classification is proposed to support parking management. An evaluation framework for bicycle hotspots is also proposed covering three aspects: demand and supply, unbalance, and land use. The methodology is applied to the case of Xiamen Island by using the trip data covering 53,629 bicycles during morning peak hours. Applying the methodology proposed, 47 pick-up hotspots and 53 return hotspots are identified, respectively. The total parking overload of return hotspots during the morning peak is 12,587 bicycles in Xiamen Island. The 53 return hotspots are classified into three clusters, including (1) hotspots where bicycle sharing is in overload status, (2) hotspots where bicycle sharing service quality needs to be improved, and (3) hotspots where bicycle sharing is in stable status. Based on the demand and land use characteristics, parking management schemes and policy implications are proposed. The result of this paper provides guidance for the layout of dockless bicycle sharing electric fences in different areas.

1. Introduction

Bicycle sharing (BS) is a popular short-distance travel mode all over the world, which promotes the healthy and low-carbon lifestyle of society. In the urban transportation system, it plays a key role in solving "the last-kilometer" problem and connecting to the public transit. Since the birth of the first generation of bike-sharing program "White Bike" in Amsterdam in 1965, BS has undergone several system updates [1]. The original bike-sharing systems were mainly docked ones, which means users need to return the bikes to fixed stations. With the development of information technology and the popularization of smart devices, the dockless bike-sharing system emerged in 2016, especially in China. The bicycle-sharing systems showed an oversupplied situation after a large number of BS operators entered the market, such as Mobike, HelloBike, and DiDiBike [2], and Xiamen is of no exception [3]. The regulatory authorities mandated in 2020 that the total number of shared bikes in Xiamen should not exceed 150,000 bikes, with 100,000 on Xiamen Island and 50,000 off the island, in order to reconcile the market order. Data shows that the average daily turnover rate of shared bicycles in Xiamen has remained at 3 to 3.5 times.

The dockless BS returning is more unregulated than docked BS, resulting in higher operation and maintenance costs. As a result of human movement patterns, the BS system demonstrated spatiotemporal unbalance, which refers to the gap between inflow and outflow in both time and space dimensions. Rebalance schemes were developed and enhanced with the use of big data models to solve the problem [4]. Rebalance implies redistribution to achieve a state where all stations have roughly equal proportions of bicycles to docks [5].

Hotspots, or places with high BS demand, received extra attention throughout the rebalancing process. In this paper, a BS hotspot is defined as an area where a large number of shared bicycles enter or exit in a short time. For BS operators, peak hour demand in hotspots is a challenge in terms of effective and timely repositioning. For BS users, the hotspots relate to the accessibility of sufficient bikes. For the government, the disordered parking situation has always been a problem. Subdivided region analysis and policy recommendations are required [6].

With the availability of high-precision positioning technology, the electric fence was introduced to BS in Xiamen on April 1, 2020. Electric fences, based on Global Position System (GPS) or Bluetooth, provide a solution for parking management. With the implementation of electric fences, however, new issues related to the BS operation arise. One of the challenges that BS is encountering is electric fence planning and design. While 15,202 electric fences are distributed in urban areas of Xiamen, there is a lack of assessment of the electric fence. For example, does the location of the electric fences matches the parking demand in the corresponding hotspot? Another issue is determining how to design a penalty for illegal parking and user incentives schemes to assist with parking management.

The main contributions of this paper are as follows:

(i) The methodology of bicycle sharing hotspot identification and classification is proposed

(ii) The hotspot evaluation framework covering demand & supply, unbalance, and land use is developed

(iii) A differentiated management strategy for bicycle sharing parking is provided

The graphical abstract is available in the supplementary file (available here). You can get a clear outline of the main findings of this paper by using this visual summary.

The remainder of this paper is organized as follows. Section 2 provides a brief overview of research related to this study. Section 3 presents the methodology framework of hotspots identification and classification. In Section 4, a case study in Xiamen Island is presented. Section 5 summarizes the main conclusions of this paper and proposes future directions.

2. Related Works

Since bicycle sharing systems have grown rapidly in the last two decades, particularly prospered again in the past five years in China, a large number of studies on BS have emerged. Recent researches primarily focus on demand prediction and fleet rebalance.

2.1. Demand Prediction. Accurate estimations for BS travel demands are crucial for successful bike management and fleet rebalance [7]. Generally, the demand prediction can be completed in two steps: the spatial unit setting and prediction algorithm design. Chen et al. [8] proposed a two-phase framework to predict overdemand clusters. In the first phase, a weighted correlation network was built to support the application of geographically constrained clustering. Feng et al. [9] developed a two-step demand prediction framework where the spectral clustering algorithm was adopted in the first step, station clustering. Similarly, in the research of Huang et al. [10], the Two-Stage Station Clustering algorithm was developed to cluster the central stations and common stations before predicting. In the context of dockless BS, Traffic Analysis Zone (TAZ) was used to predict the gap between inflow and outflow in the work of Xu et al. [11]. Traditional machine learning methods such as Random Forest (RF), Gradient Boosting Decision Tree (GBDT), and Supporting Vector Regression (SVR) were adopted early for demand prediction [9, 12]. With the development and popularization of neural network theory, deep learning methods have widely been used. Qin et al. [13] designed the Spatial-Temporal Bike Flow Prediction (ST-BFP) model, which is a convolutional network based on the residual framework. Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) were used to improve prediction performance [11, 14]. To capture the spatial correlation, Graph Convolutional Network (GCN) was introduced to build a spatiotemporal graph neural network model in work [15]. Another attempt to extract graph-based attributes was made by Yang et al. [7] to enhance short-term prediction. In summary, the spatial unit setting is the basis of demand prediction. Specifically, the spatial unit setting can be divided into three categories: station-based, station-cluster based, TAZ-based.

2.2. Bicycle Fleet Rebalance. Bicycle fleet rebalance refers to the process in which the bicycles are relocated from the oversupply area to the undersupply area. The bicycle rebalance problem (BRP) is crucial to the operation of the system.

Several studies have been devoted to the static BRP (SBRP), which means repositioning bicycles after the whole day's operation [16, 17]. SBRP is abstracted as mixed integer programming (MIP) in previous studies. According to Pal et al. [18], Mixed Integer Linear Program (MILP) could be utilized for solving the SBRP. The proposed MILP formulation can not only handle single and multiple vehicles but also allows for multiple visits to a node by the same vehicle. Liu et al. [19] applied mixed integer nonlinear programming (MINIP) formulation to solve multiple capacitated bike routing problems. Furthermore, an Adaptive Capacity Constrained K-centers Clustering (AdaCCKC) algorithm was proposed to separate outlier stations, reducing the large-scale multiple vehicle routing problems to the inner cluster one vehicle routing problem. Moreover, heuristic algorithms [20, 21] and greedy algorithms [22] have indicated the availability to optimize the rebalance process.

To further optimize the BS system and reduce maintenance costs, dynamic BRP (DBRP) is discussed and studied in depth. DBRP assumes that the whole system is constantly updated, so the rebalancing schemes should be continuously adjusted [23]. Therefore, demand prediction should be involved. Zhang et al. [24] proposed a nonlinear time-space

network flow model to fully integrate the user dissatisfaction estimating, the bicycle repositioning, and the vehicle routing. A more recent study [23] put forward a zone-based two-stage rebalancing method, dividing the research area into two types of zones (zones with deficient bikes (ZDB) and zones with sufficient zones (ZSB)). In the first stage, the ZDB only receives bikes from surrounding ZSB within the initial rebalancing range. In the second stage, based on the result of the previous stage, the remaining ZDB can receive bikes from the rebalancing range that is further away. SBRP and DBRP could be integrated for an advanced rebalance scheme. Tian et al. [25] designed a new framework to solve the rebalancing problem, which contains two aspects: dynamic rebalancing within each station and static rebalancing among stations.

User incentives-based BS fleet rebalance refers to optimizing the bike circulation by leading users to shift bikes between highly active stations and inactive ones [26]. With the interaction of sharing bicycles, users, and the environment, reinforcement learning could be adopted to determine the optimal incentive scheme [27, 28]. Otherwise, worker recruiting showed more cost-efficient compared with traditional truck-based rebalance [29].

2.3. Summary of Related Works. BS electric fences have been implemented in several cities in China. However, up to now, few works have studied electric fence planning, design, and operation. Zhang [30] firstly proposed a methodology framework to support electric fence planning. However, the study result remains at the assumption level, not under implementation. Xie et al. [31] combined the location-based electric fence and image-based bicycle parking place identification to verify the users' parking behavior. Yuan et al. [32] developed high precision virtual electric fence technology based on GNSS, intelligent terminal, and precision orientation algorithm, which achieved submeter positioning precision.

To summarize, as BS advances to a new stage of development, the variety and efficiency of system optimization methods have been improved. For both demand prediction and fleet rebalance, there is a need for accurate division and classification of the operation area, particularly the hotspots. The emergence of the electric fence shows the potential for regulating users' parking behavior. The user incentive scheme provides new ideas for user-based fleet rebalance.

However, to the best of our knowledge, there is little literature integrating the identification and classification of bicycle sharing hotspots with consideration of the electric fence. Policy guidance on the electric fence planning and parking management scheme remains unclear in response to the demands of various stakeholders.

Therefore, in this paper, a novel methodology framework for identifying and classifying BS hotspots is proposed. For the first time, electric fence data, regarded as a parking supply, is used in hotspots analysis. In addition, recommendations on the electric planning and design are given based on the hotspots classification results. The penalty for illegal parking is discussed, as well as user incentive schemes.

3. Methods

3.1. Methodology Framework. The methodology framework of this study is shown in Figure 1. Multisource data including BS trip data, electric fence data, and POI data is collected in this paper to support BS hotspots identification and classification. The methodology framework is as follows.

Task 1: BS hotspots identification. The origin (O) and destination (D) of each trip are extracted from BS trip data. To achieve high-precision recognition, the spatial resolution is set to 50 m × 50 m grids. The longitude and latitude of each grid center and the O/D number of each grid are packed as the input of DBSCAN. The determination of neighbor parameters of DBSCAN is done using the KNN method. Finally, the pick-up hotspot (PH) and return hotspot (RH) are identified separately.

Task 2: BS hotspots classification. The return hotspots (RHs) are the critical areas for rebalancing the BS fleet. Meanwhile, the government also pays great attention to the bicycle parking order in RHs. Therefore, the RHs obtained from Task 1 are chosen as the study subject of Task 2's research. To measure the RHs in BS parking demand, parking supply, unbalance, and land use, nine indicators are extracted from the collected data. RHs are classified using the Gaussian mixture clustering method based on the 9 indicators. The classification result is used to provide guidance on parking management schemes and policy recommendations.

3.2. Identification of BS Hotspots. A BS hotspot is defined here as an area where a large number of shared bicycles flow in or out in a short time. The main challenge faced by BS operators and regulators is the hotspot identification and corresponding management strategy. In this paper, we attempt to answer two questions through hotspots identification: How are hotspots distributed in space? Is there a connection between the spatial distribution of hotspots and urban structure?

The key characteristics of hotspots can be listed as follows:

(1) Closed and continuous in space

(2) High level of demand

(3) Temporal dynamic change in demand

Based on the characteristics above, the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) method is adopted with O/D (origin/destination) demand as a clustering feature to identify BS hotspots.

DBSCAN is an unsupervised machine learning algorithm. It is a density-based clustering nonparametric algorithm. The algorithm uses a simple minimum density level estimation based on a threshold for the number of neighbors, minPts, within the radius ϵ (with an arbitrary distance measure) [33]. Compared with the traditional kernel density estimation method, DBSCAN can help identify the boundary of the hotspot area accurately according to the neighbor parameters. The accuracy of identification results

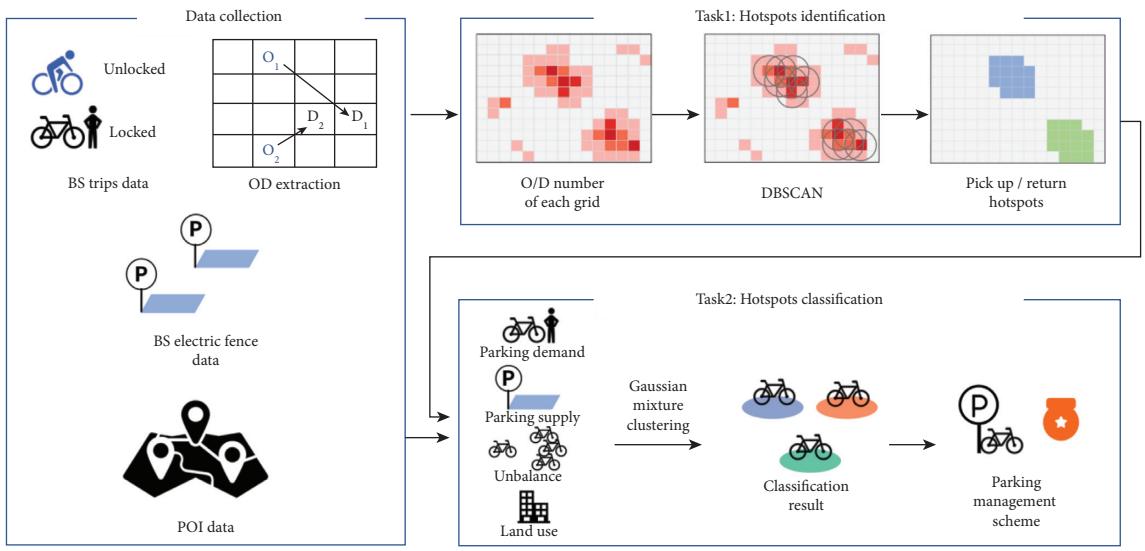


Figure 1: Methodology framework.

depends on the two neighbor parameters: the radius ϵ and the number of neighbors (or total weight), minPts.

The longitude and latitude of each grid center are packed as the training instances X, with the BS O/D (origin/destination) demand of each grid being the sample weight. For parameter estimation, firstly, the number of neighbors minPts should be given according to domain knowledge. Then the radius could be determined using the k-distance plot. The average distance corresponding to the elbow point (with the largest slope) is the optimal radius ϵ.

3.3. Evaluation Framework of BS Return Hotpots. In previous sections, BS return hotspots (RHs) could be identified with high precision. In this section, we develop an evaluation framework covering three aspects (supply and demand, unbalance, and land use) to assess the characteristics of RHs and take it as the basis for classification, as shown in Table 1. The following paragraphs explain why certain aspects are chosen.

For supply and demand, it is a crucial independent input in many BS planning problems, such as bicycle relocation, demand management [34]. Çelebi et al. [35] concluded that one of the key motivating factors for bicycle use is the convenience that the service is not interrupted because of the unavailability of empty slots for returns, emphasizing the importance of evaluating the parking supply for shared bicycles. Furthermore, demand level analysis could support the layout planning of electric fences [30].

In terms of unbalance, while BS facilitates people's first/last-mile travel, it can also have negative effects such as infringing on pedestrian rights, blocking bike paths, and impeding metro user flows [36–38], which is caused by spatial-temporal unbalance.

Regarding land use, previous studies have shown that land use factors have an effect on bicycle sharing demand [2, 39–42]. Xing et al. [43] discovered that the bike-sharing trip origin and destination could be divided into five typical groups, i.e., dining, transportation, shopping, work, and residential places on weekdays. In this paper, we choose three travel purposes from five categories: residential, work, and shopping places.

3.3.1. Demand and Supply. BS parking supply is an important index for RHs assessment. Compared with docked BS, the parking of dockless BS tends to be more dispersed. Bicycle returning has become more restricted since the implementation of electric fences in the real world. The electric fences that specify where the bikes can be returned can be regarded as the parking supply. According to *Planning Standard for Urban Pedestrian and Bicycle Transportation System* published by the Ministry of House and Urban-Rural Development (MOHURD) of China in 2021, it was suggested that the width and length of a single bicycle parking space should be 0.6–0.8 m and 2.0 m, respectively, equivalent to the area of 1.2–1.6 m^2. In this paper, we determine that the parking area for each bicycle should be 1.5 m^2. Using the area of electric fences in certain RH and the parking area for one single bicycle, the parking capacity P_c can be calculated.

BS demand is not only the basis for hotspots identification but is also a key index of classification. Although hotspots represent areas with high demand, the demand among these hotspots is still uneven. The BS demand is calculated as follows. Taking a record of OD data (bikeid, origin, destination, stime, etime) as an example, it will generate a departure demand at origin, and an arrival demand at destination. Given a time window Δt, the end of the time window t, and a hotspot H with r bikes inside, there are F_{out} BS trips departing from H, and F_{in} trips arriving in H. The parking demand P_d of hotspot H at time t can be calculated as follows:

$$P_d = r + F_{\text{in}} - F_{\text{out}}. \tag{1}$$

Then, the parking overload of hotspot H at time t can be calculated by the following equation:

Table 1: The evaluation indexes of RHs.

Aspect	Indicators	Name of indicators	Description
Supply	P_c	Parking capacity	The maximum number of bicycles that can be parked in this area
Demand	A_d	Arrival demand	The total number of trips whose destination is in this area
	D_d	Departure demand	The total number of trips whose origin is in this area
	P_o	Parking overload	The maximum difference between the number of parking bicycles and the parking capacity
Unbalance	R_m	Min-max ratio	The ratio between the minimum and the maximum of parking bicycles
	P_N	Peak time no.	The no. of the time when the number of parking bicycles reaches its peak
Land use	L_R	Residential POI	The number of residential POIs in this area
	L_O	Office POI	The number of office POIs in this area
	L_C	Commercial POI	The number of commercial POIs in this area

$$P_o = P_d - P_c, \tag{2}$$

where P_d, P_c, and P_o denote the parking demand, parking capacity, and parking overload, respectively, F_{in}, F_{out} represent BS trips arriving in hotspot H and trips departing from hotspot H, and r is the number of existing bicycles of hotspot H.

3.3.2. Unbalance. The spatial-temporal heterogeneity of BS demand and supply leads to the problem of unbalance. Peak time is defined here as the time when the number of parking bicycles reaches a maximum. Min-max ratio represents the ratio between the minimum and the maximum of parking bicycles.

To obtain a better understanding of the unbalance in hotspots during the morning peak, there is a need to study the peak time P_N and min-max ratio R_m. Given time series T_i and hotspot H, the number of parking bicycles P_d varies over time. Suppose its maximum and minimum in T_i are $P_{d\max}$ and $P_{d\min}$, the min-max ratio R_m can be calculated as follows:

$$R_m = \frac{P_{d\min}}{P_{d\max}}. \tag{3}$$

3.3.3. Land Use. BS trips, like the other modes of transportation, are activity-based. The diversity and intensity of activities may be influenced by land use factors. Previous studies have shown that the frequencies of POIs can disclose and classify urban land use [43, 44]. Therefore, incorporating land use indicators into the evaluation framework makes sense. Here, we used the number of residential POI L_R, office POI L_O, and commercial POI L_C to describe land use attributes.

3.4. Classification of BS Return Hotspots. The Gaussian mixture model (GMM) is used to classify the RHs after extracting nine indicators from multisource data. The choice of GMM clustering is motivated by two considerations. On the one hand, we preliminary investigated the performance of GMM-based classification methods on the classification problems in many application scenarios [45, 46]. Under complex data, GMM has a wider fitter range and the advantage of high generalization ability. On the other hand, inspired by traffic status classification in Liu et al.'s work [47], we chose GMM to cluster the RHs.

GMM is a common algorithm of prototype clustering in machine learning. GMM uses the Gaussian probability model as a clustering prototype. The definition of the multivariate Gaussian distribution is as follows [48]: Given random variable $\mathbf{x}$ in n-dimensional sample space χ, if $\mathbf{x}$ obeys Gaussian distribution, its probability density function is shown in the following equation:

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} e^{-1/2(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}, \tag{4}$$

where μ is an n-dimensional mean vector and Σ is $n \times n$ covariance matrix. To clearly show the dependence of Gaussian distribution on corresponding parameters, the probability density function is denoted as $p(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Then the Gaussian mixture distribution can be defined as follows:

$$p_{\mathcal{M}}(\mathbf{x}) = \sum_{i=1}^{k} \alpha_i \cdot p(x|\boldsymbol{\mu}_\mathbf{i}, \boldsymbol{\Sigma}_\mathbf{i}). \tag{5}$$

The distribution consists of k mixture components, and each mixture component corresponds to a Gaussian distribution. In equation (5), $\mu_\mathbf{i}$ and Σ_i are the parameters of the ith Gaussian mixture component, and $\alpha_i > 0$ is the corresponding mixture coefficient. It should be noted that $\sum_{i=1}^{k} \alpha_i = 1$.

The indicators of RHs are all calculated as the input of the algorithm. The RHs with nine indexes are regarded as the training set $D = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$. The random variable $z_j \in \{1, 2, \ldots, k\}$ represents the Gaussian mixture component of the sample $\mathbf{x}_j$. The objective of GMM is to divide the set D into k clusters $C = \{C_1, C_2, \ldots, C_k\}$. The cluster label λ_j of each sample $\mathbf{x}_j$ is determined as follows:

$$\lambda_j = \underset{i \in \{1,2,\ldots,k\}}{\operatorname{argmax}} \gamma_{ji}, \tag{6}$$

where $\gamma_{ji}\,(i = 1, 2, \ldots, k)$ denotes the posterior probability generated by Gaussian mixture components $p_{\mathcal{M}}(z_j = i|\mathbf{x}_j)$.

4. Results and Discussion

4.1. Case Study and Datasets. In this paper, the BS system on Xiamen Island (Xiamen, China) is selected as the study object. In December 2016, the bike-sharing service entered Xiamen and began operations. In 2017, the oversupply of

shared bikes had caused widespread concern in society. Until now, with the promulgation of legislative documents, the amount of shared bicycles has been limited to 150,000. The Xiamen Island, surrounded by the sea, can be considered as a closed area for sharing bicycle operation (Figure 2).

Multisource data including BS trip data, BS electric fence data, and POI data were used for subsequent modeling and analysis. The BS trip data and electric fence data were obtained from the Digital China Innovation Contest (DCIC 2021) official websites. A record of trip data contains the bicycle ID, latitude, longitude, lock status, and update time. When a user picks up or returns the bicycle, a record with time and location status is generated and sent to the server. The dataset includes 198,382 trips during the morning peak (6:00–10:00) of five consecutive weekdays from December 21 to 25, 2020. Although it is winter, the local average maximum temperature is around 20°C, which is suitable for cycling. Furthermore, most of the data processing in this paper is supported by the Python package TransBigData [49].

The descriptive analysis was conducted, and results showed a sharp decline in the number of orders on 2020/12/23 (Figure 3(a)), which was influenced by the moderate rain that was not suitable for cycling. Therefore, the OD data of the other four weekdays will be used for subsequent analysis. The number of orders reaches its peak at 08:10–08:20 (Figure 3(b)). 3/4 of the BS travel distances are within 1,200 meters (Figure 3(c)), which shows that BS is indeed the mode of transportation for the first/last mile. Besides, as shown in Figure 3(d), while some trips last 20 minutes or more, the majority (3/4) of trip duration is within 10 minutes.

The BS electric fence data includes the geographic boundaries of 10,471 electric fences, and each electric fence is a closed quadrilateral connected by five points (Figure 4).

Point of interest (POI) data was obtained from the Baidu web API service. The residential area and commercial & office area are the two most important factors influencing the BS usage frequency [41]. These areas are populated areas with high activity intensity. Therefore, in this paper, three types of POI, residential, commercial, and office, are used as the land use attributes of hotspots.

4.2. Results of Hotspots Identification. A BS hotspot, as defined in methods, is an area where shared bicycles flow in and out in large numbers over a specific time. We define two types of hotspots in this paper: pick-up hotspots (PHs) and return hotspots (RHs), which refer to areas with high pick-up demand and areas with high return demand, respectively.

Prior to identifying hotspots, OD pairs were assigned to the corresponding space unit, with spatial resolution set to 50 m × 50 m grids. The average number of bicycles picked up and returned in each grid was then calculated for the entire morning peak (6:00–9:59) of each weekday.

To identify BS hotspots, the two parameters of DBSCAN, epsilon (ϵ) and minimum samples (minPts), may highly affect the number and the size of hotspots, which should be determined in advance. For selecting minimum samples, there is no automatic approach. Based on domain knowledge in urban traffic, minPts was set to 150. K nearest neighbors (KNN) method was applied then to find the best ϵ value. By plotting the average k-distances (where k = minPts = 150) in ascending order (Figure 5), the distance corresponding to the elbow of the curve was the best epsilon, that is $\epsilon = 2.5$.

After DBSCAN parameter estimation, the longitude and latitude of the grids are packed as the input of the clustering algorithm. Then, pick-up and return demand are set as sample weights, respectively, to identify the BS hotspots in both situations. It reported 47 pick-up hotspots and 53 return hotspots on Xiamen Island, with different centers indicated by different colors (Figure 6).

There are some interesting findings. Figure 6 shows that 48.8 percent of pick-up hotspots and 37.7 percent of return hotspots are located near metro lines, indicating that BS plays an important role in connecting to the metro transit system and that bicycle trips beginning at metro stations are more frequent. Figure 7 shows, for example, that demand for pick-up and return is high at metro stations during the morning peak on weekdays.

While the demand for bicycle pick-up and the return has reached equilibrium in some areas, which appear in Figures 6(a) and 6(b) simultaneously, there are still some other areas with one-way bicycles inflow or outflow. For example, Figure 8(a) shows that the pick-up hotspot near *Dianqian Road 1st Road*, surrounded by multiple residential districts, failed to become a return hotspot. However, there are two return hotspots on Huli Avenue, where numerous office buildings are located nearby. One possible explanation for this phenomenon might be the separation of work and residence. During morning rush hours, a considerable number of BS users begin their journey from home to work. If the number of job opportunities mismatches the number of employed people in a certain area, the population flow may show one-way inflows or outflows, resulting in the phenomenon mentioned above. Figures 8(b) and 8(c) also present the one-way inflow and outflow in the hotspots, respectively. The observed result could explain the spatial heterogeneity to a certain extent.

4.3. Result of Hotspots Classification. Return hotspots represent the places where a substantial proportion of BS users return shared bicycles. The total parking overload of 53 RHs during the morning peak is 12,587 bicycles, while the total parking demand is 82,677 bicycles, indicating an insufficient parking supply in return hotspots. Further investigation of destination (return behavior) distribution could aid in gaining a better understanding of human mobility during morning peak hours and the role of the BS in the transportation system.

53 RHs are classified into three clusters using the Gaussian-mixture clustering algorithm. The spatial distribution of three types of hotspots is depicted in Figure 9(a). The average normalized scores of multiple indexes for the three clusters are shown in Figure 9(b). The result, depicted in Figure 9(b), indicates the heterogeneity of parking

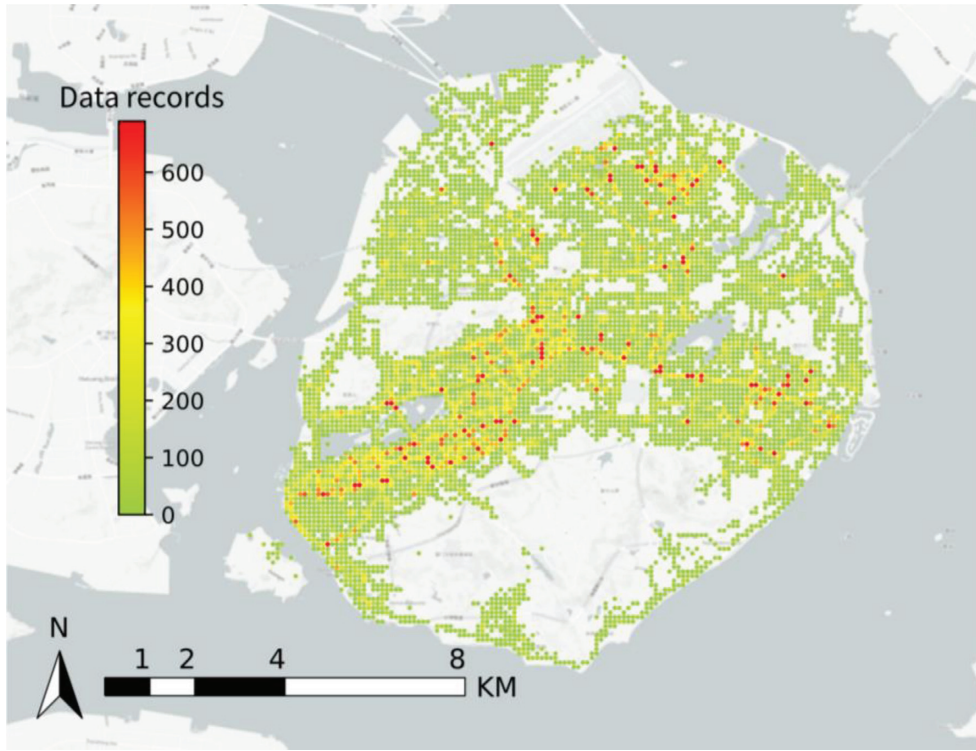


Figure 2: BS spatial distribution in Xiamen Island.

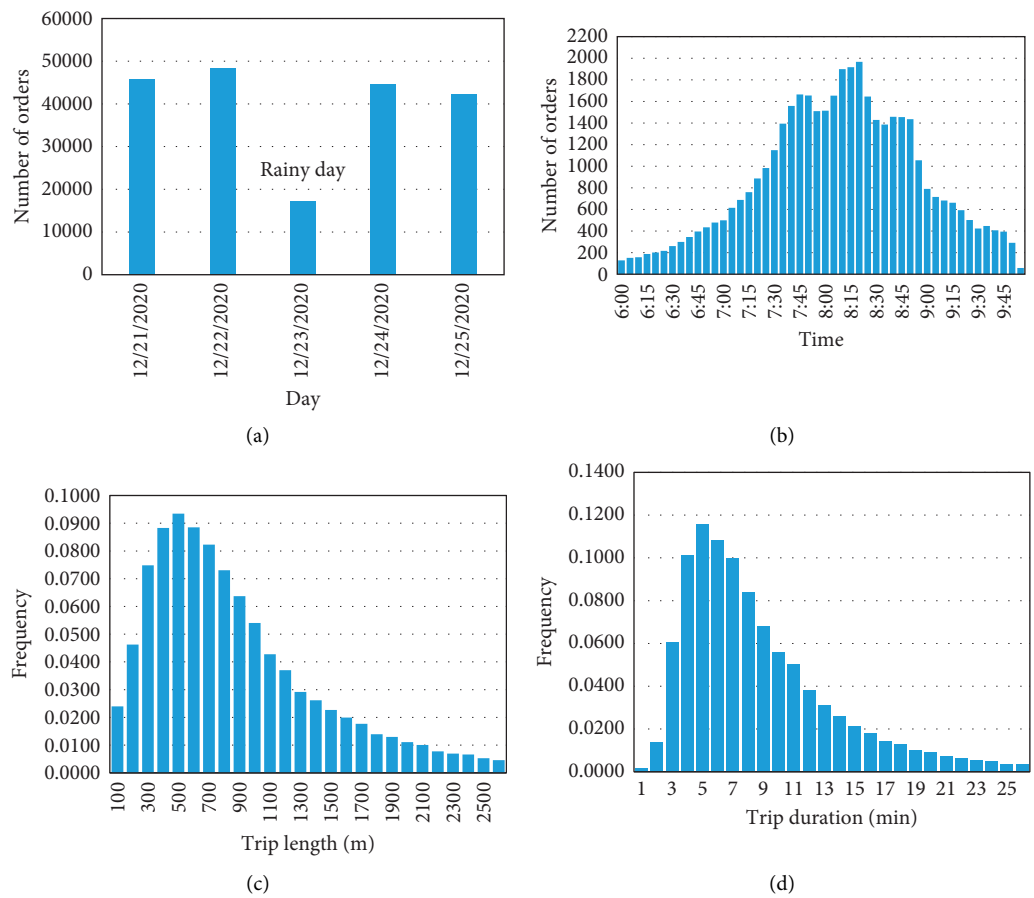


Figure 3: BS trip characteristics. (a) Daily distribution. (b) Temporal distribution. (c) Travel distance distribution. (d) Trip duration distribution.

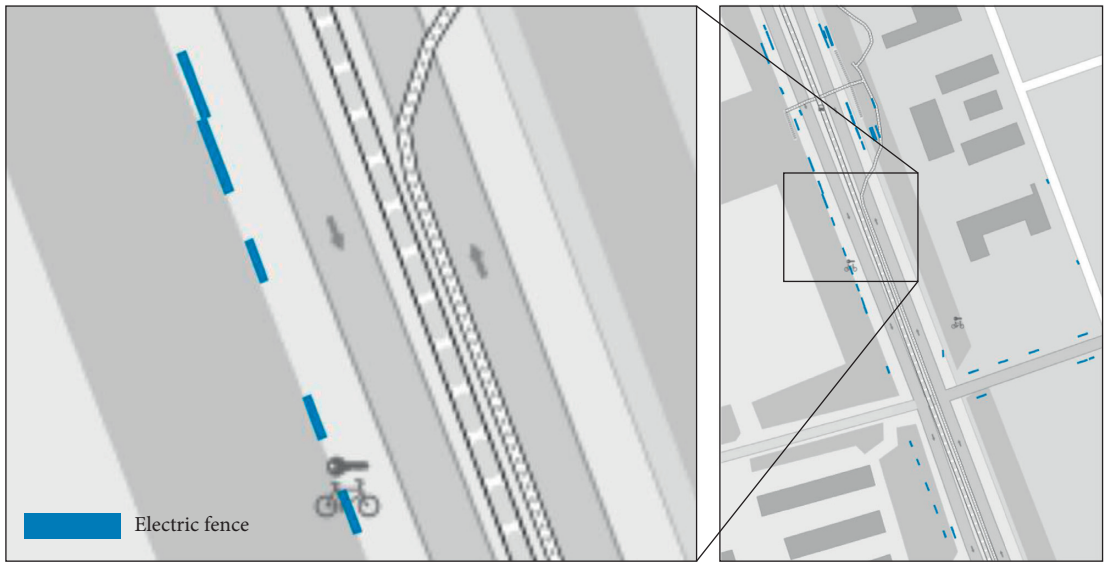


Figure 4: BS electric fence geographic data.

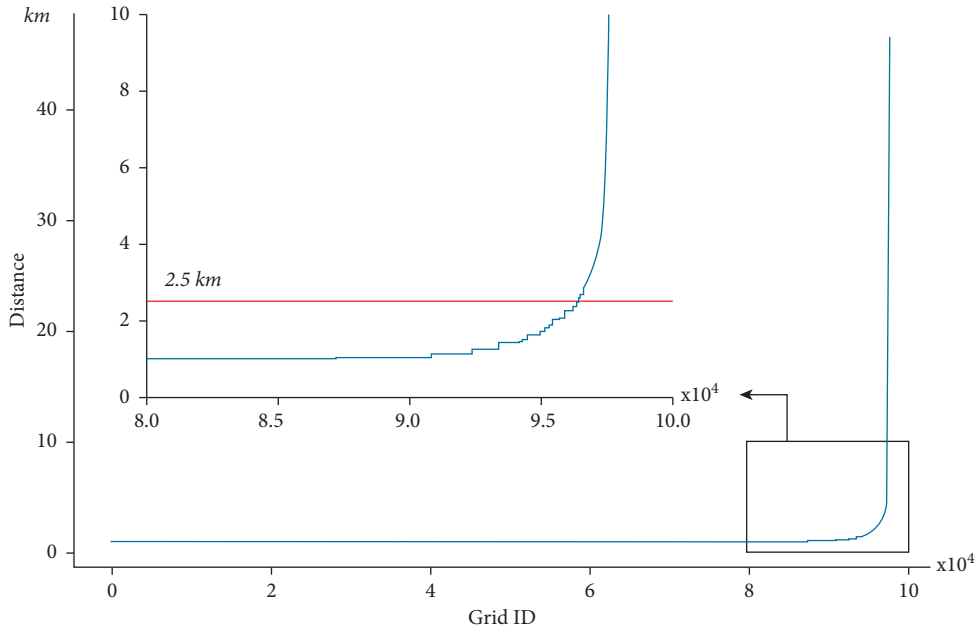


Figure 5: Points sorted by distance to the kth nearest neighbor.

demand, supply, and land use in the three return hotspot clusters. The unbalance characteristics (minmaxratio, peaktimeno) of the three clusters, on the other hand, show little variation. The gaps between the minimum and maximum BS demand during the morning peak hours, in particular, are nearly the same for the three clusters. Parking bicycles is at a high level in RHs labeled as Cluster 1 or Cluster 2, as shown in Figure 9(c), indicating that more regulatory attention should be paid to the two clusters.

4.3.1. Cluster 1: Hotspots Where BS Is in Overload Status. Cluster 1 accounts for 3.8% (2/53) of all return hotspots. These are typically areas with a high demand for BS service. Despite having a relatively large parking capacity, it is unable to meet the demand for bike-sharing. In other words, there is a great gap between parking demand and supply in these hotspots. In terms of land utilization, this category has more POI than the other two, indicating that these areas are more developed. At the same time, the proportion of office POI accounts for more than that of residential and commercial POI, resulting in high demand for the last mile to work. For example, Huli Innovation Park is home to numerous high-tech enterprises, and bike-sharing is used as the main mode for employees to enter the park.

In general, the centralized distribution of office POI, such as office buildings, and the undiversified traffic mode structure have resulted in the situation described above.

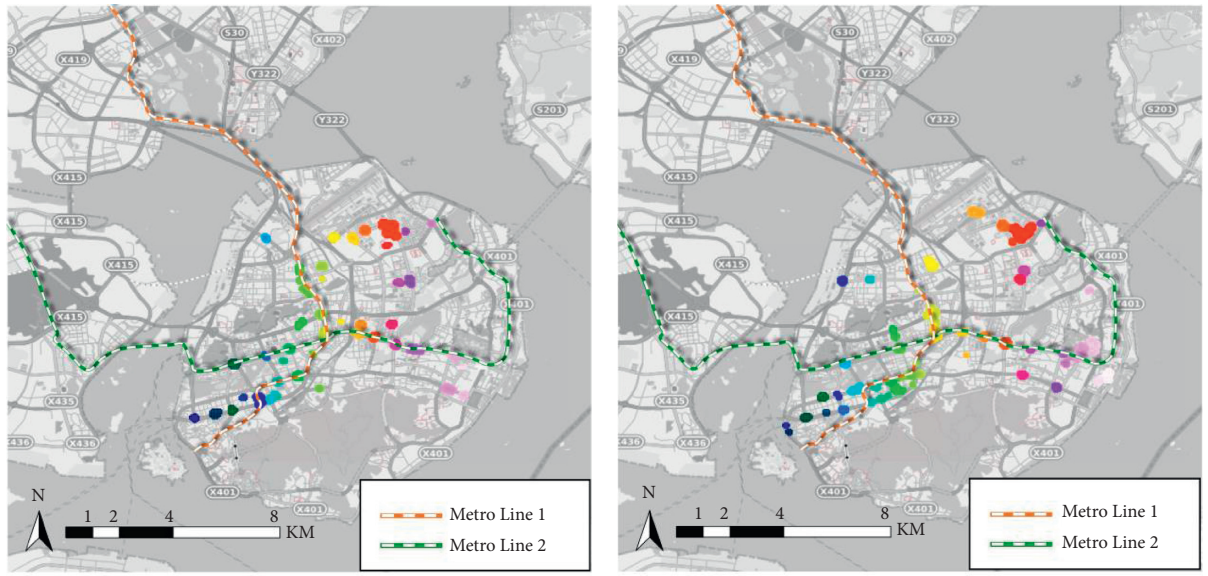


Figure 6: Pick-up hotspots and return hotpots. (a) Pick-up. (b) Return.

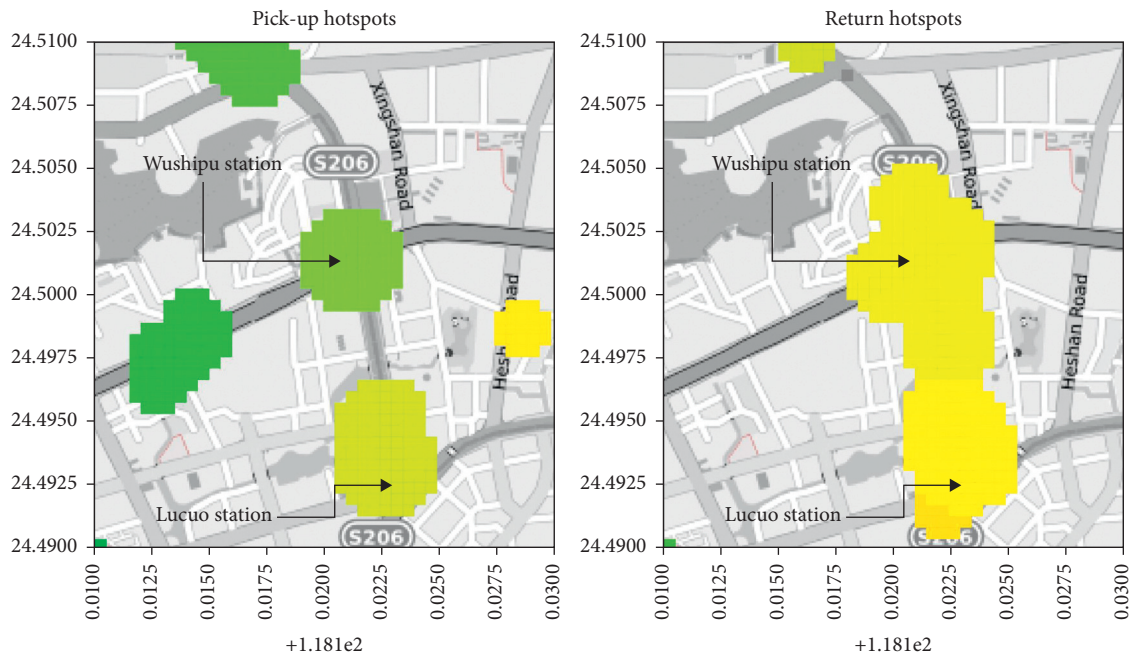


Figure 7: Hotspots near the metro station.

The BS is in overload status at this type of hotspot. It appears difficult to solve the problem solely from the perspective of BS.

4.3.2. Cluster 2: Hotspots Where BS Service Quality Needs to Be Improved. 18.9% (10/53) of return hotspots are labeled as Cluster 2. The demand for bike-sharing is medium in these return hotspots. Parking overload is also at a medium level due to a lack of electric fences (parking capacity). For example, Lücuo metro station, located in the city center, is a transfer hub with ten entrances for Lines 1 and 2. People can reach any metro station on Xiamen Island without making any stops from this metro station. As a result, a large number of BS orders were terminated at this location. On average, land use is fairly balanced. The proportion of residential POI is similar to that of commercial POI but slightly higher than that of office POI.

To summarize, the designated electric fence in Cluster 2 cannot meet the parking demand due to limited roadside

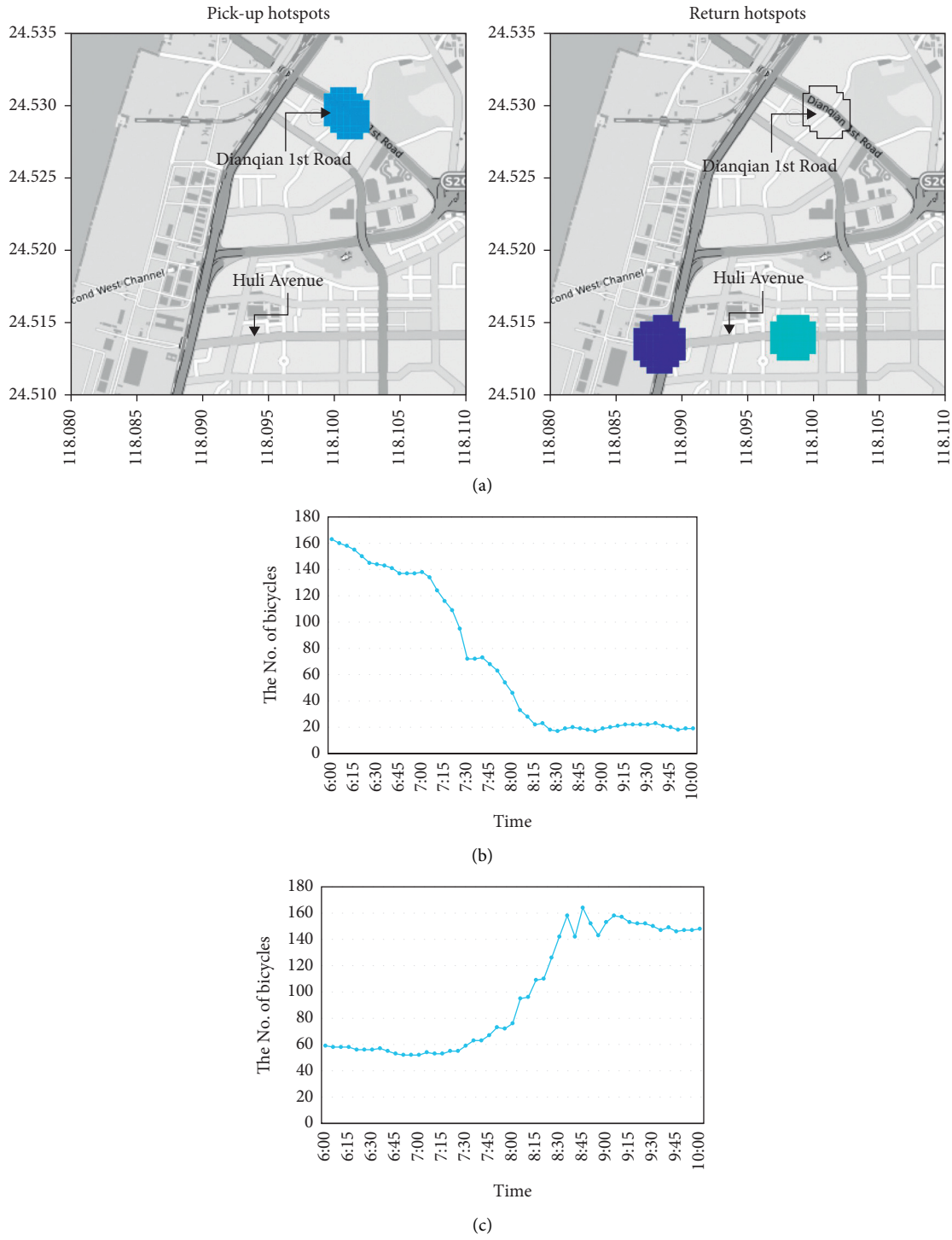


Figure 8: Outflow and inflow of shared bicycles in Dianqian 1st Road and Huli Avenue. (a) Hotspots near Dianqian 1st Road and Huli Avenue. (b) The no. of bicycles near Dianqian 1st Road. (c) The no. of bicycles near Huli Avenue.

space. This type of hotspot is the most vulnerable to BS electric fence supply. In other words, if more or larger electric fences could be assigned to the hotspot, parking congestion would be significantly reduced. BS service quality should be improved in Cluster 2 return hotspots by increasing parking supply and strengthening parking management. Otherwise, if users are unable to return their bikes to the electric fences, there is a risk that shared bicycles will occupy the pedestrian space.

4.3.3. Cluster 3: Hotspots Where BS Is in Stable Status. 77.3% (41/53) of return hotspots are classified into Cluster 3. In such areas, the demand level for bike-sharing is relatively low. Although the parking capacity is insufficient, the

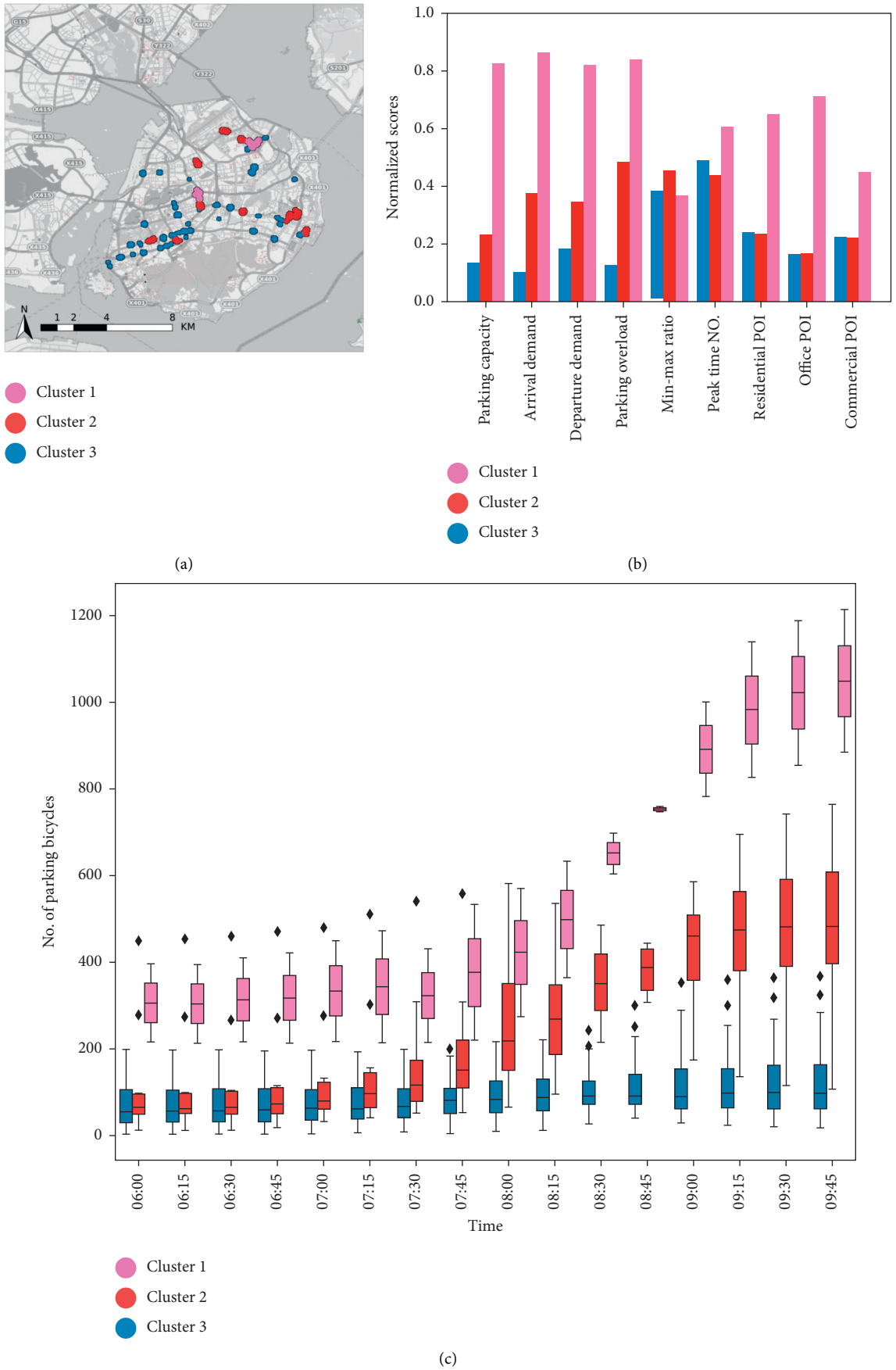


Figure 9: Hotspots classification result. (a) Spatial distribution of different types of hotspots. (b) Average normalized scores of indexes. (c) No. of parking bicycles in RHs.

parking gap is still small. As shown in Figure 9(a), the size of this type of hotspot is rather smaller than the other two. The average hotspot in this category is equivalent to a circle with a radius of 185 m. In addition, the land use of hotspots is nearly balanced, similar to Cluster 2.

Overall, in Cluster 3 hotspots, the BS is in stable status with little parking overload. However, the dispersion of return hotspots tends to increase potential operation and maintenance costs.

4.4. Discussion of Parking Management Schemes and Policy Implications. The emergence of electric fences shows the potential for regulating parking behavior among users. At the moment, the shape of electric fences is uniform, with each fence being a closed quadrilateral connected by five points. The planning and designing of electric fences need to be optimized in combination with the supply and demand characteristics of hotspots.

Here, according to the demand intensity and management requirement, we propose three recommended electric fence layouts (Figure 10): the first type is a large continuous block, the second is a continuously banded area, and the third is a dispersed banded area.

Parking management schemes are classified into two types: illegal parking penalties and user incentives. The penalty for illegal parking means that if a user fails to return the bicycle to the electric fence, the BS App will notify the user. If the user continues to park outside the electric fence, a fine will be imposed. User incentives imply that, prior to a bicycle-sharing trip, the app can recommend an electric fence with adequate parking ability near the destination, supplemented by certain incentive rewards (such as free tickets, coupons, etc.). It should be noted that parking management schemes should be integrated with the design of electric fences.

The recommended electric fence layout and parking management schemes for three clusters of RHs will be described in the below sections.

4.4.1. Cluster 1: Hotspots Where BS Is in Overload Status. For BS operators, the electric fences could be designed as large continuous areas to deal with parking demand during morning peak hours. On the other hand, the inelastic demand for bicycle sharing may lead to the failure of parking management schemes. It's worth noting that there are more office POIs in this type of hotspot, which may cause the tidal-flow phenomenon. Considering the bicycles stuck here may affect the sidewalk, it is recommended to move the bicycles out of the area after the morning peak to avoid high parking overload.

For the government, considering the great demand for short-distance travel in this area, in addition to bicycle-sharing, alternative modes of transportation such as shuttle buses could be introduced to serve the residents in this type of area.

4.4.2. Cluster 2: Hotspots Where BS Service Quality Needs to Be Improved. Due to the concentrated parking of bike-sharing in such areas, it is more appropriate to set up a continuously banded area of electric fences according to the roadside environment. In terms of parking management, a parking incentive scheme (coupons, free of charge, etc.) could be used to guide users to park bicycles to the adjacent areas where electric fences are still available. For operators, this type of area is suitable for the development of sharing bicycle services. Therefore, it is necessary to strengthen the dispatching management and match the travel demand in this type of area.

For the government, the active traffic environment in such areas should be further improved, such as expanding the width of bicycle lanes. Furthermore, the bicycle-riding space should be isolated from the vehicle space for safety.

4.4.3. Cluster 3: Hotspots Where BS Is in Stable Status. Given the dispersed nature of shared bicycle parking in this area, the location of electric fences should be integrated with the bus station and metro station. Furthermore, electric fences should be set as a dispersed band along the roadside, where maintenance personnel can easily collect bicycles. It is also recommended that parking penalty schemes be implemented to regulate users' parking behavior, particularly parking outside the electric fence, which can be considered illegal parking behavior.

Despite the demand level of this type of area being low, the peak hours and demand valley-peak ratio are similar to the other two types. As far as the government is concerned, this kind of hotspot is the vast area where the "the first/last mile" problem exists. The government should supervise BS enterprises to launch appropriate bicycles in these places to ensure the public's short-distance travel.

4.5. A Case Study of BS Return Hotspot Parking Management Scheme. The classification of return hotspots based on supply-demand relationships, unbalanced characteristics, and land use attributes offer the potential for more precise parking management schemes. Take Huli Innovation Park as an example, which houses 6,869 businesses and provides 66,000 jobs. The total number of bicycles parked in this area is 1802. Figure 11(a) depicts the spatial distribution of parking congestion.

As shown in Figure 11(b), there are three clusters of RHs in this park. In the hotspot labeled as Cluster 1 inside the park, commuters return shared bicycles here, resulting in a sharp increase in the number of parking bicycles after 8:00 AM (Figure 11(c)). It should be noted that this trend is opposite to the hotspot near the Xiamen Air Bike Lane entrance. This case might reveal the phenomenon that the shared bicycles flow in the office park during the morning peak. Moreover, as can be seen in Figure 11(a), the parking overload appears to be higher at the gate of office buildings. This finding is likely to be related to the closed management of office buildings. Xiamen Air Bike Lane entrance, marked as Cluster 2, is located at the edge of the park. In this hotspot, people can get access to bicycle service through the Air Bike Lane, which is an elevated road only for bicycles. Therefore, the availability of shared bicycles leads to a growth of parking bicycles before 8:00 AM. However, the number of

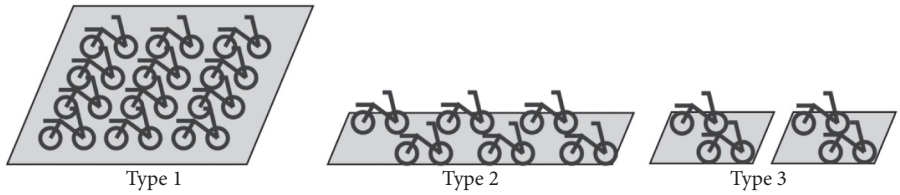


Figure 10: Recommended electric fence layout.

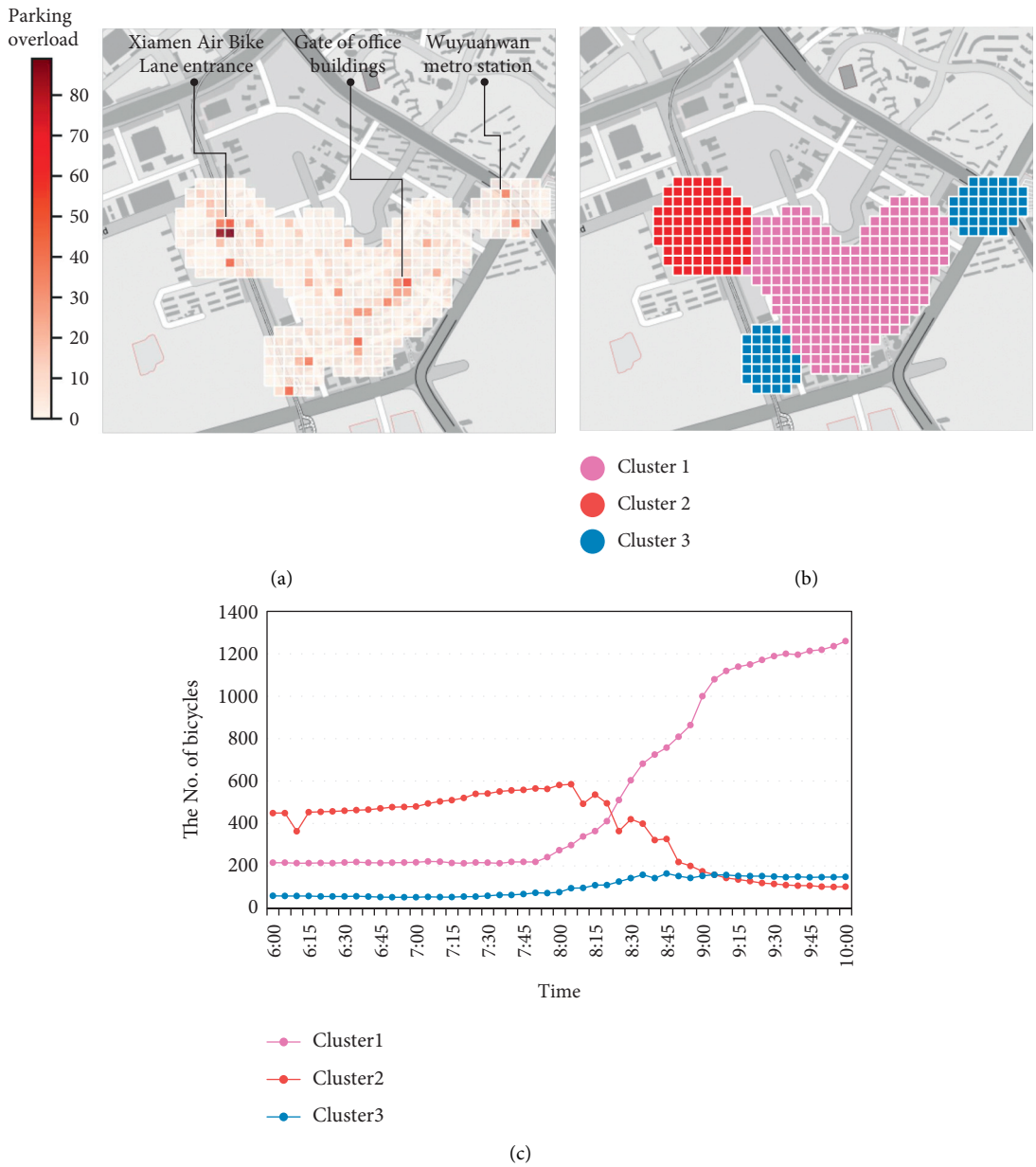


Figure 11: Hotspots in Huli Innovation Park. (a) Parking overload. (b) RHs classification result. (c) Parking bicycles.

bicycles decreased sharply from 585 to 102 until 10:00 AM, as shown in Figure 11(c). Two hotspots classified into Cluster 3 are also at the edge of the park. Considering the time-varying characteristics of parking bicycles, the BS is in stable status without much parking overload. The number of bicycles parked in these two hotspots reached its peak at 8:45 AM (Figure 11(c)), then remained at around 150 bicycles.

In the hotspot labeled as Cluster 1 inside the park, there is a need to design large continuous areas as electric fences in such hotspots to facilitate commuters' parking. Particularly at the gate of office buildings, the real-world double-layer bicycle rack could be introduced to expand the capacity of the electric fences. Moreover, based on the phenomenon that parking bicycles continue to increase during the morning peak, BS

enterprises could transfer bicycles out of the hotspot to reduce the parking overload. For the hotspot labeled as Cluster 2, the availability of the Air Bike Lane enhances the usage frequency of shared bicycles. To avoid high parking overload, the electric fences here should be expanded in combination with the location of the entrance of the Air Bike Lane. Besides, some bicycles could be supplemented to the hotspot to meet the needs of commuters to enter the park by bike. In the two hotspots classified into Cluster 3, it is not necessary to greatly expand electric fences. Instead, the electric fence should be placed near the bus station and the metro station to promote transfer connection with public transit. In addition, part of the electric fences in the hotspot could be set closed so that bicycles would be returned in a few opened electric fences.

5. Conclusions and Future Directions

In this paper, a novel framework for integrating multisource data is proposed to support BS parking system design and management. The definition of a BS hotspot is given. DBSCAN method is adopted with O/D (origin/destination) demand as a clustering feature to identify BS hotspots. In particular, we identify 47 pick-up hotspots (PHs) and 53 return hotspots (RHs), respectively, and conduct a comparative analysis. An evaluation framework covering three aspects (demand and supply, unbalance, and land use) is developed to assess the characteristics of RHs. The evaluation framework consists of 9 indexes derived from multisource data. Furthermore, for the first time, real-world bicycle-sharing electric fence data is used as a parking supply. The investigation of demand and supply in return hotspots shows that the total parking overload of 53 RHs during the morning peak is 12,587 bicycles. Using the Gaussian-mixture clustering algorithm, 53 RHs are classified into three clusters, including hotspots where BS is in overload status, hotspots where BS service quality needs to be improved, and hotspots where BS is in stable status. Overall, there is a lack of parking capacity in RHs on Xiamen Island. Based on the demand&supply pattern, the land utilization, and the unbalance feature, parking management schemes and policy implications are proposed.

The main findings of this paper are as follows:

(i) The study has found that bicycle sharing hotspots do exist from the perspective of bicycle pick-up and bicycle return.

(ii) The results of hotspots identification have revealed that bicycle sharing plays an important role in the connection to the metro transit. Additionally, bicycle trips starting from metro stations are more frequent.

(iii) The difference in the spatial distribution between pick-up hotspots and return hotspots provides insights into the one-way bicycle inflow or outflow phenomenon. An analysis of return hotspots reveals that the parking supply is insufficient.

(iv) The classification result of 53 return hotspots has found that the supply and demand patterns in hotspots are heterogeneous.

The findings in hotspots identification are consistent with previous studies, such as the spatial feature of hotspots [2, 50]. The findings of this study provide recommendations for the placement of BS electric fences in various areas. Some limitations, however, should be acknowledged. BS trip data only include weekday morning peak trips. As a result, the extraction of BS's temporal dynamics is unavailable. Parking management schemes only take into account illegal parking penalties and user incentives. The effect of parking management schemes on BS service should be studied further in future studies using multiscenario simulation. The real-world survey data should be used to calibrate the users' attitudes toward the electric fence and the corresponding penalty schemes.

Data Availability

BS trip data and electric fence geographic data were obtained from Digital China Innovation Contest, DCIC 2021 (https://data.xm.gov.cn/contest-series/digit-china-2021/#/3/competition_data).

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

Study conception and design were contributed by Yingkun Xie, Qing Yu, Xiaolei Liu, and Xinyuan Wang. Data collection, analysis and interpretation of results, and draft manuscript preparation were done by Yingkun Xie and Qing Yu. Review and editing were performed by Ying Hui and Qing Yu. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 51978475).

Supplementary Materials

Illustration of graphical abstract: (1) Origin/destination events in each grid extracted from the bicycle-sharing trip dataset are packed as the input of the DBSCAN cluster algorithm. (2) It reported 47 pick-up hotspots and 53 return hotspots on Xiamen Island. (3) Using electric fence data and point of interest (POI) data, an evaluation framework covering three aspects (supply and demand, unbalance, and land use) are developed to assess the characteristics of return hotspots. (4) Using the Gaussian-mixture clustering algorithm. 53 return hotspots are classified into three clusters, including hotspots where BS is in overload status, hotspots where BS service quality needs to be improved, and hotspots where BS is in stable status. (5) Parking management schemes and policy implications are proposed based on the classification result. (*Supplementary Materials*)

References

[1] H. Si, J.-g. Shi, G. Wu, J. Chen, and X. Zhao, "Mapping the bike sharing research published from 2010 to 2018: a scientometric review," *Journal of Cleaner Production*, vol. 213, pp. 415–427, 2019.

[2] Y. Zhang, D. Lin, and X. C. Liu, "Biking islands in cities: an analysis combining bike trajectory and percolation theory," *Journal of Transport Geography*, vol. 80, Article ID 102497, 2019.

[3] H. Luo, F. Zhao, W.-Q. Chen, and H. Cai, "Optimizing bike sharing systems from the life cycle greenhouse gas emissions perspective," *Transportation Research Part C: Emerging Technologies*, vol. 117, Article ID 102705, 2020.

[4] J. Wu, "Challenges and opportunities in algorithmic solutions for rebalancing in bike sharing systems," *Tsinghua Science and Technology*, vol. 25, no. 6, pp. 721–733, 2020.

[5] C. Médard de Chardon, G. Caruso, and I. Thomas, "Bike-share rebalancing strategies, patterns, and purpose," *Journal of Transport Geography*, vol. 55, pp. 22–39, 2016.

[6] Q. Yu, W. Li, D. Yang, and Y. Xie, "Policy zoning for efficient land utilization based on spatio-temporal integration between the bicycle-sharing service and the metro transit," *Sustainability*, vol. 13, no. 1, p. 141, 2020.

[7] Y. Yang, A. Heppenstall, A. Turner, and A. Comber, "Using graph structural information about flows to enhance short-term demand prediction in bike-sharing systems," *Computers, Environment and Urban Systems*, vol. 83, Article ID 101521, 2020.

[8] L. Chen, D. Zhang, L. Wang et al., "Dynamic cluster-based over-demand prediction in bike sharing systems," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 841–852, New York, NY, USA, September 2016.

[9] S. Feng, H. Chen, C. Du, J. Li, and N. Jing, "A hierarchical demand prediction method with station clustering for bike sharing system," in *Proceedings of the 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC)*, pp. 829–836, Guangzhou, China, June 2018.

[10] J. Huang, X. Wang, and H. Sun, "Central station based demand prediction in a bike sharing system," in *Proceedings of the 2019 20th IEEE International Conference on Mobile Data Management (MDM)*, pp. 346–348, Hong Kong, China, June 2019.

[11] C. Xu, J. Ji, and P. Liu, "The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 47–60, 2018.

[12] T. Xu, G. Han, X. Qi, J. Du, C. Lin, and L. Shu, "A hybrid machine learning model for demand prediction of edge-computing-based bike-sharing system using internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7345–7356, 2020.

[13] R. Qin, L. Kong, M. Guo, B. Yao, and M. Guizani, "Rebalance modern bike sharing system: spatio-temporal data prediction and path planning for multiple carriers," in *Proceedings of the 2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS)*, pp. 1–6, Singapore, December 2018.

[14] B. Wang and I. Kim, "Short-term prediction for bike-sharing service using machine learning," *Transportation Research Procedia*, vol. 34, pp. 171–178, 2018.

[15] R. Guo, Z. Jiang, J. Huang et al., "BikeNet: accurate bike demand prediction using graph neural networks for station rebalancing," in *Proceedings of the 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, pp. 686–693, Leicester, UK, August. 2019.

[16] M. Liu and X. Xu, "Dockless bike-sharing reallocation based on data analysis: solving complex problem with simple method," in *Proceedings of the 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC)*, pp. 445–450, Guangzhou, China, June 2018.

[17] P. Leclaire and F. Couffin, "Method for static rebalancing of a bike sharing system," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1561–1566, 2018.

[18] A. Pal and Y. Zhang, "Free-floating bike sharing: solving real-life large-scale static rebalancing problems," *Transportation Research Part C: Emerging Technologies*, vol. 80, pp. 92–116, 2017.

[19] J. Liu, L. Sun, W. Chen, and H. Xiong, "Rebalancing bike sharing systems," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1005–1014, New York, NY, USA, August 2016.

[20] J. Schuijbroek, R. C. Hampshire, and W.-J. van Hoeve, "Inventory rebalancing and vehicle routing in bike sharing systems," *European Journal of Operational Research*, vol. 257, no. 3, pp. 992–1004, 2017.

[21] L. Shi, Y. Zhang, W. Rui, and X. Yang, "Study on the bike-sharing inventory rebalancing and vehicle routing for bike-sharing system," *Transportation Research Procedia*, vol. 39, pp. 624–633, 2019.

[22] Y. Duan, J. Wu, and H. Zheng, "A greedy approach for vehicle routing when rebalancing bike sharing systems," in *Proceedings of the 2018 IEEE Global Communications Conference (globecom)*, Abu Dhabi, United Arab Emirate, December 2018.

[23] X. Zhang, H. Yang, R. Zheng, Z. Jin, and B. Zhou, "A dynamic shared bikes rebalancing method based on demand prediction," in *Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 238–244, Auckland, New Zealand, October 2019.

[24] D. Zhang, C. Yu, J. Desai, H. Y. K. Lau, and S. Srivathsan, "A time-space network flow approach to dynamic repositioning in bicycle sharing systems," *Transportation Research Part B: Methodological*, vol. 103, pp. 188–207, 2017.

[25] Z. Tian, J. Zhou, W. Y. Szeto, L. Tian, and W. Zhang, "The rebalancing of bike-sharing system under flow-type task window," *Transportation Research Part C: Emerging Technologies*, vol. 112, pp. 1–27, 2020.

[26] J. Hu, Z. Yang, Y. Shu, P. Cheng, and J. Chen, "Data-driven utilization-aware trip advisor for bike-sharing systems," in *Proceedings of the 2017 IEEE International Conference on Data Mining (ICDM)*, V. Raghavan, S. Aluru, G. Karypis, L. Miele, and X. Wu, Eds., pp. 167–176pp. 167–, New Orleans, LA, USA, November 2017.

[27] Y. Duan and J. Wu, "Optimizing rebalance scheme for dockless bike sharing systems with adaptive user incentive," in *Proceedings of the 2019 20th IEEE International Conference on Mobile Data Management (MDM),*, pp. 176–181, Hong Kong, China, June. 2019.

[28] L. Pan, Q. Cai, Z. Fang, P. Tang, and L. Huang, "A deep reinforcement learning framework for rebalancing dockless bike sharing systems," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 1393–1400, 2019.

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

[30] Y. Zhang, D. Lin, and Z. Mi, "Electric fence planning for dockless bike-sharing services," *Journal of Cleaner Production*, vol. 206, pp. 383–393, 2019.

[31] S. Xie, Y. Li, Q. Xu, F. Fang, and L. Li, "Image-based parking place identification for regulating shared bicycle parking," in *Proceedings of the 2018 15th International Conference On Control, Automation, Robotics And Vision (ICARCV)*, pp. 1709–1714, Singapore, November. 2018.

[32] X. Yuan, J. Senlin, Z. Chongxia, L. Wei, H. Yan, and Y. Hejun, "High accuracy virtual electronic fence management technique based on GNSS," in *Proceedings of the 2017 13th IEEE International Conference on Electronic Measurement Instruments (ICEMI)*, pp. 79–83, Yangzhou, China, October. 2017.

[33] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "DBSCAN revisited, revisited," *ACM Transactions on Database Systems*, vol. 42, no. 3, p. 19, 2017.

[34] C. S. Shui and W. Y. Szeto, "A review of bicycle-sharing service planning problems," *Transportation Research Part C: Emerging Technologies*, vol. 117, Article ID 102648, 2020.

[35] D. Çelebi, A. Yörüsün, and H. Işık, "Bicycle sharing system design with capacity allocations," *Transportation Research Part B: Methodological*, vol. 114, pp. 86–98, 2018.

[36] S. Chang, R. Song, S. He, and G. Qiu, "Innovative bike-sharing in China: solving faulty bike-sharing recycling problem," *Journal of Advanced Transportation*, vol. 2018, pp. 1–10, 2018.

[37] Z. Chen, D. van Lierop, and D. Ettema, "Dockless bike-sharing systems: what are the implications?" *Transport Reviews*, vol. 40, no. 3, pp. 333–353, 2020.

[38] J.-g. Shi, H. Si, G. Wu, Y. Su, and J. Lan, "Critical factors to achieve dockless bike-sharing sustainability in China: a stakeholder-oriented network perspective," *Sustainability*, vol. 10, no. 6, p. 2090, 2018.

[39] E. Eren and V. E. Uz, "A review on bike-sharing: the factors affecting bike-sharing demand," *Sustainable Cities and Society*, vol. 54, Article ID 101882, 2020.

[40] A. Li, P. Zhao, Y. Huang, K. Gao, and K. W. Axhausen, "An empirical analysis of dockless bike-sharing utilization and its explanatory factors: case study from Shanghai, China," *Journal of Transport Geography*, vol. 88, Article ID 102828, 2020.

[41] Y. Du, F. Deng, and F. Liao, "A model framework for discovering the spatio-temporal usage patterns of public free-floating bike-sharing system," *Transportation Research Part C: Emerging Technologies*, vol. 103, pp. 39–55, 2019.

[42] Y. Shen, X. Zhang, and J. Zhao, "Understanding the usage of dockless bike sharing in Singapore," *International Journal of Sustainable Transportation*, vol. 12, no. 9, pp. 686–700, 2018.

[43] J. Bao, C. Xu, P. Liu, and W. Wang, "Exploring bikesharing travel patterns and trip purposes using smart card data and online point of interests," *Networks and Spatial Economics*, vol. 17, no. 4, pp. 1231–1253, 2017.

[44] Y. Xing, K. Wang, and J. J. Lu, "Exploring travel patterns and trip purposes of dockless bike-sharing by analyzing massive bike-sharing data in Shanghai, China," *Journal of Transport Geography*, vol. 87, Article ID 102787, 2020.

[45] H. Alizadeh, A. Khoshrou, and A. Zuquete, "Traffic classification and verification using unsupervised learning of Gaussian Mixture Models," *2015 IEEE International Workshop on Measurements & Networking (M&N)*, in *Proceedings of the 2015 IEEE International Workshop on Measurements Networking (M N)*, pp. 1–6, Coimbra, Portugal, October 2015.

[46] H. Alizadeh, H. Vranken, A. Zuquete, and A. Miri, "Timely classification and verification of network traffic using Gaussian mixture models," *IEEE Access*, vol. 8, Article ID 91287, 2020.

[47] X. Liu, L. Pan, and X. Sun, "Real-time traffic status classification based on Gaussian mixture model," in *Proceedings of the 2016 IEEE First International Conference on Data Science in Cyberspace (DSC)*, pp. 573–578, Changsha, China, June. 2016.

[48] Z. Zhou, *Machine Learning*, Tsinghua University Press, Beijing, China, 2016.

[49] Q. Yu and J. Yuan, "TransBigData: a Python package for transportation spatio-temporal big data processing, analysis and visualization," *Journal of Open Source Software*, vol. 7, no. 71, p. 4021, 2022.

[50] Z. Liu, Y. Shen, and Y. Zhu, "Where will dockless shared bikes be stacked?" in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 566–575, New York, NY, USA, July 2018.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 1116221, 18 pages
https://doi.org/10.1155/2022/1116221

Research Article

Parking Permit Scheme for Morning Commute considering Parking Search

Duo Xu and Huijun Sun

Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Duo Xu; 516472407@qq.com

Received 16 December 2021; Revised 23 January 2022; Accepted 26 January 2022; Published 22 March 2022

Academic Editor: Wenxiang Li

Improving parking efficiency is essential to promoting the reform in urban transportation. But the large amount of deadweight costs caused by the parking is often underestimated because it is difficult to measure. Based on the existing investigations from the small fraction of cruising vehicles, this paper explores the influencing factors of the parking issue and describes it by the user equilibrium model. Then, two types of permit management schemes were proposed, lot-based and spot-based. By analyzing their performance in reducing system cost, three conclusions were drawn. Firstly, parking search leads to traveler's schedule and location adjustments, raises the trip cost, reduces the parking lot occupancy, and makes the parking issues "invisible." Secondly, permit scheme levels up managers' control, and it performs well in reducing deadweight loss, but only by eliminating the search cost, the deadweight loss can be fundamentally reduced. Thirdly, reducing parking search needs information guidance; with the rapid growth of urban parking demand, managers should make a transition to the permit scheme with parking information.

1. Introduction

Electrification and automated and shared vehicles will likely revolutionize urban transport and further promote travel efficiency, energy conservation, and emission reduction. Until then, parking issues will remain an inherent and rather unpleasant attribute of car travel in metropolitan regions [1].

Studies show that although green travel has been widely advocated, private cars are still the main way of urban travel. In the city of Beijing, private cars account for 30% of the travel demand, and each car travels more than 3 times with about 50 km distance a day on average [2]. However, as the endpoint of car travel, finding a parking spot often constitutes an appreciable fraction of the total travel time and contributes to traffic congestion and negative externalities [3–5]. According to the survey data, the time cost that a vehicle looks for a vacant spot accounts for 30–50% of the total travel time, nearly 70% of the direct travel cost, and 30% of traffic jams [6–9]. In Chicago, cruising for parking produces a total of 63 million miles of distance and 48,000 tons of carbon dioxide annually [10].

Over the years, charging (includes static, multi-phases, or dynamically varied with time and space) plays an important role in regulating demand, and it is widely adopted in practice [11–13]. However, it has limitations in dealing with parking searches. First, it cannot regulate demand accurately to meet the full utilization of parking lots with no vehicles cruising. Second, the charge itself does not contain information to help drivers avoid the search. In practice, managers often raise charges to keep a small number of parking spots vacant (5%–15%) and consider it an effective compromise [14, 15].

With the development of information and communication technology, quantity control approaches based on permits are introduced into the field of traffic demand management in recent decades. They originated from the regulation approach in environmental externalities [16, 17] and have developed to several forms, e.g., the license-plate rationing, the appointment-based schemes, the cap-and-auction scheme, the cap-and-trade scheme, etc. Permit scheme applied in the parking field is meaningful; first, it

performs well in matching supply and demand, avoiding leap growth in costs when demand overflows; second, it is compatible with identity authentication, information release, and flexible price adjustment and is suitable for the integration of private and incomplete open parking resources, e.g., residential parking spots. They provide a potential opportunity to realize smart, shared parking.

In the empirical study of parking, the existing knowledge is strongly biased towards the reaction of drivers to parking prices [18–22], while the other critical factors, e.g., the occupancy rate, delay, and distance to destination, remain obscure [1]. In recent years, various methods were used to observe parking searching and cruising, e.g., follow vehicles, park-and-visit tests, in-car video, and GPS tracking [4, 23–31], but they did not reach completely consistent conclusions in studying travelers' search behavior and its impact on the system.

In the theoretical research field, the literature on parking permits is in two categories. One is the parking problem itself and to study the solution of reducing travelers' delay time or queuing time; the other is to make full use of multiple trip modes or road networks and to study how to reasonably allocate travel demand [32–39]. However, the parking search has not been fully considered. Search is more complicated than queuing. First, queuing cost is linear and definite, and travelers get parking services after waiting for a certain period. The parking lot supply is also fully utilized. But the search cost is nonlinear, and it increases with the reduction of vacant parking spots [6, 14, 40–42]. When it is higher than travelers' expectations (equilibrium point of the system), parking facilities may not be fully used, and the excess demand will expand the parking range. Second, queuing is affected by travelers' departure/arrival time, which can be reduced by dispersing travelers' departure/arrival. But searching is affected by parking lot utilization (essentially the arrival order). The time is longer for the late arrivals, which cannot be solved by decentralizing the departure time (as the current permit does).

Therefore, in order to explore solutions to the parking search problem, this paper summarizes the results of existing empirical research. First, parking search and cruise are common in the parking process, and in most cases, they take 3 to 5 minutes. Second, by comparing the data of city center and suburban, the scarcity of parking facilities affects the cruising time. Third, drivers who are more familiar with the road environment spent less time in cruising, indicating that the information is helpful for parking. Fourth, some recent studies have found that the search time and distance become shorter, and the proportion of travelers who stop before the destination increases. It is inferred that travelers can perceive the use of parking facilities at various locations more accurately and make trade-offs in advance. Accordingly, the parking problem is summarized into four aspects, the parking search cost, associated trip delay caused by departure time adjustment, the extra walking cost caused by parking location adjustment, and underutilization of the parking lot.

On this basis, this paper establishes a user equilibrium model to study the relationship between parking lot utilization and parking choices (parking position and departure time selection), then proposes permit schemes based on quantity control, and explores their feasibility in solving the parking search problem. Correspondingly, three schemes are introduced.

No Permit Scheme. Managers only charge and do not use permits (static charge is adopted). Travelers compete from the dimension of temporal and spatial. This scheme is set as a benchmark.

Parking Lot Permit Scheme. Manager issues parking permits based on independent parking lots, and travelers with permits can park anywhere in the corresponding parking lot. It is theoretically similar to the permit proposed by Zhang et al. [32] that does not distinguish between parking spaces. In practice, it is similar to the offline permit scheme (regularly in a parking facility), but the cycle is short (e.g., within a day).

Parking Spot Permit Scheme. Manager issues parking permits based on specific parking spots, and travelers with parking permits need to park in a designated spot. It is theoretically similar to the permit applied in the field of shared parking (the opening time of the supplier needs to match the parking time of the demander). In practice, it is applied in the management of private spots in a residential area or some spots reserved and bound to specific vehicles (parking spot marked with license-plate number in the office building) (Table 1).

The remainder of this paper is organized as follows. Section 2 presents the assumptions and analyzes the components of the trip costs. Section 3 deduces the user equilibrium in a no permit scheme. Two parking lot permit schemes are elaborated in Section 4. In Section 5, a numerical experiment is provided. Section 6 concludes the paper.

2. Assumptions

This paper concentrates on parking issues in a business district (not unique) during the morning peak, mainly considering the following.

Firstly, commuters can accumulate experience from daily trips, and they try to achieve user equilibrium. Secondly, commuters arrive at the parking lot in a concentrated time, which leads to a rapid and continuous change in parking lot occupancy rate from low to high in a short period. The impact of search problems is obvious. Thirdly, commuting trips have long parking times in general, and the problem can be simplified to a one-time parking problem (commuters do not leave away in the study period).

To facilitate the presentation of the essential ideas without loss of generality, we make the following basic assumptions.

A1: Network. As shown in Figure 1, in a linear network structure, the business district (office) is at O, and the residential area (home) is at the other end H of the city. A

Table 1: Description of three schemes.

Scheme name	Abbreviation	Description
No permit	Scheme n	The manager sets a single charge rate on the basis of first come first serve.
Parking lot permit	Scheme p	The manager issues permit for different parking lots. Commuters park in a lot by the permit of that lot.
Parking spot permit	Scheme r	The manager issues permit for every spot. Commuters park in a spot by the permit of that spot.

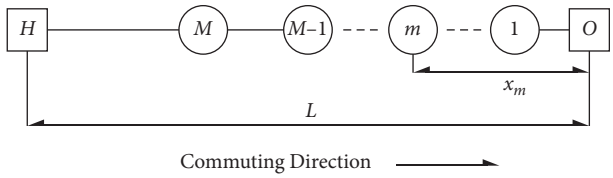


Figure 1: Linear urban structure and parking facilities.

highway with length L connects H and O (the capacity is large enough, and there is no congestion).

Assumption 1 is the simplification of a certain OD pair in a city. (1) Vehicle drivers prefer a short walking distance (generally no more than 0.3 km–0.6 km or within one block) after parking [43]. (2) The commuter number to a certain district is far less than the total commuter number in the city. Vehicle velocity is mainly affected by the external traffic; we set the capacity of the road network as large enough, and vehicle velocity is constant for simplicity.

A2: Commuters. There are N commuters who travel from H to O once a day, they choose the departure time and parking location, and their working start time is t^*. Their trip is composed of three parts: (1) drive to the parking lot; (2) search for a vacant spot in the lot; (3) walk to O. Their time value is α (during a trip) and β (arriving early or late), respectively.

Assumption 2 studies morning peak commuting. (1) Trips are concentrated, and parking competition is obvious. (2) The relatively long parking time can avoid the impact of leaving halfway.

A3: Parking Lot. There are M off-street parking lots distributed in a straight line within an acceptable walking distance approximately to O. The capacity and location of the parking lot m $(m = 1, 2, \ldots, M)$ are k_m and x_m, respectively. We set the walking distance inside the parking lot as 0. To satisfy both the scarcity of parking spots near O and the satisfiability of parking demand, we set $k_m < N < \sum_M k_m$.

Assumption 3 describes the current parking situation in urban areas. (1) Because the parking supply close to the destination is insufficient, some commuters may not be able to park at the destination but can park within a certain distance (e.g., in Beijing, there are more than ten parking lots in a square kilometer in the central area of the city). (2) Parking lot usually has more than one entrance and exit, and its internal spots are nonlinearly distributed. Therefore, we ignore the spots' differences for simplicity.

Further, the rest of the notations throughout the paper are listed in Table 2, and we will directly use them in the following sections.

Correspondingly, the trip cost includes travel cost (car driving and walking), search cost, delay cost, and parking charge.

2.1. Travel Cost. When commuter parks at x_m, his/her drive time and walk time are $(L - x_m)/v_c$ and x_m/v_w, respectively $(v_c > v_w)$. Then, the travel cost is

$$C^t(x_m) = \alpha\left[\frac{(L - x_m)}{v_c} + \frac{x_m}{v_w}\right]. \tag{1}$$

2.2. Search Cost. The number of occupied parking spots is no more than the capacity, and the formula is $n(x_m, t) \le k_m$. If there is no parking guidance information, commuters will search for vacant parking spots one by one in the parking lot. The expected number of spots searched before finding a vacant one is $k_m/[k_m - n(x_m, t)]$ [40]. The expected search time satisfies $t_s(x_m, t) = \lambda k_m/[k_m - n(x_m, t)]$ $(\lambda \ll C_t(0))$. If commuters receive parking information in advance, they will go directly to the designated spot. Considering that the car speed will not be reduced with the spot information, but the search distance is much longer, we set $\pi > \lambda$. The expected search cost in the scheme j is

$$C_j^s(x_m, t) = \begin{cases} \dfrac{\alpha\lambda k_m}{k_m - n(x_m, t)}, & \text{scheme } j \text{ without information,} \\ \alpha\pi, & \text{scheme } j \text{ with information.} \end{cases} \tag{2}$$

2.3. Delay Cost. The delay cost is the product of the delay time and the time value β. The delay cost of commuters who park at lot m and time t is

$$C^d(x_m, t) = \beta\left|t^* - t - t_s(x_m, t)\frac{x_m}{v_w}\right|. \tag{3}$$

2.4. Parking Charge. We set all parking lots to have nonnegative parking charges (the charging price in the no permit scheme and the permit price in the parking lot and parking spot scheme), and there is $p_j(x_m, t) \ge 0$.

In conclusion, the total trip cost of the commuter park at lot m and time t in the scheme j is

$$C_j(x_m, t) = C^t(x_m) + C_j^s(x_m, t) + C^d(x_m, t) + p_j(x_m, t). \tag{4}$$

Table 2: Nomenclature.

Notation	Interpretation
x_m	Location of the parking lot m
x_j^*	Furthest parking location in scheme j
t_{j,x_m}^s	The first arrival time to the parking lot m in the scheme j
t_{j,x_m}^e	The last arrival time to the parking lot m in the scheme j
t_{j,x_m}^i	The $i-th$ arrival time to the parking lot m in the scheme j
$t_s(x_m,t)$	The expected search time at the parking lot m and time t
$n(x_m,t)$	Occupied parking spot number at parking lot m and time t
λ	Unit search time without information guidance
π	Average search time with spot information
v_c	Car velocity
v_w	Walking velocity
C	Trip cost
C^t	Travel cost
C_j^s	The expected search cost in the scheme j
C^d	Delay cost
$p_j(x_m,t)$	Parking charge at the parking lot m and time t in the scheme j
C^*	Trip cost under user equilibrium
τ	The minimum distinguishable time interval between two adjacent arrivals

Among the costs in (4), the travel cost is related to the parking location x_m, the delay cost is related to the parking time t and the parking location, and the search cost and charging price are also related to the scheme j.

3. User Equilibrium in the No Permit Scheme

At present, the most common form of urban parking management is to set a certain amount of charging fees (without spot information) in advance. They are generally applicable to a wide temporal and spatial range (e.g., day or night, suburban, or city center). When we study a small range (e.g., a traffic zone during morning peak), charges can be regarded as constant.

Since the total parking spot number is greater than the total commuter number, the trip time (cost) has a boundary. Rational commuters will weigh various factors and try to minimize trip costs by individually adjusting their strategy (e.g., schedule and parking location). If no one can reduce his/her trip cost by changing strategy individually, the system reaches user equilibrium.

3.1. Parking Lot Choice of the Commuter. In the scheme n, parking charges do not vary with time and location. We set parking charges that are all equal to zero for simplicity. Commuters' choices involve two dimensions, spatial (where to park) and temporal (when to park). In this section, we focus on the spatial dimension: (1) derive the user equilibrium between parking lots; (2) calculate the parking number in each parking lot and get the spatial parking distribution.

According to (2), when the parking number is close to the lot capacity, the commuter's search cost becomes so large that it will exceed the equilibrium cost. From this point, we get Remark 1.

Remark 1. If commuters are rational and punctual, there will be no traveler to arrive at the office after the time t^* under user equilibrium.

Proof 1. If a commuter arrives at O later than t^*, the positive delay cost can be replaced by a higher search cost. It means that the parking lot can attract more commuters who planned to park elsewhere and reduce their travel cost, which contradicts the assumption.

This means that the parking lot can attract more commuters who originally planned to park in other locations.

According to Remark 1, the final parking number of the parking lot m can be represented by $n(x_m,t^*)$.

Given the time point when the system reaches the final state, we further verify the existence and uniqueness of the user equilibrium of the scheme n. We construct the following model based on the Beckman programming model (Beckman programming model (proposed in 1956) which satisfies Wardrop's user equilibrium criteria):

$$\begin{aligned} &\min Z(N) = \sum_M \int_0^{n(x_m,t^*)} C_n(x_m,t^*,w)\mathrm{d}w, \\ &s.t. \sum_M n(x_m,t^*) = N, \\ &\quad n(x_m,t^*) \geq 0. \end{aligned} \tag{5}$$

When commuters are traveling without congestion, the parking lots they choose can be seen as routes without overlapping arcs in the road network. The final number of parking lots $n(x_m,t^*)$ is equivalent to the route flow, which satisfies the conservation condition.

The Lagrangian function of (5) is

$$La = Z(N) + \mu\left[N - \sum_M n(x_m,t^*)\right], \tag{6}$$

where μ is the Lagrangian multiplier. Omitting proof and derivation, the Kuhn–Tucker condition of (5) is

$$n(x_m, t^*) \cdot [C_n(x_m, t^*) - \mu] = 0, \quad C_n(x_m, t^*) - \mu \geq 0. \tag{7}$$

Therefore, the minimum trip cost C^* exists and equals the Lagrangian multiplier μ. From (11), if $n(x_m, t^*) > 0$, $C_n(x_m, t^*) = C_n^*$, and if $n(x_m, t^*) = 0$, $C_n(x_m, t^*) \geq C_n^*$.

According to the cost settings in (2) and (4), trip costs are related to the parking number in the objective lot, so there is $\partial C_n(x_m, t^*)/\partial n(x_\chi, t^*) = 0$ (χ is also the serial number of the parking lot, and it is independent of m), and

$$\frac{\partial^2 Z}{\partial n(x_m, t^*) \partial n(x_\chi, t^*)} = \begin{cases} \dfrac{dC_n(x_m, t^*)}{dn(x_\chi, t^*)}, & x_m = x_\chi, \\ 0, & x_m \neq x_\chi. \end{cases} \tag{8}$$

According to equation (2), $dC_n(x_m, t^*)/dn(x_m, t^*) > 0$. The Hessian matrix of the objective function (5) is a positive definite matrix. It has a unique solution C_n^*.

The UE programming represented by (5) has a convex objective function (the cost of each parking location is a strictly increasing function of parking number), but the parking number at each location is not necessarily convex. Therefore, the programming problem is nonlinear, and there may be multiple solutions. Fortunately, different from the general traffic allocation problems, parking spot users instinctively prefer parking close to the destination (the destination is attractive for traffic flow). Based on this, we continue to discuss whether there is a unique parking distribution.

Since the total parking spot number is greater than the commuter's number, the parking distribution has a boundary in the scheme n. We denote the farthest parking location as x_n^*; then, for any $x_m > x_n^*$, there is $n(x_m, t^*) = 0$.

From (4) and (7), we get Remark 2 and Remark 3. □

Remark 2. Commuter who parks at the farthest parking location x_n^* does not bear the delay cost, and his/her search cost is $\alpha\lambda$.

Proof 2. If the sum of search cost and delay cost is greater than $\alpha\lambda$ at the location x_n^*, a commuter parked in a further location $x_n^* + \Delta x$ ($\Delta x \longrightarrow 0$) has a lower trip cost, which contradicts the assumption. □

Remark 3. In the scheme n, when the system reaches user equilibrium, parking distribution (or the utilization of each parking lot) is unique.

Proof 3. From Remark 2 and equation (4), the equilibrium cost C_n^* can be expressed as

$$C_n^* = \alpha\left[\frac{(L - x_n^*)}{v_c} + \lambda + \frac{x_n^*}{v_w}\right]. \tag{9}$$

From Remark 2 and (9), for the parking lot m with positive $n(x_m, t^*)$, there is $C^t(x_m) + C_n^s(x_m, t^*) = C_n^*$.

The search cost satisfies $C_n^s(x_m, t^*) = C_n^s(x_m, t^* - x_m/v_w)$ because all the commuters finish their parking at the time $t^* - x_m/v_w$ and there will be no commuters parked between $t^* - x_m/v_w$ and t^*. The final parking number $n(x_m, t^*)$ at each location satisfies $n(x_m, t^*) = n(x_m, t^* - x_m/v_w)$. Combined with (9), there is

$$n(x_m, t^*) = k_m - \frac{\lambda k_m}{(x_n^* - x_m)/v_w - (x_n^* - x_m)/v_c + \lambda}. \tag{10}$$

Combining conservation condition (5), there is

$$\sum_M \left[k_m - \frac{\lambda k_m}{(x_n^* - x_m)/v_w - (x_n^* - x_m)/v_c + \lambda}\right] = N. \tag{11}$$

Theoretically, we assume that the number of parking lots M is large enough and the distance between them is close enough; then, we can rewrite (11) as the integral form.

$$\int_{x_n^*} \left[k_x - \frac{\lambda k_x}{(x_n^* - x)/v_w - (x_n^* - x)/v_c + \lambda}\right] dx = N. \tag{12}$$

We denote k_x as a constant instead of $k(x)$ because parking capacity is independent of location x. Take the derivative of the LHS of (12) with x, and we get $\lambda k_x (1/v_c - 1/v_w)/[(x_n^* - x)/v_w - (x_n^* - x)/v_c + \lambda]^2 < 0$. The LHS of (12) is a monotonically decreasing function, so there is a unique solution. Similarly, the discrete function (11) has a unique solution. □

3.2. Parking Time Choice of the Commuter. In Section 3.1, we discussed the commuters' choice in spatial dimension as well as the final utilization of each parking lot at the moment t^*. In this section, we are concerned about the parking time choices within the parking lot (temporal dimension). We take the last parker in each parking lot as a benchmark and extend the research to the time dimension to study commuters' arriving time (to the parking lot).

Under user equilibrium, the sum of search and delay costs is equal for commuters arriving at different times in the same parking lot. Referring to the cost of the last arrival, there is $C_n^s(x_m, t) + C^d(x_m, t) = C_n^s(x_m, t^*)$ (excluding the same travel cost). Then, we get

$$\frac{\alpha\lambda k_m}{k_m - n(x_m, t)} + \beta \cdot \left[t^* - t - \frac{x_m}{v_w} - \frac{\lambda k_m}{k_m - n(x_m, t)}\right] = \frac{\alpha\lambda k_m}{k_m - n(x_m, t^*)}. \tag{13}$$

For the first arriver, his/her search time is λ, and we calculate his/her arrival time at the lot m:

$$t_n^{s,x_m} = t^* - \frac{\alpha}{\beta} \cdot \frac{\lambda[n(x_m, t^*) - 1]}{k_m - n(x_m, t^*)} - \frac{x_m}{v_w} - \lambda. \tag{14}$$

For the last arriver, his/her search time is $\lambda k_m/[k_m - n(x_m, t^*) + 1]$, and we calculate his/her arrival time to the parking lot m:

$$t_n^{e,x_m} = t^* - \frac{x_m}{v_w} - \frac{\lambda k_m}{k_m - n(x_m, t^*) + 1}. \quad (15)$$

All the other commuters' arrival time t_n^{i,x_m} ($i \in [1, n(x_m, t^*)]$) is between t_n^{s,x_m} and t_n^{e,x_m}.

3.3. Graphic Representation of Equilibrium. In Sections 3.1 and 3.2, we discussed the equilibrium of commuters' choices between parking lots and at different times, respectively. In this section, we draw figures to describe the equilibrium of multiple dimensions.

According to (2), we present the curve of search cost in Figure 2. The search cost keeps at a low level when the occupancy rate of parking spots is low (e.g., under 80%) and then increases significantly when the occupancy rate approaches 100% (it is assumed that commuter spends one unit time searching for one spot; when the occupancy rate reaches 0.8, the expected search time will increase to 5 times, and when it reaches 0.95, the cost of parking search will increase to 20 times).

3.3.1. Occupancy-Time Dimension. Figure 3 shows the relationship between a cumulative arrival rate and delay time. It takes four ultimate occupancy rates (0.96, 0.9, 0.8, and 0.5) as an example (blue, orange, yellow, and purple curves, respectively). It shows the parking lots with a high occupancy rate; most of the commuters will bear more delay time. When the system reaches equilibrium, the sum of search cost and delay cost of travelers in the same parking lot is equal; the high delay time compensates for the high search time of late arrivals. Parking lots with a higher occupancy rate (the curve deviating from the origin) are the locations close to the destination O.

3.3.2. Time-Distance Dimension. To fully demonstrate the influence of distance factors on the occupancy rate of parking lots, we transform the discrete distribution model of the parking lot into the continuous distribution model and make a theoretical analysis. The distances between any two adjacent lots satisfy $x_{m+1} - x_m \longrightarrow 0$.

According to Remark 2, the trip costs of the last parkers only include travel cost and search cost, so they all depart from H at $t^* - C^*/\alpha$. We further derive the latest time of (1) departing from residence, (2) arriving at parking lots, and (3) leaving from parking lots, which are, respectively, plotted in Figure 4 by orange dot solid line (LT (depart)), orange solid line (LT (arrive lot)), and the red dotted line (LT (leave lot)). For the earliest arrivals at each location, the cost only includes the travel cost and delay cost, and the sum is fixed. According to equations (14) and (15), we derive the earliest time of (1) departing from residence H and (2) arriving at parking spots, which are, respectively, plotted by blue dot solid line (ET (depart)) and blue solid line (ET (arrive lot)).

As shown in Figure 4, the part between the red dotted line and the black horizontal solid line t^* is commuters' walk time, the part between the red dotted line and the solid orange line is the maximum search time, and the part between the red dotted line and the solid blue line is the sum of the maximum delay time (includes a unit search time). At

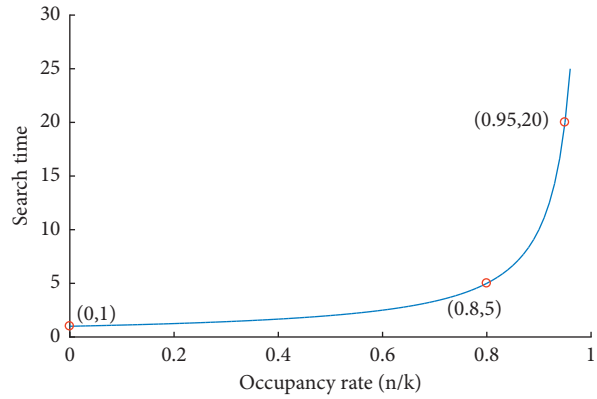


Figure 2: Search time varying with the occupancy rate.

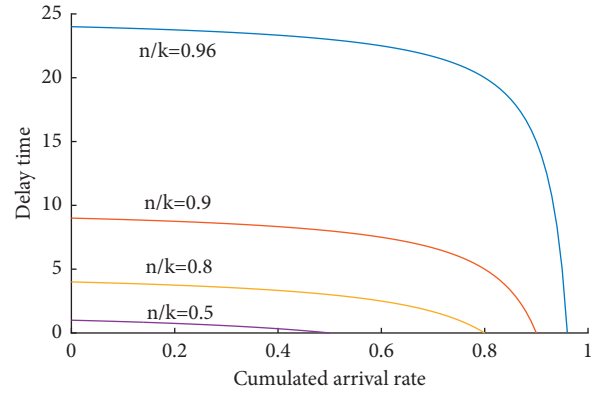


Figure 3: Delay time of locations with different n/k.

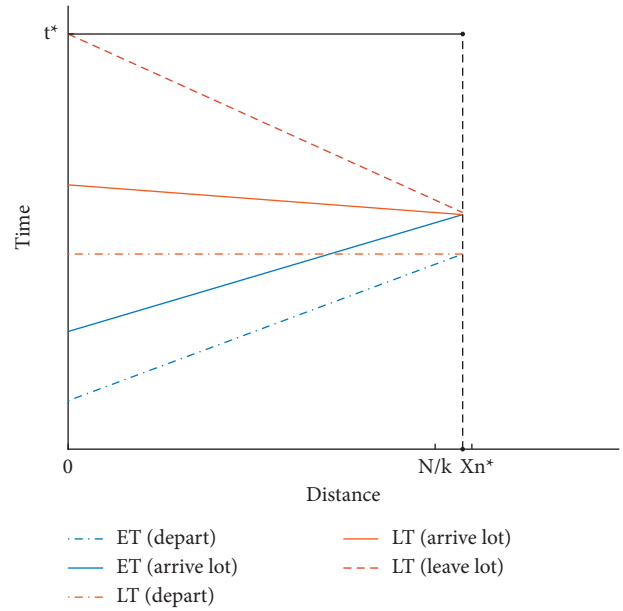


Figure 4: Trip schedule in scheme n.

the farthest location, the first arriver is also the last, so the blue and orange lines intersect x_n^*.

3.3.3. Occupancy-Distance Dimension. According to Remarks 1 and 2, the sum of travel cost and search cost for all the last commuters at each location is the same. We can calculate the final parking number at each location $n(x_m, t^*)$ from (10). Figure 5 shows the relationship between the occupancy rate of each parking lot with distance at the time t^*. The gentle change near the office shows that when the value of occupancy rate $n(x_m, t^*)/k_m$ is large, the linear variation in the travel cost can be offset by a small change in it (the small change produces a relatively large search cost variation). Similarly, the fast descending part near the farthest parking location x_n^* shows that when the value of occupancy rate is small, the linear variation in the driving cost needs to be offset by a large change in it.

Figures 2–5 indicate that the deadweight loss cost associated with parking for commuters in the scheme n is mainly reflected in search and delay. When the occupancy rate of the parking lot is high, the search time of commuters is longer. Under user equilibrium, commuters arriving earlier in the same parking lot also bear longer delays to avoid searching. In Figure 5, most parking lots within acceptable walking distance are with a high occupancy rate; only a small number of commuters choose farther parking lots to avoid the search cost and the associated delay cost. The parking problem is a combination of high search, delay, and travel costs, but its appearance may be as shown in Figure 5, where parking near O is not fully utilized. This can even give managers the illusion that parking spots are not scarce.

3.3.4. Time-Distance-Occupancy Dimension. By integrating time, distance, and occupancy, we draw a three-dimensional figure to depict temporal and spatial parking behavior (distribution), as shown in Figure 6. Its projection on the "time-distance" coordinate plane in Figure 4, that on the "occupancy-distance" coordinate plane in Figure 5, and that on the "occupancy-time" coordinate plane represent the cumulative arrival of each location over time (the upper yellow part of the occupancy rate no longer extends to time t^*, indicating that it has reached the final value at that time).

Figure 6 shows the following. (1) Search cost leads to the low efficiency of spot utilization while the occupancy rate of all locations is below 1, and it decreases as the distance increases. (2) The cumulative arrival surface reflects most commuters arriving early in nearly all parking locations, and the arrival rate tends to decrease with time.

Figure 6 is consistent with the parking survey results mentioned in Section 1. It provides a visual representation of parking utilization (individual parking utilization and parking range) and delays.

Therefore, in the most commonly adopted static charge scheme (single charge within a certain space), there are not only an explicit parking search but also some implicit issues,

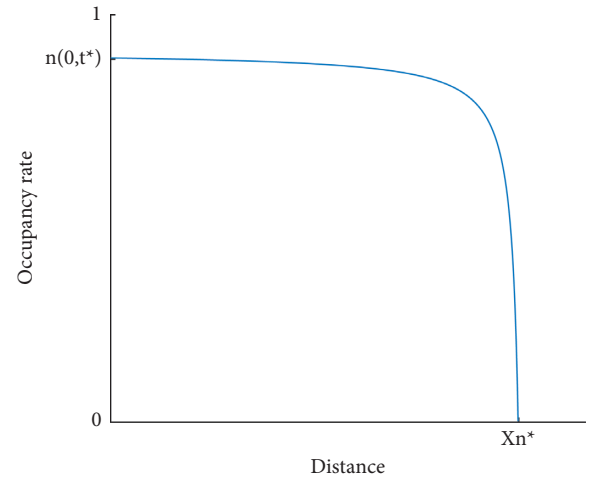


Figure 5: The final occupancy rate of all locations.

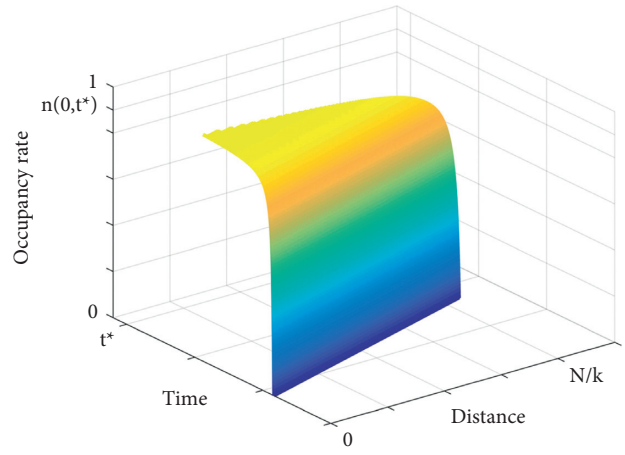


Figure 6: Arrivals at parking lots in scheme n.

e.g., a wide-range and large-scale delay cost occurs when parking spots are short in supply.

4. Improved Schemes of Parking Permit

Section 3 depicts the series of parking problems, including parking search, schedule delay, travel cost increase, and insufficient use of parking lot. In this section, two parking permit schemes are proposed, the parking lot permit scheme and the parking spot permit scheme. Both of them have basic practical application prototypes. The parking lot permit scheme is similar to a membership system, where travelers can purchase parking services from a parking lot for some time (weeks, months, and so on). It is usually for business uses, e.g., in institutions, companies, communities, etc. The parking spot permit scheme is similar to a reservation system, in which travelers reserve a designated parking spot where they can park directly. It is often used in a ticket booking system (containing the time and location

information), e.g., in shared parking management or the airplane ticket or train ticket booking.

Aiming at reducing the total cost of the system and the equilibrium cost of commuters, this section explores the effect and limitations of permits in regulating each sort of cost. Combined with the actual parking process, we design a parking spot allocation mechanism (the number of open spots varies over time and space) and set prices to achieve the optimal parking distribution from the system optimal perspective.

4.1. Parking Lot Choice of the Commuter. To design the optimal pattern of permit issuance, an analysis of the components of travel costs is necessary so that we can find out which parts can be improved in the scheme p. According to Section 2, parking costs include four parts; among them, travel cost and search cost, respectively, depend on the location and occupancy rate (or arrival sequence) of parking spots; we define them as "invariable costs" (if the location and parked number are fixed, the costs are unique); delay cost and parking charge are "variable costs" which can be adjusted by the manager or commuters themselves.

Parking permits have two functions: one is to ensure parking for commuters and the other is to adjust commuters' parking behavior by setting the validity period and charging price. Referring to the method in Section 3, we still start by studying the final equilibrium state of the system and analyze the cost composition of the last arrival of each parking lot.

To begin with, we derive the equilibrium of the scheme p. As discussed in Section 3, we start with finding the farthest parking location and derive the equilibrium of the system.

According to Remark 2, all the $n(x_m, t^*) - th$ commuters at each lot in the scheme n only experience travel costs and search costs, both of which are invariable costs and cannot be internalized by charging. Therefore, we draw a further remark.

Remark 4. If the invariable costs of the system remain constant, the parking lot permit cannot reduce the equilibrium cost.

Proof 4. According to equations (4) and (9), the cost of the $n(x_m, t^*) - th$ commuter at each lot is

$$C_p\left(x_m, t^{e,x_m}\right) = \alpha\left[\frac{\lambda k_m}{k_m - n\left(x_m, t^*\right)} + \frac{L - x_m}{v_c} + \frac{x_m}{v_w}\right] + p_p\left(x_m, t^{e,x_m}\right). \tag{16}$$

The equilibrium cost is equal to the sum of travel cost, search cost, and parking charge; when the charge $p_p(x_m, t^{e,x_m}) > 0$, the cost will exceed the equilibrium cost C_n^*. In addition, if the manager alleviates searching by reducing the permit issuance, then some commuters will move to other parking lots, $C_p^* > C_n^*$, which will appear in at least one parking lot, which contradicts the assumption.

Therefore, the final parking numbers of each parking lot $n(x_m, t^*)$ remain still, as shown in Figure 5. The equilibrium cost in the scheme p satisfies $C_p^* = C_n^*$.

By contrast, delay cost is variable. As shown in Figure 6, most commuters experience high delay costs in the spots with a high occupancy rate. If the manager raises the charge, commuters have less incentive to depart early and compete for parking spots under user equilibrium. Meanwhile, the deadweight loss of the system is also converted into internal revenue.

Theoretically, taking a parking charge which is close to the commuter's delay cost seems well in internalizing a great proportion of the system deadweight loss. However, there are two constraints.

First, if all commuters at the same parking lot do not delay or their arrival times are in a very small range, commuters arrive at nearly the same time (or in a very short time interval), but they strictly keep a specific arrival order. It is hard to implement.

Second, as the last commuters in each location have the longest search time, other commuters cannot arrive later than them. Therefore, commuters' delay time at the lot m is no less than the search time differences between them.

The search cost is determined by the order of arrival at the parking lot. To achieve the theoretical optimal charge, the manager needs to set different charges for every arrival time. We set the minimum distinguishable time interval between two adjacent arrivals in the scheme p as τ; then, the latest arrival time of the $i - th$ commuter at location x_m, t_p^{i,x_m} is

$$t_p^{i,x_m} = t_p^{e,x_m} - \left[n\left(x_m, t^*\right) - i\right] \cdot \tau. \tag{17}$$

As discussed in Section 3.1, when commuters leave the parking lot before $t^* - x_m/v_w$, they experience delay costs. The delay cost of the $i - th$ commuter satisfies

$$C_p^d\left(t_p^{i,x_m} + \frac{\lambda k_m}{k_m - i + 1} + \frac{x_m}{v_w}\right) = \beta\left(t^* - t_p^{i,x_m} - \frac{\lambda k_m}{k_m - i + 1} - \frac{x_m}{v_w}\right). \tag{18}$$

Since the last commuter at the lot m finishes parking at a time $t^* - x_m/v_w$ and arrives O on time, we calculate the delay cost of the $i - th$ commuter in the scheme p according to the difference between his/her leaving time (from the parking lot) and $t^* - x_m/v_w$.

Substituting (17) into (18), the delay cost satisfies

$$C_p^d\left(t_p^{i,x_m} + \frac{\lambda k_m}{k_m - i + 1} + \frac{x_m}{v_w}\right) = \frac{\beta\lambda k_m}{k_m - n\left(x_m, t^*\right) + 1} - \frac{\beta\lambda k_m}{k_m - i + 1} + \left[n\left(x_m, t^*\right) - i\right] \cdot \beta\tau, \tag{19}$$

where $t_p^{i,x_m} = t_p^{n(x_m,t^*),x_m} - [n(x_m, t^*) - i] \cdot \tau$. When $\tau \longrightarrow 0$, $t_p^{i,x_m} > t_n^{i,x_m}$, commuters in the scheme p arrive later than before. The delay cost is equal to the equilibrium cost minus the travel and search costs.

$$C_p^d\left(t_p^{i,x_m}+\frac{\lambda k_m}{k_m-i+1}+\frac{x_m}{v_w}\right) = C_p^*-\alpha\left(\frac{x_m}{v_w}+\frac{L-x_m}{v_c}+\frac{\lambda k_m}{k_m-i+1}\right). \tag{20}$$

$C_p^d(t_p^{i,x_m}) \le C_n^d(t_n^{i,x_m})$ when $\tau \longrightarrow 0$.

If and only if $i = n(x_m, t^*)$, $C_p^d(t_p^{i,x_m} + \lambda k_m/k_m - i + 1 + x_m/v_w) = C_n^d(t_n^{i,x_m} + \lambda k_m/k_m - i + 1 + x_m/v_w)$.

Therefore, the charge fee (zero in the scheme n) equals the difference between the search costs of the scheme p and scheme n.

$$p_p\left(x_m, t_p^{i,x_m}\right) = C_p^d\left(t_p^{i,x_m}+\frac{\lambda k_m}{k_m-i+1}+\frac{x_m}{v_w}\right) - C_n^d\left(t_n^{i,x}+\frac{\lambda k_m}{k_m-i+1}+\frac{x_m}{v_w}\right). \tag{21}$$

The arrival time t_p^{i,x_m} of the $i-th$ commuter is stipulated to satisfy

$$t_p^{i,x_m} = t^* - \frac{x_m}{v_w} - \frac{\lambda k_m}{k_m - n(x_m,t^*)+1} - [n(x_m,t^*)-i]\cdot\tau. \tag{22}$$

The arrival time of the first and last arrivals t_p^{s,x_m} and t_p^{e,x_m}, respectively, satisfies

$$t_p^{s,x_m} = t^* - \frac{x_m}{v_w} - \frac{\lambda k_m}{k_m - n(x_m,t^*)+1} - [n(x_m,t^*)-1]\cdot\tau,$$

$$t_p^{e,x_m} = t^* - \frac{x_m}{v_w} - \frac{\lambda k_m}{k_m - n(x_m,t^*)+1}. \tag{23}$$

To conclude, the manager in the scheme p should issue $n(x_m, t^*)$ parking permits to the parking lot m between t_p^{s,x_m} and t_p^{e,x_m} for $x_m \le x_p^*$. The time interval of the two adjacent issuances is τ. No parking permits will be issued at other times and parking lots.

Combined with theoretical analysis, we further draw Figure 7 for the scheme p (the legend is identical to Figure 3).

Figure 7 shows the impact of charging on commuters' schedules in the scheme p. According to Remark 1, there is no delay cost for the last parker in each lot, and his/her search cost (time) is the largest among all parkers in the lot. On the other hand, the search cost (time) is determined by the arrival order. Therefore, if the search time of the last commuter is $t_s(x_m, t)$, then the sum of delay time and search time of other commuters are no less than $t_s(x_m, t)$ in the scheme p.

The inflection points of each curve in Figure 7 represent the delay time of the penultimate commuter. It is affected by the search time of the last commuter (the longer the search, the greater the search time difference between adjacent commuters and the longer the delay for the penultimate commuter). The curve part is consistent with Figure 3 described in the scheme n. Before the inflection point, the time interval for each commuter is τ, so it is a straight line with a slope of $-k_m\tau$. The multiplier $-k_m$ is caused by narrowing the definition field from $(0, k_m]$ to $(0,1]$.

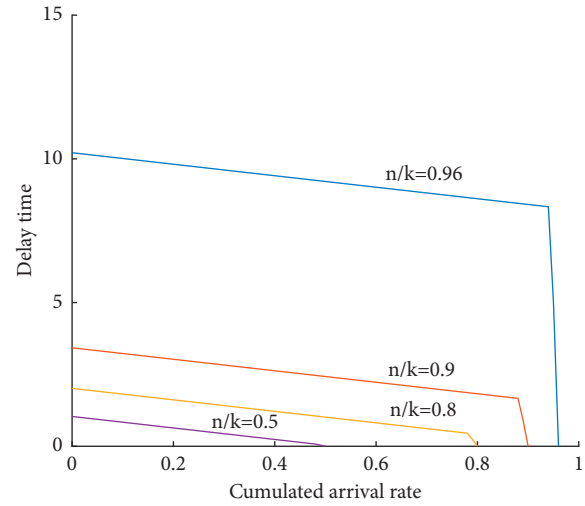


Figure 7: Delay time of different locations in scheme p.

The first $n(x_m, t^*) - 1$ commuters who arrive early will still experience nonnegative delay costs because the search time of the latest arrivals is not eliminated.

To better reflect the influence of the distance factor on the occupancy rate of parking lots, we make the discrete distribution of parking lots continuous.

Figure 8 has the same legend as Figure 4; it shows that for all parking locations, as τ is small enough, the blue dotted line (ET (depart)) is close to the orange dotted line (LT (depart)) and the blue solid line (ET (arrive lot)) is also close to the orange solid line (LT (arrive lot)), which depicts that the time composition of the first $n(x_m, t^*) - 1$ commuters can approach the last commuter before they arrive. The delay time of the first $n(x_m, t^*) - 1$ commuters has not been eliminated by permits; on the contrary, it is slightly higher than the search time of the last commuter.

From the dimensions of time, distance, and occupancy, the scheme p is depicted as three-dimensional in Figure 9. Compared with the scheme n described in Figure 6, the arrival time (to the office) of the majority of commuters is closer to t^*, and the change is more significant as the lot is closer to O.

We have analyzed the performance of the parking lot permit scheme. To approach the theoretical system optimally, it needs (1) the manager to distinguish the arrival time of each commuter and (2) all the commuters to arrive on time, which is difficult to implement in practice. As an alternative, the manager can set up a small number of groups (each group of commuters has the same charge price and expiration time). If there are η commuters in a group, the workload will be reduced to $1/\eta$ times. However, as a trade-off, the manager can only issue permits based on the last arrivals of each group and ignore the cost differences within the group. It is easy to prove that when $\eta = 1$, the scheme is equal to the scheme p; when $\eta = n(x_m, t^*)$, the scheme is equal to the scheme n. We will examine this compromise scheme in Section 5. □

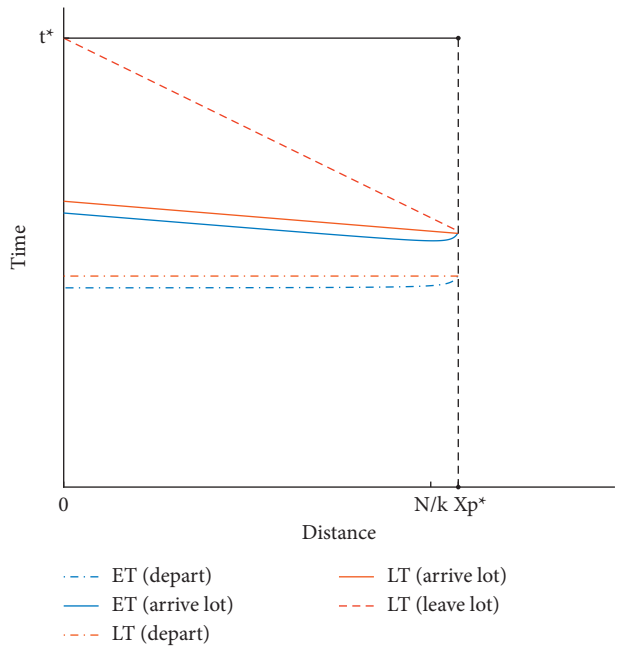


Figure 8: Trip schedule in scheme p.

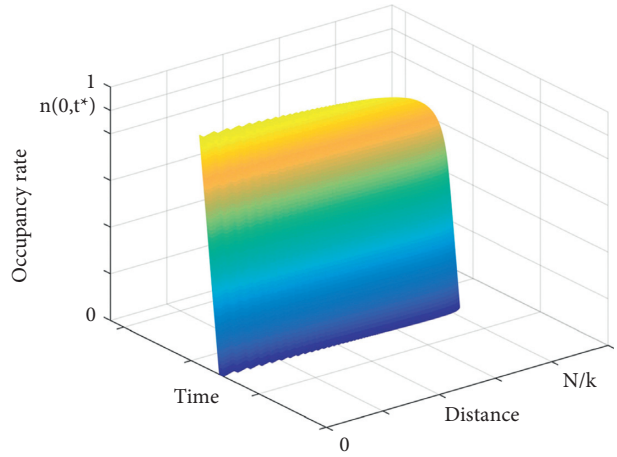


Figure 9: Arrivals at parking lots in scheme p.

4.2. Parking Spot Permit Scheme. According to Section 4.1, the contribution of the parking lot permit scheme is that a certain percentage of the delay cost can be internalized through the charging. However, it still has limitations. It cannot eliminate search cost and the efficiency loss directly related to it (e.g., early arrivals and insufficient utilization of parking lots).

Firstly, the scheme p does not change the equilibrium; the utilization of parking lots is the same as in the scheme n; despite the high demand for parking near O, the problem of parking lots not being fully utilized remains. Some commuters continue bearing more travel costs because they have to park in a farther location. The total travel cost in the scheme n is $\sum_{\{x_m | x_m \le x_n^*\}} n(x_m, t^*) C^t(x_m)$, which is higher than the total travel cost when spots are fully utilized $(\sum_{\{x_m | \sum_{x_m} k_m \le N\}} k_m C^t(x_m) + (N - \sum_{x_m} k_m) C^t(x_{m+1}))$.

Secondly, it is not able to effectively reduce the total delay cost of the system. The arrival of all commuters at one location is limited by both the arrival time and search time of the last (or last several) commuters, and the delay cost cannot be eliminated. Compared with the ideal case, it reduces the system's total delay cost by less than $1 - \beta/\alpha$.

The commuters resort to parking search when they cannot get information about the parking spots. They have to travel through the whole parking spot (or parking lot). To address this problem, we propose the parking spot permit scheme r. In this scheme, the manager issues parking permits with specific spot information. Permits are set and charged according to the location and serve time. Commuters get the spot information in advance (before departure) and go to the target spot directly. Their average search time is π.

Different from the schemes n and p, commuters' search cost no longer exceeds the equilibrium cost when $n(x_m, t) \longrightarrow k_m$ (they are not influenced by other parkers in the scheme r). Therefore, we infer that at the time t^*, all parking lots whose parking demand is greater than supply satisfy $n(x_m, t) = k_m$. The commuters park at these lots only have travel cost and fixed search cost pi. Denote the farthest parking location of the scheme r as x_r^*; according to (4), the trip cost is

$$C_r\left(x_r^*, t^* - \frac{x_r^*}{v_w} - \pi\right) = C^t(x_r^*) + \alpha\pi + p_r\left(x_r^*, t^* - \frac{x_r^*}{v_w} - \pi\right). \tag{24}$$

From (24), we can get the following remark.

Remark 5. When the scheme r achieves maximum efficiency, the charging price at x_r^* satisfies $p_r(x_r^*, t) = 0$. If the system achieves equilibrium, the minimum equilibrium cost is $C_r^* = C^t(x_r^*) + \alpha\pi$.

Proof 5. Excluding parking charges, the highest parking cost is at the farthest parking location x_r^*. If the charging price $p_r(x_r^*, t) > 0$, then commuters' cost at x_r^* and the equilibrium cost will increase, which contradicts the assumption.

We now study the utilization of each parking location. As we have discussed, the difference between the schemes r and p is that commuters in the scheme r do not need to search for parking spots when they are parking; their cost has no relationship with the current utilization of the parking lot or the order of arrival. For the lot in which m satisfies $x_m \le x_r^*$, the last commuters do not experience delay cost, and their trip cost is

$$C_r\left(x_m, t^* - \frac{x_m}{v_w} - \pi\right) = C^t(x_m) + \alpha\pi + p_r\left(x_m, t^* - \frac{x_m}{v_w} - \pi\right). \tag{25}$$

In the scheme r, parking lot m (used) is still attractive to the $k_m - th$ commuter when the parking charge is appropriately set (e.g., no more than $C^t(x_r^*) - C^t(x_m)$). In addition, the costs excluding charging of the last commuters in each parking lot are no longer equal, so when the system is in

equilibrium, the charges for parking lots $x_m \le x_r^*$ are positive.

In the discrete model, if the location of the $m' - th$ parking lot $x_{m'} = x_r^*$ in the scheme r, then m' satisfies

$$\begin{cases} \sum_{m'-1} k_m x_m < N, \\ \sum_{m'+1} k_m x_m \ge N. \end{cases} \tag{26}$$

(26) indicates that when parking lots are discretely distributed, the $m' - th$ parking lot may not be fully used.

The equilibrium cost C_r^* of the scheme r satisfies

$$C_r^* = \alpha\left[\frac{(L - x_r^*)}{v_c} + \frac{x_r^*}{v_w} + \pi\right]. \tag{27}$$

According to (24), the charging fee for the last arrivers at x_m is

$$p_r\left(x_m, t^* - \frac{x_m}{v_w} - \pi\right) = \alpha\left(x_r^* - x_m\right)\left(\frac{1}{v_w} - \frac{1}{v_c}\right). \tag{28}$$

All commuters in the same parking lot arrive at the same time. Therefore, all commuters arrive at the same time (in theory), and the first commuter is also the $i - th$ or the last, and it satisfies $t_r^{s,x_m} = t_r^{e,x_m} = t_r^{i,x_m}$ and equation (30):

$$t_r^{i,x_m} = t^* - \frac{x_m}{v_w} - \pi. \tag{29}$$

To conclude, the manager in the scheme r should issue all k_m parking permits to the parking lot m at the time $t^* - x_m/v_w - \pi$ for $x_m \le x_r^*$. No parking permits will be issued at other times and parking lots.

We continue to use the continuous distribution model of parking lots to ensure the accuracy of theoretical analysis. From the above derivation, we represent the trip schedule of commuters in Figure 10 (the legend is consistent with Figures 4 and8, and parking lots are continuously distributed). In the scheme r, the utilization rate of each parking lot is 100%, and the search time is π.

Different from Figures 4 and8, in Figure 10, (1) the farthest parking location x_r^* satisfies $\sum_{\{m|x_m \le x_r^*\}} k_m x_m = N$, which is nearer than the previous two schemes; (2) two series of lines (blue and orange dotted lines and blue and orange solid lines), respectively, represent that the earliest and latest departing and arrival times overlap with each other, which shows that the time composition of commuters who park at the same location tends to be the same and the delay time is minimized; (3) the latest arrival (orange solid) line is close to the latest leave lot (red dot) line, which shows that the search time is minimized.

From the dimensions of time, distance, and occupancy, parking distribution in the scheme r further develops to Figure 11.

Compared with Figure 9, the scheme r continues to put off the departing and arrival times of commuters. The change nearly equals the change from the scheme p to the scheme n.

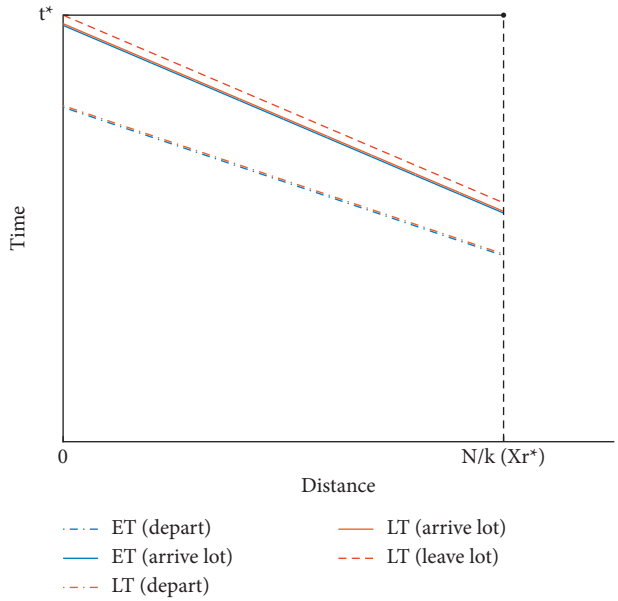


Figure 10: Trip schedule of different locations in scheme r.

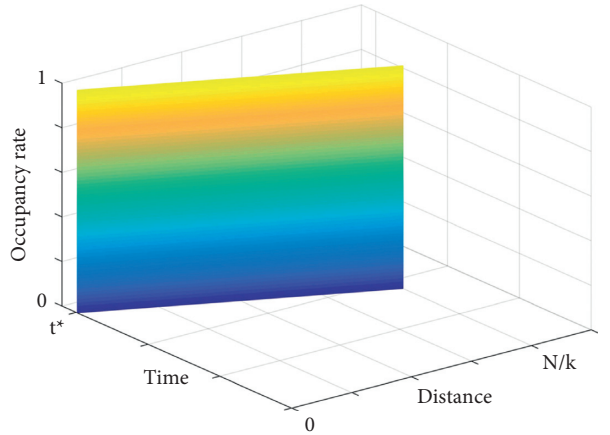


Figure 11: Arrivals at parking lots in the scheme r.

Compared with the scheme p, the delay cost in the scheme r is reduced by nearly two times, and the per-capita search cost changes to a constant value π. On the other hand, the surface changes into a plane; its projection on the "time-distance" coordinate plane is a straight line, depicting that the schedule of all commuters parked at the same location is consistent. Its projections on "occupancy-time" and "occupancy-distance" coordinate planes are rectangles, which show that all the parking lots are fully utilized in both temporal and spatial dimensions. □

5. Numerical Experiment

In this section, we provide a numerical experiment to analyze parking costs and user equilibrium of the scheme n and then quantitatively examine the effectiveness of the schemes p and r. We set the parameters as follows:

Parking Lots. Parking lot number is set as $M = 7$. The distance of the $m - th$ parking lot from the destination O is $x_m = 0.15 * (m - 1)$ kilometers.

Commuters. The commuter's number is set as $N = 1000$. The unit search time is set as $\lambda = 30$ seconds [14]. In the parking spot permit scheme r, commuters' search time under the guidance information π is set as 60 seconds according to the empirical research on factors of cruising speed and distance range [44–50].

The time value α is 1 cent per second, and that of β is 0.5 cents per second (see [50]). Commuters drive at 16 km/h and walk at 5 km/h, respectively.

Manager. The manager charges for parking spots. For simplicity, we set the parking charge as 0 yuan in the scheme n. In the schemes p and r, the manager charges variable fees according to the location and parking time.

Scenarios. The acceptable walking distance is generally within 600 meters. Therefore, we focused on the parking behavior within the scope of 1 km (L) from the destination O. From scarce to sufficient, we set four parking supply scenarios, corresponding to 7 parking lots with capacities of 200, 250, 400, and 800, respectively.

5.1. Equilibrium in the Scheme n. According to (7) and Remark 2, when the system gets to user equilibrium, the cost of all used parking lots will be equal, and when it is minimum, there is $n(x_m, t)[C_n(x_m, t) - C_n^*] \geq 0$. In theory, there is the farthest parking location x_n^*, the parking lots beyond which will not be used. Therefore, not all seven parking lots may be utilized by commuters.

From (10) and (11), the equilibrium and conservation conditions of the utilized parking lots are

$$C_n^* = \alpha\left[\frac{L - x_m}{v_c} + \frac{x_m}{v_w} + \frac{\lambda k_m}{k_m - n(x_m, t^*)}\right], \tag{30}$$

$$N = \sum_M n(x_m, t^*). \tag{31}$$

Substitute (29) into equation (30), and the equilibrium cost C_n^* can be calculated.

Based on the known conditions, we calculated the equilibrium occupancy rate $(n(x_m, t)/k_m)$ of each parking lot in four scenarios, as shown in Table 3.

The first two columns of Table 3 are the parking lot numbers and the distances from the destination O, and the last four columns represent the occupancy rate of four scenarios.

It shows that in four scenarios, the parking lot equilibrium occupancy rate decreases with the distance increase. For example, in scenario 1, the occupancy rate of parking one nearest to O is 0.932, and that of parking lot six farthest away is 0.589. Also, not all parking lots are used by commuters. With the increase in parking lot capacity, commuters tend to concentrate on parking lots closer to the destination. For example, only parking lot seven is not used in scenario 1, and the number of unused parking lots in scenarios 2–4 gradually increases.

We plot the information in Table 3 as Figure 12. It is consistent with Figure 5 (the four curves from far to near the origin, respectively, represent the occupancy rate of each parking lot when the capacity of a single parking lot is 200, 250, 400, and 800). Although the discretely distributed parking lots affect the smoothness of the curve, it still depicts the trend that the parking lot occupancy rate decreases with the increase of distance.

We further calculate the average occupancy rate of the utilized parking lots, equilibrium cost, per-capita travel cost, and per-capita deadweight loss (the sum of search and delay cost) of four scenarios in Table 4.

The first two columns of Table 4 are the scenario numbers and the single parking lot capacity; the third column is the average occupancy rate of the utilized parking lots, and the last three columns are commuters' equilibrium cost, per-capita travel cost, and per-capita deadweight loss.

In Table 4, with the increase of parking lot capacity, the equilibrium cost C_n^* decreases. The reasons are as follows. (1) There are more parking spots near the destination O that can accommodate commuters; thus, the travel costs are reduced. (2) More parking spots also reduce the commuters' search cost and the delay cost associated with it and then reduce the deadweight loss.

We can also see from Table 4 that the decline in the figures in column 5 is lower than the decline in the figures in column 6. It depicts that the increased parking capacity plays a more important role in the reduction in deadweight loss cost than the reduction in travel cost. From this point of view, it is reasonable for the manager to keep some parking spots close to the destination vacant and reduce the search cost (e.g., SFpark) and let some commuters park in distant locations.

Affected by the discrete distribution of parking lots, the average utilization rate does not decrease strictly, but we can infer from the decrease of deadweight loss that the parking supply is gradually getting ampler.

5.2. Improvement of the Parking Lot Permit Scheme. In the scheme p, because parking permits without information cannot eliminate the search cost, the trip schedule of the last commuter in each lot is the same as that in the scheme n. The utilization of each lot at the time t^* is consistent with the situation described in Table 3 and Figure 11.

Different from the scheme n, the manager in the scheme p can set time-varying charges to reduce the delay cost. Based on (2) and Section 3.2, commuters' search cost is determined by the occupied spot number, namely, the arrival order. The arrival time difference between two adjacent commuters is α/β times their search time difference. Then, we calculated the arrival time of each commuter from the search cost of the last commuters.

Table 3: The equilibrium occupancy rate of four scenarios in the scheme n.

Lot number	Distance	Scenario 1 occupancy	Scenario 2 occupancy	Scenario 3 occupancy	Scenario 4 occupancy
1	0	0.932	0.917	0.890	0.772
2	0.15	0.919	0.896	0.849	0.478
3	0.30	0.899	0.860	0.761	0
4	0.45	0.865	0.785	0	0
5	0.60	0.796	0.542	0	0
6	0.75	0.589	0	0	0
7	0.90	0	0	0	0

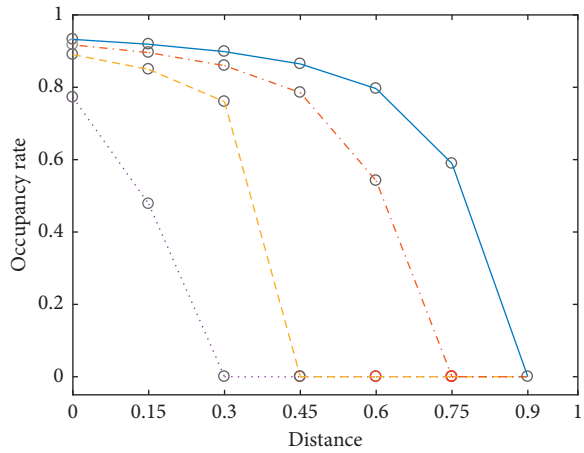


Figure 12: Occupancy rates of four scenarios (scheme n).

Table 4: The equilibrium occupancy and the costs of four scenarios in the scheme n.

Scenario number	k_m	Average occupancy	C^*	C^t	Deadweight loss
1	200	83.33%	6.69	3.95	2.74
2	250	80.00%	5.87	3.58	2.29
3	400	83.33%	4.99	2.95	2.04
4	800	62.50%	3.57	2.53	1.04

To compare with the scheme n, we also set the charge of the last commuter in the farthest lot as 0 and define two cases. Case 1 is theory-based, and τ_1 (from τ in (17)) satisfies $\tau_1 \longrightarrow 0^+$, which means the manager can accurately distinguish the arrival order of commuters and set a time-varying charge for each commuter. Case 1 can be regarded as the upper limit of system improvement. Case 2 is practice-based, and τ_2 satisfies $\tau_2 > 0$, which means it is difficult for the manager to accurately distinguish the arrival order. Instead, the manager set a small number of expirable times (set up groups include multiple commuters accordingly) and charged different parking fees. When the number of groups is 1, the scheme p is the same as the scheme n, which can be regarded as the lower limit of the system improvement (as shown in Table 4). τ_2 represents the cases between one group and $n(x_m, t^*)$ groups.

In Case 1, the manager reduces the delay time by charging the difference between the original arrival time of the commuter to the parking lot and the search time of the last commuter. Based on this, we calculate the revenue collected by charging under four scenarios.

Columns 3 and 4 in Table 5 are equal to columns 4 and 5 in Table 4. Column 6 is the parking per-capita charge in the scheme p; it is equal to the difference between the deadweight loss of scheme n (column 5 in Table 4) and scheme p (column 5). Column 7 is the percentage of parking charges to equilibrium costs, which represents the percentage of total costs that the system can reduce through charges (e.g., in scenario 1, the average parking charge is 1.01 yuan, accounting for 15.1% of the equilibrium cost of 6.69 yuan).

Table 5 depicts that although the permit scheme does not reduce the actual cost of commuters, it saves a certain percentage of the deadweight loss by charging from a system perspective. We can also see from Table 5 that with the increase of parking supply, the proportion of cost internalization by charging decreases, which is mainly due to the ease of parking search, and the delay cost associated with it also reduced.

In Case 2, we set the manager groups every 20 commuters in a parking lot and set charges based on the arrival time of the last commuter in each group. This setting greatly reduces the number of parking charges, and they are implemented simply. In this case, the deadweight losses internalized by charging of four scenarios are shown in Table 6.

The meanings of each column in Table 6 are the same as those in Table 5; since the 20 commuters are benchmarked against the last commuter in each group, their delay costs were not eliminated (only the delay cost of the last commuters in each group is eliminated); instead, the delay times are the arrival time differences between them and the last commuter.

By comparing columns 5–7 in Tables 5 and 6, the amount of deadweight loss that the manager internalizes through charges is reduced in Case 2. The effect is more obvious in scenarios 1 and 2, where parking spots are relatively scarce (with a decrease of 1.1%), while in scenarios 3 and 4, the decrease is 0.4% and 0.3%, respectively. The reason is that the increase in parking lot utilization leads to an increase in the search (delay) time for adjacent commuters.

Table 5: The equilibrium costs of four scenarios in the scheme p (Case 1).

Scenario number	km	C^*	Ct	Deadweight loss	Pp	Charge proportion
1	200	6.69	3.95	1.73	1.01	15.1%
2	250	5.87	3.58	1.48	0.81	13.8%
3	400	4.99	2.95	1.36	0.68	13.6%
4	800	3.57	2.53	0.78	0.26	7.3%

Table 6: The equilibrium costs of four scenarios in schemes s and p (Case 2).

Scenario number	k_m	C^*	C^t	Deadweight loss	p_p	Charge proportion (%)
1	200	6.69	3.95	1.80	0.94	14.1
2	250	5.87	3.58	1.54	0.75	12.8
3	400	4.99	2.95	1.38	0.66	13.2
4	800	3.57	2.53	0.79	0.25	7.0

Table 7: The equilibrium occupancy rate of four scenarios in the scheme r.

Lot number	Distance	Scenario 1 occupancy	Scenario 2 occupancy	Scenario 3 occupancy	Scenario 4 occupancy
1	0	1	1	1	1
2	0.15	1	1	1	0.25
3	0.30	1	1	0.5	0
4	0.45	1	1	0	0
5	0.60	1	0	0	0
6	0.75	0	0	0	0
7	0.90	0	0	0	0

For the same group of 20 commuters, their arrival time is relatively dense if they arrive earlier and relatively dispersed if they arrive later.

5.3. Improvement of Parking Spot Permit Scheme. In the scheme r, commuters are guided to parking spots by information other than searching for vacant spots. Therefore, the search time for each commuter is a fixed time. According to Section 4.3, the scheme r shortens commuters' parking search time and delay time and enhances the occupancy rate of parking spots.

To compare with the other two schemes, we set the charge of the commuters in the farthest lot as 0. In the scheme r, based on the known conditions, we calculated the equilibrium occupancy rate $(n(x_m, t)/k_m)$ of each parking lot in four scenarios, as shown in Table 7.

The meanings of each column in Table 7 are the same as those in Table 3; it shows that the parking distribution of commuters is more compact in the scheme r. Except for the farthest parking lots (the number of demands cannot be divided by the supply number), all the parking lots are fully used.

We plot the information in Table 7 as Figure 13. The four curves from far to near the origin, respectively, represent the occupancy rate of each parking lot when the capacity of a single parking lot is 200, 250, 400, and 800.

In Figure 13, the parking spot permit scheme changes the spatial parking distribution. Different from Figure 12, most parking lots have occupancy rates of 0 or 1. Commuters are closer to the destination O than scheme n and scheme p. More distant parking lots are left to be vacant.

According to Table 7, we further calculate the average occupancy rate of the utilized parking lots, equilibrium cost, per-capita travel cost, per-capita deadweight loss (search and delay cost), and per-capita parking charge in four scenarios in Table 8 (columns 3–7).

Compared with Table 4, a new column (column 7) (per-capita charge) is added to Table 8. In Table 8, later arrivals in each parking lot no more experience a relatively long search time, and the occupancy rate of the utilized parking lot increases. In scenarios 1 and 2, the occupancy rate of the parking lot reaches 100%; in scenarios 3 and 4, due to the discrete distribution of the parking lot, the occupancy rate remains unchanged. Compared with Table 4, the equilibrium cost and deadweight loss are all reduced in the scheme r (e.g., for scenario 1, the equilibrium cost in the schemes n and p is 6.69, while it is 5.52 in the scheme r; the deadweight loss is 2.74 in the scheme n, and 1.73 and 1.80 in two cases of the scheme p, while it is 0.60 in the scheme r). Additionally, with the increase in parking supply, the parking space span of commuters' parking decreases, so the commuters' travel cost and manager's charge also decrease with the increase of the case number.

Table 9 describes the comparison of various costs of commuters between scheme r and scheme n.

In Table 9, Columns 1–3 correspond to columns 1, 2, and 4 in Table 8, and columns 4 and 6, respectively, show the reduction of commuter's equilibrium cost and travel cost

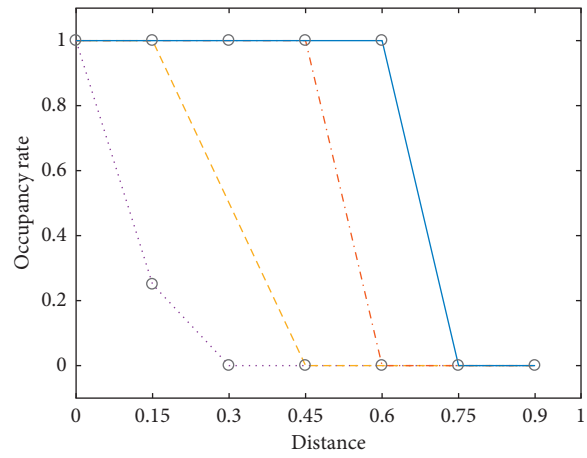


Figure 13: Occupancy rates of four scenarios (scheme r).

Table 8: The equilibrium occupancy and the costs of four scenarios in the scheme r.

Scenario number	k_m	Average occupancy	C_r^*	C^t	Deadweight loss	p_r
1	200	100%	5.52	3.74	0.60	1.18
2	250	100%	4.78	3.36	0.60	0.82
3	400	83.33%	4.04	2.84	0.60	0.60
4	800	62.50%	3.29	2.40	0.60	0.29

Table 9: Improvement of four scenarios in the scheme r.

Scenario number	k_m	C_r^*	C^* reduction	C^t	C^t reduction	p_r	System cost reduction
1	200	5.52	21.2%	3.74	5.3%	1.18	54.1%
2	250	4.78	22.8%	3.36	6.1%	0.82	48.2%
3	400	4.04	23.5%	2.84	3.7%	0.60	45.1%
4	800	3.29	8.5%	2.40	5.1%	0.29	19.0%

compared with the scheme n. Column 8 shows the reduction of system cost from the perspective of the manager (excluding commuters' parking charge). Different from schemes n and p, the scheme r brings benefits to both commuters and the system and reduces the equilibrium point.

Combined with Tables 9 and 5, in the scheme r, the equilibrium cost of commuters is significantly reduced (8.5%–23.5%) compared with the scheme n, which is about 1.2 to 1.7 times as effective as the scheme p (7.0%–14.1%). Compared with the schemes n and p, the travel cost of commuters is also reduced by 3.7%–6.1%. From the perspective of the system, the reduction of per-capita cost is more obvious (19.0%–54.1%), which is about 2.7 to 3.8 times the effect of the scheme p.

6. Conclusion and Discussion

At present, many reforms have been gradually implemented and are committed to building a smart, efficient, green, and shared urban transport system. As one of the most important trip modes in big cities, it is of great significance to improve the efficiency of parking management. However, due to the limited public parking spots in the downtown area, parking demand often exceeds parking supply. Searching for vacant spots not only raises travelers' trip costs but also intensifies the competition for parking spots, which makes travelers depart earlier or park further to guarantee a vacant spot, resulting in delay cost. In addition, the parking lot is surprisingly underutilized.

This paper studies the parking equilibrium and distribution of morning commuting and assumes all parking spots are utilized only once (in fact, the downtown parking spots often have high saturation and turnover rate, in which case the system deadweight loss is much higher).

This paper first analyzes the components of the trip cost that derives the user equilibrium and parking distribution in the no permit scheme. In that scheme, (1) parking spots cannot be fully utilized due to the high search cost at the last moment; (2) a large number of commuters arrive early, and commuters who arrive on time bear high search costs.

In response to this problem, this paper proposes two improved schemes based on the quantity control approach. The first is the parking lot permit scheme. By analyzing the causes of various costs, we classify them into two categories (invariable cost and variable cost) and find that (1) permit without parking information can only guarantee parking but eliminates the search cost depending on the utilization of the

parking spots and (2) the delay cost can be reduced, but due to the arrival order, the early arrivals have to experience at least the same delay time as the search time spent by the last arrival. Therefore, the effect of the scheme p is restricted.

We further propose a parking spot permit scheme with precise spot information and find that information plays an important role in eliminating search costs as well as reducing system deadweight loss. Commuters in the scheme r can directly go to the target parking spot instead of going through all spots to find a vacant spot, they are no longer restricted by the current parking situation, and the interaction between commuters disappears as well. In the case of insufficient parking supply, it effectively reduces the high search and delay costs. By deriving the new equilibrium, we find that (1) as the search costs are controlled (they are no longer higher than equilibrium cost), the occupancy rate of the parking spots increases and the farthest parking location gets closer to the office, and the equilibrium cost reduces as well; (2) delay is no longer related to search cost, and it can be converted into the system revenue completely by charging fees; (3) for the system, the cost is only the travel cost C^t and the average search cost $\alpha\pi$.

A numerical example is introduced to study the costs within the last kilometer. It shows the following. (1) The parking lot permit scheme improves the system by obtaining parking revenue (its compromise also works well), but the user equilibrium cost does not change. (2) The parking spot permit scheme reduces the total cost of the system to a greater extent by both improving the occupancy rate of the parking lot and reducing the user equilibrium cost. This is consistent with the conclusion of the previous analysis.

With the development of the economy and society, urban parking resources are becoming increasingly scarce; fortunately, as the information collection and transmission become more convenient, the accuracy of obtaining parking information is greatly improved. The parking permit is not only a measure but also an information carrier. It (1) guarantees the parking right in advance and prevents travelers from going back and forth between parking lots (this does not occur in the theoretical equilibrium in this paper, but it may occur when there is demand disturbance in practice) and (2) adjusts parking effectively by setting different open times and charging prices. Compared with the no permit scheme, it changes the regulation mode of supply and demand based on "price" to that based on "quantity and price," which is more accurate. In addition, the parking permits establish a one-to-one relationship between users and parking lots/spots (which is suitable for authentication) and make the management scheme feasible in parking lots of all sorts (e.g., residential shared parking lots). It will promote the development of parking management towards intelligence, sharing, and automation.

We analyze the allocation and charging mechanism under the framework of two improvement schemes for system optimum purpose (minimizes the total system cost and user equilibrium cost at the same time). However, there are still some important issues unanswered. One of them is that parking management is managed not only by the government but also by the operators for the profitable purpose; there is a difference between the purpose of profit and the purpose of publicity (higher charge and lower occupancy in general). How to bridge the gap needs to be further studied. The other parking difficulty is not only reflected in a one-time parking process but also in finding vacant spots among many repeated used ones. When the parking lot keeps on both a high occupancy rate and turnover rate, the deadweight loss cost of the system will also be much higher, especially under schemes n and $p\pi$. Parking pricing and user equilibrium should be calculated based on both the situations of the previous period and the current time interval. It is worthy of further research.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (91846202, 71890972 / 71890970, 72171020), and the 111 project (b20071).

References

[1] N. Fulman, I. Benenson, and E. Ben-Elia, "Modeling parking search behavior in the city center: a game-based approach," *Transportation Research Part C: Emerging Technologies*, vol. 120, Article ID 102800, 2020.

[2] Beijing Transport Institute, "2021 beijing transport development annual Report (in Chinese)," 2021, https://www.bjtrc.org.cn/List/index/cid/7.html.

[3] T. Litman, *Parking management comprehensive implementation guide*, Victoria Transport Policy Institute, Victoria, BC, 2011.

[4] R. C. Hampshire, D. Jordon, and O. Akinbola, "Analysis of Parking Search Behavior with Video from Naturalistic Driving," *Transportation. Research. Board*, vol. 2543, no. 1, pp. 152–158, 2016.

[5] E. Inci, J. N. Van Ommeren, and M. Kobus, "The external cruising costs of parking," *Journal of Economic Geography*, vol. 17, no. 6, pp. 1301–1323, 2017.

[6] K. Axhausen, J. Polak, M. Boltze, and J. Puzicha, "Effectiveness of the parking guidance information system in frankfurt am main," *Traffic Energy & Control*, vol. 35, no. 5, pp. 304–309, 1994.

[7] R. Arnott and J. Rowse, "Modeling parking," *Journal of Urban Economics*, vol. 45, pp. 97–124, 1999.

[8] V. R. Vuchic, *Transportation for Livable Cities*, Routledge, New York, NY, 1999.

[9] D. Shoup, "Cruising for parking," *Transport Policy*, vol. 13, pp. 479–486, 2006.

[10] D. Ayala, O. Wolfson, B. Xu, B. Dasgupta, and J. Lin, "Parking slot assignment games," in *Proceedings of the GIS '11 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 299–308, Illinois, IL, USA, November 2011.

[11] R. Arnott, A. De Palma, and R. Lindsey, "A temporal and spatial equilibrium analysis of commuter parking," *Journal of Public Economics*, vol. 45, no. 3, pp. 301–335, 1991.

[12] W. H. Lam, Z. C. Li, H. J. Huang, and S. C. Wong, "Modeling time-dependent travel choice problems in road networks with multiple user classes and multiple parking facilities," *Transportation Research Part B: Methodological*, vol. 40, no. 5, pp. 368–395, 2006.

[13] X. Zhang, H. J. Huang, and H. M. Zhang, "Integrated daily commuting patterns and optimal road tolls and parking fees in a linear city," *Transportation Research Part B: Methodological*, vol. 42, no. 1, pp. 38–56, 2008.

[14] Z. Qian and R. Rajagopal, "Optimal dynamic parking pricing for morning commute considering expected cruising time," *Transportation Research Part C*, vol. 48, pp. 468–490, 2014.

[15] Z. Y. Gu, Najmi Ali, Saberi Meead, W. Liu, and H. Rashidi Taha, "Macroscopic parking dynamics modeling and optimal real-time pricing considering cruising-for-parking," *Transportation Research Part C*, vol. 118, Article ID 102714, 2020.

[16] T. D. Crocker, "The structure of atmospheric pollution control system," *The Economic of Air Pollution*, vol. 61, pp. 81–84, 1966.

[17] J. H. Dales, "Pollution, property, and prices: an essay in policy-making and economics," *Canadian Journal of Political Science*, vol. 2, no. 3, pp. 386-387, 1969.

[18] I. Benenson, K. Martens, and S. Birfir, "PARKAGENT: an agent-based model of parking in the city. Computers," *Environment and Urban Systems*, vol. 32, no. 6, pp. 431–439, 2008.

[19] K. Dieussaert, K. Aerts, T. Steenberghen, S. Maerivoet, and K. Spitaels, "SUSTAPARK: An agent-based model for simulating parking search," in *Proceedings of the AGILE International Conference on Geographic Information Science*, pp. 1–11, Hannover Germany, June 2009.

[20] N. Levy, K. Martens, and I. Benenson, "Exploring cruising using agent-based and analytical models of parking," *Transportmetrica*, vol. 9, no. 9, pp. 773–797, 2013.

[21] M. Karaliopoulos, K. Katsikopoulos, and L. Lambrinos, "Bounded rationality can make parking search more efficient: the power of lexicographic heuristics," *Transportation Research Part B*, vol. 101, pp. 28–50, 2017.

[22] S. Lehner and S. Peer, "The price elasticity of parking: a meta-analysis," *Transportation Research Part A*, vol. 121, pp. 177–191, 2019.

[23] J. N. Van Ommeren, D. Wentink, and P. Rietveld, "Empirical evidence on cruising for parking," *Transportation Research Part A*, vol. 46, no. 1, pp. 123–130, 2012.

[24] J. B. Lee, D. Agdas, and D. Baker, "Cruising for parking: new empirical evidence and influential factors on cruising time," *Journal of Transport Land Use*, vol. 10, no. 1, pp. 931–943, 2017.

[25] F. Alemi, C. Rodier, and C. Drake, "Cruising and on-street parking pricing: a difference-in-difference analysis of measured parking search time and distance in San Francisco," *Transportation Research Part A*, vol. 111, pp. 187–198, 2018.

[26] R. C. Hampshire and D. Shoup, "What share of traffic is cruising for parking?" *Journal of Transport Economics and Policy*, vol. 52, no. 3, pp. 184–201, 2018.

[27] A. Millard-Ball, R. C. Hampshire, and R. Weinberger, "The curious lack of cruising for parking," *Land Use Policy*, vol. 91, Article ID 103918, 2019.

[28] A. Behrang, B. Douglas, and A. Paz, "Searching for on-street parking: an empirical investigation of the factors influencing cruise time," *Transport Policy*, vol. 97, pp. 186–196, 2020.

[29] G. D. Chiara and A. Goodchild, "Do commercial vehicles cruise for parking?empirical evidence from Seattle," *Transport Policy*, vol. 97, pp. 26–36, 2020.

[30] R. R. Weinberger, A. Millard-Ball, and R. C. Hampshire, "Parking search caused congestion: where's all the fuss?" *Transportation Research Part C*, vol. 120, Article ID 102781, 2020.

[31] J. van Ommeren, M. McIvor, I. Mulalic, and E. Inci, "A novel methodology to estimate cruising for parking and related external costs," *Transportation Research Part B*, vol. 145, pp. 247–269, 2021.

[32] X. N. Zhang, H. Yang, and H. J. Huang, "Improving travel efficiency by parking permits distribution and trading," *Transportation Research Part B*, vol. 45, no. 7, pp. 1018–1034, 2011.

[33] H. Yang, W. Liu, X. L. Wang, and X. N. Zhang, "On the morning commute problem with bottleneck congestion and parking space constraints," *Transportation Research Part B*, vol. 58, pp. 106–118, 2013.

[34] W. Liu, H. Yang, Y. Yin, and F. Zhang, "A novel permit scheme for managing parking competition and bottleneck congestion," *Transportation Research Part C*, vol. 44, pp. 265–281, 2014.

[35] T. Akamatsu and K. Wada, "Tradable network permits: a new scheme for the most efficient use of network capacity," *Transportation Research Part C*, vol. 79, pp. 178–195, 2017.

[36] M. Xu and S. Muller, "Tradable credits scheme on urban travel demand: a linear expenditure system Approach and simulation in beijing," *Transportation Research Procedia*, vol. 25, pp. 2934–2948, 2017.

[37] J. Wang, X. N. Zhang, and H. M. Zhang, "Parking permits management and optimal parking supply considering traffic emission cost," *Transportation Research Part D*, vol. 60, pp. 92–103, 2018.

[38] J. Wang, X. N. Zhang, H. Wang, and M. Zhang, "Optimal parking supply in bi-modal transportation network considering transit scale economies," *Transportation Research Part E*, vol. 130, pp. 207–229, 2019.

[39] J. Wang, H. Wang, and X. N. Zhang, "A hybrid management scheme with parking pricing and parking permit for a many-to-one park and ride network," *Transportation Research Part C*, vol. 112, pp. 153–179, 2020.

[40] S. P. Anderson and A. De Palma, "The economics of pricing parking," *Journal of Urban Economics*, vol. 55, no. 1, pp. 1–20, 2004.

[41] Z. C. Li, H. K. William, S. C. Lam, Z. D. L. Wong, and H. J. Huang, "Modeling park-and-ride services in a multimodal transport network with elastic demand," *Transportation Research Record*, vol. 14, pp. 101–109, 2007.

[42] A. Horni, L. Montini, R. Waraich, and K. Axhausen, "An agent-based cellular automaton cruising-for-parking simulation," *The International Journal of Transportation Research*, vol. 5, no. 4, pp. 167–175, 2013.

[43] H. Ding, *Analysing and modeling on-street parking choice behavior. M.S. thesis*, Beijing Institute of Technology, Beijing, China, 2016.

[44] Y. Yang, Z. Z. Yuan, J. Y. Li, and Y. H. Wang, "Multimode public transit OD prediction and scheduling model," *Advances in Transportation Studies*, vol. 3, pp. 133–146, 2018.

[45] P. van der Waerden, H. Timmermans, and L. Van Hove, "GPS data and car drivers' parking search behavior in the city of

Turnhout, Belgium," *Geoinformatics for Intelligent Transportation*, pp. 247–256, Springer International Publishing, 2015.

[46] R. Liu, Y. Yang, D. Kwak, D. Hang, and B. Nath, "Your search path tells others where to park: towards fine-grained parking availability crowdsourcing using parking decision models," in *Proceedings of the ACM on Interactive, Mobile, Wearable Ubiquitous Technology*, vol. 1, no. 3, pp. 1–27, 2017.

[47] W. Li, L. Bao, and Y. Li, "Comparative evaluation of global low-carbon urban transport," *Technological Forecasting and Social Change*, vol. 143, pp. 14–26, 2019.

[48] W. Li, Z. Pu, and Y. Li, "How does ridesplitting reduce emissions from ridesourcing? a spatiotemporal analysis in chengdu, china," *Transportation Research Part D*, vol. 96, Article ID 102885, 2021.

[49] Y. Yang, Z. Z. Yuan, J. J. Chen, and M. Z. Guo, "Assessment of osculating value method based on entropy weight to transportation energy conservation and emission reduction," *Environmental Engineering & Management Journal*, vol. 16, no. 10, pp. 2413–2424, 2017.

[50] K. A. Small, "The scheduling of consumer activities: work trips," *The American Economic Review*, vol. 72, no. 3, pp. 123–132, 1982.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 1785199, 14 pages
https://doi.org/10.1155/2022/1785199

Research Article

Pricing Method of the Flexible Bus Service Based on Cumulative Prospect Theory

Wanjing Ma,[1] **Yuhang Guo**,[1] **Kun An**,[1] **and Lei Wang**[1,2]

[1]*The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai 201804, China*
[2]*College of Transport and Communications, Shanghai Maritime University, Shanghai, China*

Correspondence should be addressed to Kun An; kunan@tongji.edu.cn

Received 26 October 2021; Revised 11 January 2022; Accepted 5 February 2022; Published 16 March 2022

Academic Editor: Ren-Yong Guo

Pricing directly affects the sustainable development of the flexible bus service. This study proposes a profit maximization model and a social welfare maximization model for the flexible bus operator based on the cumulative prospect theory. Fares and uncertain travel time due to unforeseen detours in serving passengers jointly affect passengers' mode choice. On the other hand, fares and passengers' probabilistic choices over the flexible bus jointly determine the profits of the flexible bus company and social welfare. This study explores the relationship between the probability of passengers choosing the flexible bus, trip fares, and uncertain travel time. Serving more passengers indicates more profits, which also results in longer detour time thus decreasing the probability of passengers choosing the flexible bus. Considering the interactive influence among passengers, we further calculate the detour time distribution. Finally, a pricing model is established to compensate for the side effects of the detour. The results show that heterogeneous fares can help the flexible bus company to obtain higher profits but have negligible influence on social welfare. In addition, the development of long-distance services and regulations over the detour time can also help to obtain more profits.

1. Introduction

The flexible bus can be regarded as a bridging mode between the conventional fixed-route transit and dial-a-ride services. The flexible bus typically uses small or medium-sized vehicles to provide door-to-door travel services subject to passenger demand [1]. Since taxies are way more expensive despite their high level of service, the flexible bus service with its high flexibility and customization capability is an attractive alternative for high-to-medium income customers [2]. It is critical for operators to seek an appropriate pricing method for the flexible bus service. On the one hand, reasonable pricing can affect the behavior of travelers, reduce the demand for private cars/taxies, and transfer it to public transit, thereby alleviating urban traffic congestion. On the other hand, pricing also directly affects the financial sustainability of the flexible bus company.

In recent years, a growing number of studies investigated the pricing problem for many emerging travel modes such as flexible bus and customized bus. Li et al. [3] formulated a competitive game model in which the objective was to maximize the profits of customized bus services and ride-sharing based on passengers' transport mode choices. Sayarshad and Gao [4] developed a dynamic pricing scheme that utilized a balking rule considering socially efficient levels and revenue-maximizing price. An equilibrium-joining threshold was obtained by imposing a toll on the customers who joined the on-demand mobility system. Kaddoura et al. [5] proposed an agent-based transport simulation to investigate different design concepts for the demand-responsive transit. The simulation results show that a small service area and low prices may result in an unwanted mode shift effect from walk and bicycle to Demand Responsive Transit. Gong et al. [6] constructed a game theory model between the customized city bus service and the conventional urban bus transportation to maximize the profits of the two transportation modes. Wang et al. [2] integrated the disaggregated trip choice model with the

vehicle routing model to determine incentive schemes. They concluded that passengers' sensitivity towards incentives is decisive to the result.

In addition, there are many studies on the pricing for other on-demand responsive travel modes, such as ride-hailing and ride-pooling. Ozkan [7] studied the interrelationship between pricing and matching decisions of a ride-sharing firm. He formulated a stylized ride-sharing model that captured customer and driver behaviors considering the geospatial nature of the system. The results showed that optimizing the pricing decisions alone with fixed matching rules did not increase the number of matchings in general. Yan et al. [8] designed and implemented a matching and pricing algorithm at scale to strike a balance between model complexity and accurate description of the marketplace dynamics. Wang et al. [9] found that if the platform offers the carpool service option, they can achieve a larger market coverage and the riders can enjoy more affordable rides without compromising on service quality. Bai et al. [10] considered an on-demand service platform using earning-sensitive independent providers with heterogeneous reservation prices (for work participation) to serve its time and price-sensitive customers with the heterogeneous valuation of the service. Zhong et al. [11] examined how an on-demand ride-hailing platform in competition with the traditional taxi industry designs its pricing strategies under unregulated and regulated pricing scenarios. They found that the monopolistic on-demand ride-hailing platform's price rate and profit under the unregulated pricing scenario are relatively higher than those under the regulated pricing scenario.

Detours are a problem faced by both the flexible bus and carpooling companies. However, few studies considered how the detour time would affect the pricing strategy of the flexible bus. Ke et al. [12] innovatively established a set of nonlinear equations to explore the relationships between the platform decision variables (i.e., trip fare) and endogenous variables (e.g., actual detour time) in ride-sourcing markets with and without on-demand ride-pooling services. By considering the extra detour time experienced by passengers and drivers, they found that decrease in trip fare not only directly increases passenger demand due to negative price elasticity, but also reduces actual detour time, which in turn increases passenger demand. Zhang and Nie [13] established a market equilibrium based on a spatial driver–passenger matching model that determines the passenger wait time for both solo and pooling rides. They found the system's benefit diminishes quickly as the average en-route detour time increases.

Both ride pooling and flexible bus are travel demand-responsive travel modes and how to pricing for both of them considering the uncertainty of the detour time is a challenging problem. However, a flexible bus can serve more passengers at the same time, making possible detours outcomes more complicated. The uncertainty in detour time affects passengers' travel decisions to a greater extent, thereby affecting demand. Only few pricing studies considered travel decisions under uncertain travel time conditions. Choi et al. [14] applied the mean-risk theory to analytically explore how the risk attitude of customers affects the optimal service pricing decision of the on-demand platform. Wu et al. [15] proposed a choice-based framework for modeling the supply/demand interaction in risky choice contexts. The model allowed the system operator to set an optimal pricing strategy regardless of whether user risk preferences are risk-seeking or risk-averse.

There are few studies on the relationship between pricing, detours, and passenger decision-making under uncertain conditions. Instead, most existing studies considered the relationship between price and demand from a macro perspective. Flat and homogenous fare is provided to all passengers [3, 6]. When passengers receive different service levels in terms of detour level, their fares should be different.

Cumulative prospect theory (CPT) is widely used to solve travel decision-making problems and transportation pricing problems. As a descriptive decision-making model under uncertainty, the prospect theory was proposed by Kahneman and Tversky [16] as a critique of the expected utility theory. In 1992, Kahneman and Tversky [17] extended their model to CPT. CPT is based on the assumption of bounded rationality of decision-makers and can describe decision-makers' behavior under uncertain conditions more accurately. Katsikopoulos et al. [18] found that drivers were risk-averse when choosing among routes in the gain domain and risk-seeking in the loss domain. It is consistent with the view of CPT. Sepehr et al. [19] introduced a CPT-based framework for mode choice modeling using observational data. The framework utilized CPT for modeling reliability in the trip-based models. CPT has also been widely applied to congestion pricing. Liu et al. [20] considered the psychological factors of passengers in the congestion pricing model and verified the feasibility of the model based on user equilibrium and CPT. Xu et al. [21] developed an optimal congestion pricing model in which user equilibrium was adopted to capture travelers' response to pricing signals under risk based on CPT. However, the application of CPT in the pricing of the flexible bus is limited.

We aim to study the pricing of the flexible bus service considering the interaction between passenger travel behavior and fares and thus propose a profit-maximization pricing strategy for the flexible bus system based on CPT. Considering the travel time of the flexible bus is uncertain due to its high route flexibility, we can apply CPT to our pricing model. The relationship of detour time distribution, fares, and acceptance of detours (choosing to take the flexible bus) are explored. On this basis, considering the detour problem of the flexible bus, we build a pricing model under static demand for the flexible bus company with the goal of expected profit maximization and social welfare maximization.

In summary, the aims of this research are as follows:

(1) Under uncertain travel time of trips by the flexible bus, we explore the influence of fares and uncertain travel time on passengers' travel mode choices based on CPT

(2) We establish two customized pricing models with the objective of profit maximization for the flexible

bus company and social welfare, considering the impacts of trip detours on-demand loss

The remainder of this paper is organized as follows. Section 2 establishes the mode choice model and the pricing model. The parameters are calibrated through a stated preference (SP) survey. Section 3 describes the design of two case studies and analyzes the results. The last section summarizes the innovations and conclusions.

2. Methodology

In this section, we first introduce the hypotheses used in the flexible bus pricing problem. Based on prospect theory, we calculate the perceived travel time utility of passengers. Taking the travel time utility and out-of-pocket costs as the components in the utility function, we can obtain the probability of passengers taking flexible buses for a given detour time. In addition, considering the mutual influence among passengers in each other's travel time, we develop a series of equations to obtain the detour time distribution. Finally, a flexible bus pricing model is constructed, which aims at maximizing the bus company's profit and maximizing the social welfare respectively. The model parameters are calibrated through the SP survey and Maximum Likelihood Estimation (MLE). The main ideas are shown in Figure 1.

2.1. Problem Description. We consider regional flexible bus service in this study [22], which allows vehicles to operate in a demand-responsive way through reservation within the service area. Passengers can only board or alight in the predefined service areas. The pick-up region and drop-off region are connected by a non-stop fast route. It can be regarded as a variant of customized buses. Figure 2 illustrates the operation of the regional flexible bus. In this paper, we aim to investigate the pricing problem for this regional flexible bus service.

The flexible bus adopts small and medium-sized vehicles and provides door-to-door services based on reservation. In the responsive area, the stops and bus routes depend on the demand of passengers, and the demand of passengers is elastic which is subject to the utility of the flexible bus. Compared with the taxi, the flexible bus will detour when delivering multiple passengers. More passengers to be served will probably lead to longer detour time. To passengers, a longer travel time would decrease the utility of the flexible bus thus affecting his/her mode choice probability. It is necessary to make reasonable pricing to compensate for the side effects of detours so as to make flexible buses competitive with taxies. In addition, as flexible buses may adjust routes dynamically while en-route, the travel time of passengers on-board may be prolonged and thus is uncertain. Passengers' risk attitude and personal preference also influence passengers' decision-making in the case of uncertain travel time. Therefore, it is necessary to accurately measure the impact of these factors on passenger decision-making to obtain the probability of passengers choosing the flexible bus. The mode choice probability and ticket price jointly affect the expected profit of the flexible bus company and the social welfare. This section solves the pricing problem considering the interactions among passengers in the system. The price for each passenger is optimized to maximize the expected profit and the social welfare respectively.

Due to the similarity of flexible bus and taxi services, we assume passengers can choose between the flexible bus and the taxi for their trip. They mainly consider the travel time and the fare assuming the other influencing factors are the same. The flexible bus knows the demand of passengers in advance through the reservation, but it is uncertain whether passengers will take it in the end. After knowing the passengers' demand, the flexible bus will plan the route. If the passenger eventually chooses to get on the bus, the bus will pick him up. According to the probability of passenger boarding, the flexible bus can estimate the travel time through the probability of passengers getting on the bus. Before boarding, the flexible bus displays passengers the fare, travel time, possible detour time, and corresponding probabilities (in this paper, they are captured by the detour time distribution). Detour time refers to the increased time for passengers due to picking up other passengers (the time loss for waiting for boarding, etc.) after boarding compared with the shortest travel time from the origin to the destination. The shortest travel time is equal to the travel time of a taxi. The distribution of detour time refers to the detour time and its corresponding probability.

2.2. Behavior of Passengers. When passengers make mode choice decisions, they will be affected by many factors such as travel time, cost, and uncertainty of travel time [5]. In addition, passengers will also be influenced by their risk attitudes and preferences [21]. Considering the characteristics of the flexible bus, serving other passengers leads to the increase of uncertain detour time for the loaded passengers. The uncertain travel time and fare jointly affect the passengers' mode choice decisions. This part explores the interrelationship between the distribution of detour time, the fare, and the acceptance of detours under the condition of uncertain travel time of the flexible bus.

2.2.1. Perceived Time Utility. Most of the traditional travel decision-making models are based on the expected utility theory, assuming that the decision-maker is entirely rational. In fact, people in the process of making travel decisions are often affected by the traveler's habits, attitude towards risk, preferences, and other factors. Decision-makers often can-no't be entirely rational.

CPT was developed by Kahneman and Tversky [17] based on prospect theory (PT), which is different from the traditional expected utility theory (EUT). CPT has three main observations about an individual's bounded rationality. (i) People usually consider possible outcomes relative to a certain reference point (x_0) rather than to the final state. The payoffs are defined as the gains or losses relative to x_0 before making choices. The payoffs that people perceive can be described by a concave function for gains and a convex function for losses which results in diminishing sensitivity.

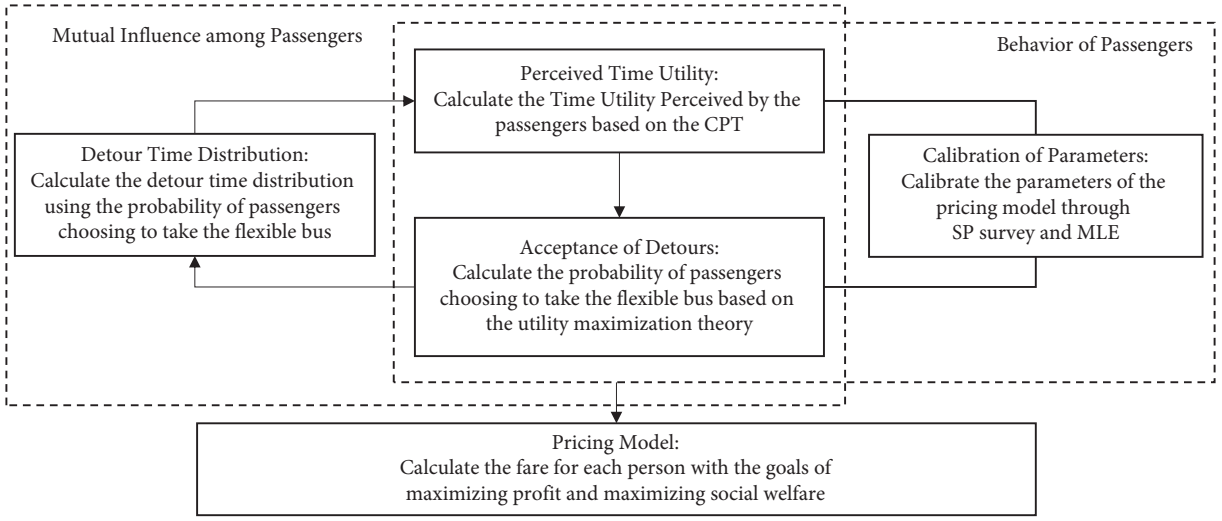


Figure 1: Main idea of the method.

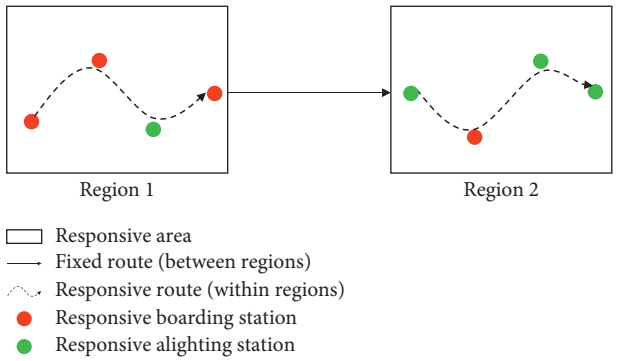


Figure 2: Operation of the regional flexible bus.

(ii) People tend to be loss aversion. People are more sensitive to losses than gains. (iii) People tend to overweight the probability of extreme, but rare events, and underweight more common events. People's decisions are affected by the decision weights, not actual probabilities. In conclusion, CPT argues that decision-makers perceive values $v(x_i)$ (travel time in the case of this paper) differently from the actual values x_i, and probabilities P_i (probability of experiencing a certain travel time for a given trip) are converted to decision weights $w(P_i)$.

Although some studies have considered the influence of detour time, they have not considered the uncertainty of detour [12]. Due to demand uncertainty, the travel time of the flexible bus is uncertain. Therefore, the choice between flexible public transportation and other modes of transportation is a typical travel decision-making problem under uncertain conditions. Given that CPT provides a well-supported descriptive paradigm for individuals' decision-making under risk or uncertainty, we can apply CPT to measure travelers' perceived utility of time.

We assume that a passenger intends to choose between a flexible bus and a taxi for a trip. Choosing each alternative can lead to k different outcomes, quantified by $x_1, x_2, \ldots, x_k$, The outcome i happens with probability P_i. Each alternative can be seen as a prospect (x_i, P_i). The following value function proposed by Kahneman and Tversky [17] is used in this paper:

$$v(x_i) = \begin{cases} (x_i - x_0)^{\alpha}, x_i \geq x_0, \\ -\lambda (x_0 - x_i)^{\beta}, x_i < x_0, \end{cases} \quad \text{where } 0 < \alpha, \beta < 1, \lambda \geq 1, \tag{1}$$

where x_0 is the reference point; α is the exponent of the value function over the gain region, and β is the exponent of the value function over the loss region. $v(x_i)$ is the value function, reflecting an individual's perceived value. The parameters α and β measure the degree of diminishing sensitivity of the value function. λ is the loss aversion coefficient, indicating that individuals are more sensitive to losses than gains.

The following probability weighting function by Kahneman and Tversky [17] is used:

$$w(P_i) = \frac{P_i^{\gamma}}{\left[P_i^{\gamma} + (1 - P_i)^{\gamma}\right]^{1/\gamma}} \text{ where } 0 < \gamma < 1, \tag{2}$$

where γ is the probability weighting parameter, representing the level of distortion in probability judgment in the decision-making process. $w(P_i)$ is the probability weighting function, reflecting an individual's perception of probability. Figures 3 and 4 show the value function and the weighting function. As shown in Figures 3 and 4, the value function $v(x_i)$ is S-shaped, concave in the gain region and convex in the loss region. Besides, individuals are more sensitive to losses than gains. Furthermore, the probability weighting function $w(P_i)$ expands the influence of rare events and shrinks the influence of common events.

The cumulative prospect value is defined as

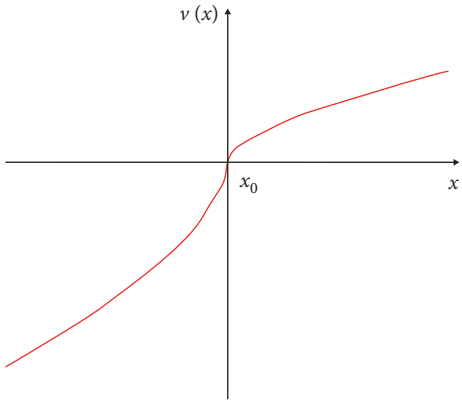


Figure 3: Value function.

$$CPV = \sum_{k} v(x_i)\pi(P_i), \tag{3}$$

$$\pi(P_i) = w(P_i + P_{i+1} \cdots + P_k) - w(P_{i+1} + P_{i+2} \cdots + P_k), \quad where\ i = 1, 2, \ldots, k-1, \tag{4}$$

$$\pi(P_k) = w(P_k), \tag{5}$$

The decision weight $\pi(P_i)$ is calculated based on the cumulative distribution function. All potential outcomes are ranked in increasing order in terms of preference. In these equations, k is the number of all potential outcomes, and i denotes a generic outcome.

For mode-alternative m, the travel time can be $t_{m,1}, t_{m,2}, \cdots, t_{m,k}$ with probabilities $P_{m,1}, P_{m,2}, \cdots, P_{m,k}$, respectively. To apply CPT, we need to define the reference point, the value function, and the weighting function first. In the case of the flexible bus detour problem, people will pay more attention to the increase in detour time compared with the shortest travel time when making travel decisions. Therefore, the shortest travel time t_0 (equal to taxi travel time) is selected as the reference point. (1) is assumed to be the value function. (2) is assumed to be the probability weighting function. Since $t_{m,i} \geq t_0, \forall i \in k$, we only model in the loss domain.

The calculation process of perceived time utility $CPV(t_m)$ is as follows:

$$v(t_{m,i}) = -\lambda(t_{m,i} - t_0)^{\beta}, t_{m,i} \geq t_0, \tag{6}$$

$$w(P_{m,i}) = \frac{P_{m,i}^{\gamma}}{\left[P_{m,i}^{\gamma} + (1 - P_{m,i})^{\gamma}\right]^{1/\gamma}}, \tag{7}$$

$$\pi(P_{m,i}) = w(P_{m,i} + P_{m,i+1} \cdots + P_{m,k}) - w(P_{m,i+1} + P_{m,i+2} \cdots + P_{m,k}), \quad where\ i = 1, 2, \ldots, k-1 \tag{8}$$

$$\pi(P_{m,k}) = w(P_{m,k}), \tag{9}$$

$$CPV(t_m) = \sum_{i=1}^{k} v(t_{m,i})\pi(P_{m,i}), \tag{10}$$

where $CPV(t_m)$ is the perceived time utility of mode m. (6) calculates the perceived utility $v(t_{m,i})$ corresponding to the travel time $t_{m,i}$. (7), (8), (9) calculate the decision weight $\pi(P_{m,i})$ corresponding to the travel time $t_{m,i}$. (10) calculates the total perceived time utility of mode m.

2.2.2. Acceptance of Detours. We assume that the speeds of the flexible bus and the taxi are equal. The difference in travel time between the flexible bus and the taxi is because that the flexible bus needs to detour to serve multiple passengers in one trip. Therefore, we define the probability of passengers choosing a flexible bus as the acceptance of detours. According to the assumption, the fare will also have an important impact on passengers' decision-making.

When travelers make travel decisions, they always choose the one with a larger utility, which is called the utility maximization theory. The formula of the utility maximization theory is as follows:

$$V_m = \sum_{A} \theta_m y_m, \tag{11}$$

where y_m is the characteristic variable and θ_m is the parameter reflecting travelers' perceived relative importance of different attributes. According to the discussion above, travel time and travel cost are selected as characteristic variables. Considering that travel time is uncertain, the utility of travel time is more suitable to be described by CPT. The travel cost is a deterministic factor. Therefore, the probability estimation and risk attitude of travelers will not affect the utility of the travel cost. The modified utility function is

$$MV_m = \beta_0 \sigma_m + \beta_t CPV(t_m) + \beta_c C_m, \tag{12}$$

where C_m is the travel cost of mode m; $CPV(t_m)$ is the perceived travel time utility of mode m; σ_m is a dummy variable indicating whether a particular mode-alternative is the flexible bus or not; β_0 is the mode-specific constant added to capture the effect of unforeseen variables; β_c is the coefficient of cost, and β_t is the coefficient of time. CPT can be used to calculate the value of the modified utility (MV_m) of each mode (the flexible bus and the taxi) and the mode decision between the two modes. Acceptance of detours can be described by a Logit function:

$$p = \frac{\exp(MV_{fb})}{\exp(MV_{fb}) + \exp(MV_{taxi})}, \tag{13}$$

$$p = \frac{1}{1 + \exp\left(-\beta_0 + \beta_t\left(CPV(t_{taxi}) - CPV(t_{fb})\right) + \beta_c \Delta C\right)}, \tag{14}$$

$$\Delta C = C_{taxi} - C_{fb}, \tag{15}$$

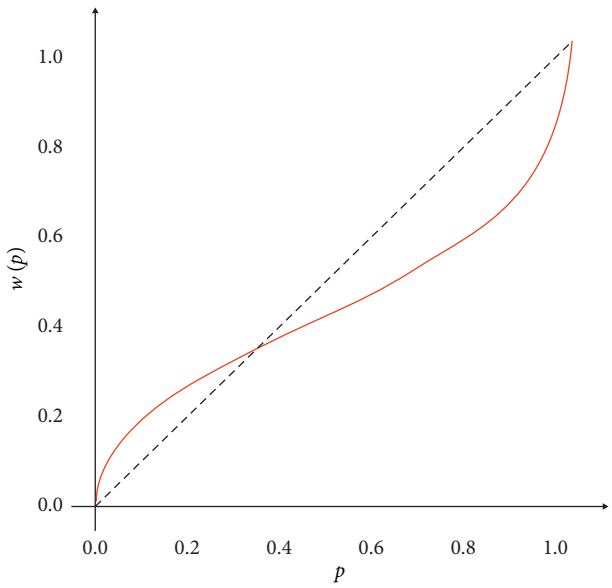


Figure 4: Weighting function.

where p is the acceptance of detours; fb refers to the flexible bus; MV_{fb} is the modified utility of the flexible bus; MV_{taxi} is the modified utility of the taxi, and ΔC is the difference between taxi cost and flexible bus cost. $CPV(t_{taxi})$ is equal to 0 and $CPV(t_{fb})$ can be calculated by (6), (7), (8), (9), and (10). Therefore, passengers' acceptance of detour can be simplified as

$$p = \frac{1}{1 + \exp\left(-\beta_0 + \beta_t \lambda \sum_{i=1}^{k} (\Delta t_i)^{\beta} \pi(P_i) + \beta_c \Delta C\right)}, \quad (16)$$

$$\Delta t_i = t_i - t_0, \quad (17)$$

where Δt_i is the detour time for taking the flexible bus.

2.3. Detour Time Distribution. In (16), the probability of taking the flexible bus is affected by the perceived time utility and fare. The perceived time utility is jointly affected by the detour time and its corresponding probability. Calculating the perceived time utility of a passenger requires the detour time distribution. By analyzing the reasons for the detour, we propose a method to calculate the time distribution of the detour. For a certain passenger, the detour he/she faced is caused by serving other passengers during his/her ride. Generally speaking, the more stations served, the longer the detour time the passenger suffers. However, the detour does no't necessarily happen. It is related to the probability of passengers choosing to take the flexible bus. Considering that the passenger demand is usually densely distributed in a certain area, it can be assumed that, for each passenger, the detour time increases by T for serving every extra station (including the boarding station and the alighting station).

To explain the method, we consider a case of five passengers, A-E, and they get on and off at given stations. Figure 5 shows everyone's boarding and alighting stations. Since the flexible bus provides door-to-door service, the boarding and alighting stations are also the origin and the destination respectively. Take A as an example, if B-E choose to take the flexible bus, A will be affected by the detour caused by serving B-E. If B-E do no't take the flexible bus, A will not suffer a detour. This is related to the probability that B-E choose to take the flexible bus. The probabilities of B, C, D, and E taking the flexible bus are p_B, p_C, p_D, and p_E respectively. They further affect the probability of the flexible bus serving Stations 2–5. The probabilities of serving Stations 2–5 are sp_2, sp_3, sp_4, and sp_5 respectively, which can be calculated as follows:

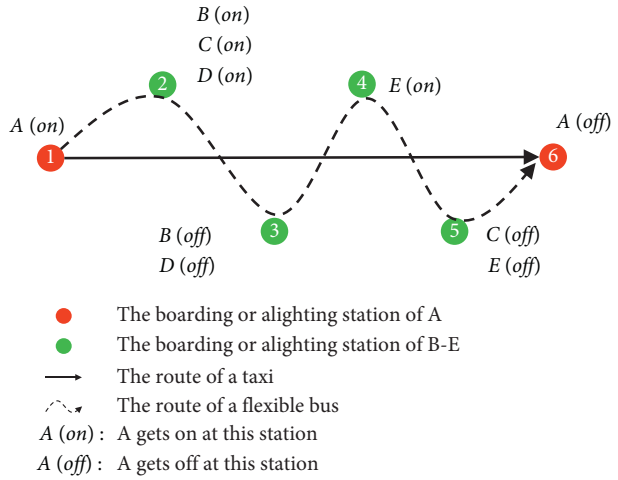


Figure 5: Boarding and alighting stations of A-E.

$$sp_2 = 1 - (1 - p_B)(1 - p_C)(1 - p_D), \quad (18)$$

$$sp_3 = 1 - (1 - p_B)(1 - p_D), \quad (19)$$

$$sp_4 = p_E, \quad (20)$$

$$sp_5 = 1 - (1 - p_C)(1 - p_E), \tag{21}$$

According to the assumption, the detour time increases by T for serving every extra station. For A, the detour time can be 0, T, $2T$, $3T$, and $4T$ with probabilities P_1, P_2, P_3, P_4, and P_5, respectively. By combining Stations 2–5, the corresponding probability of the detour time can be obtained. Taking the detour time $\Delta t = 0$ as an example, the corresponding probability is calculated as follows:

$$P_1 = P(\Delta t = 0) = (1 - sp_2)(1 - sp_3)(1 - sp_4)(1 - sp_5), \tag{22}$$

where sp_2, sp_3 sp_4, and sp_5 can be calculated by (18), (19), (20), and (21). P_2, P_3, P_4, and P_5 can be calculated in the same way. The possible detour time for A is (0, T, $2T$, $3T$, $4T$, $5T$). Then, we can get the detour time distribution of A. Similarly, we can use this method to calculate the detour time distribution for customers B-E. This idea of calculating detour time distribution can then be extended to a flexible public transport system with Q individuals. In general, for a passenger, his/her detour time is related to the number of other stations served by the flexible bus. The probability corresponding to the detour time is determined by the probability of passengers choosing to take the flexible bus in other stations.

2.4. Pricing Model

2.4.1. Service Provider Profit. The profit of flexible buses is the fare revenue minus the operating cost. The expected fare revenue of passenger is

$$r_j = x_j \cdot p_j, \tag{23}$$

where x_j is the fare of the flexible bus for passenger j and p_j is the probability of passenger j taking the flexible bus. According to (12), (13), (14), (15), (16), and (17), p_j is related to detour time and the difference between taxi fare c and flexible bus fare x_j:

$$p_j = \frac{1}{1 + \exp\left(-\beta_0 + \beta_t \lambda \sum_{i=1}^{k} \left(\Delta t_{i,j}\right)^{\beta} \pi\left(P_{i,j}\right) + \beta_c\left(c - x_j\right)\right)}, \tag{24}$$

Taxi fare c is related to the travel distance of the passenger. The travel distance is composed of the travel distance within the region and the travel distance between regions. The travel distance between regions is a constant, while travel distance within the region is related to the origin and the destination of the passenger. Considering that the travel distance between regions accounts for the main part, we assume that ticket fare and the distance between regions are positively correlated. The taxi fare c is:

$$c = \mu s, \tag{25}$$

where μ is the fare of taxi per kilometer and s is the travel distance between regions.

The operating cost is assumed to be L, which is the system parameter related to operating distance, the vehicle types, fuel consumption, and other factors of the flexible bus system. The calculation of operating cost is similar to the calculation method of taxi fare. The operating distance is composed of the operating distance within the region and the operating distance between regions. The operating distance between regions is a constant, while operating distance within the region is related to the spatial distribution of passengers. Besides, detours within the region will also affect the operation costs. The operating cost L is

$$L = \varphi(s + v\overline{\Delta t}), \tag{26}$$

where φ is the fare of taxi per kilometer; s is the travel distance between regions; v is the average speed of the flexible bus and $\overline{\Delta t}$ is the expected detour time. $\overline{\Delta t}$ can be calculated by the distribution of detour time mentioned in Section 2.4.

The profit of the service provider Z_1 can be expressed as

$$Z_1 = \sum_{j=1}^{Q} x_j \cdot p_j - L, \tag{27}$$

where L is the operating cost; p_j is the probability of passenger j taking the flexible bus, and c is the fare of the taxi.

2.4.2. Passenger Surpluses. The passenger surplus is the fare that a passenger is willing to pay for the flexible bus service minus the actual out of pocket costs. The fare that a passenger is willing to pay for the flexible bus service can be expressed as:

$$W = c - U_{\Delta t}, \tag{28}$$

$$U_{\Delta t} = \beta_t \lambda \sum_{i=1}^{k} \left(\Delta t_i\right)^{\beta} \pi\left(P_i\right), \tag{29}$$

where W is the fare that a passenger is willing to pay, and $U_{\Delta t}$ is the disutility caused by detours.

The passenger surpluses Z_2 can be expressed as:

$$Z_2 = \sum_{j=1}^{Q} \left(W_j - x_j\right), \tag{30}$$

2.4.3. Optimization Model of Fares. We consider two objectives (profit maximization and social welfare maximization) to optimize the fare of the flexible bus.

(i) Expected profit maximization: compared with the taxi, the flexible bus will detour, so the fare of the flexible bus will not exceed the fare of the taxi. The profit maximization model of the flexible bus can be expressed as

$$\max Z_1 \tag{31}$$

s.t. (12)-(27)

$$0 \le x_j \le c, \tag{32}$$

(ii) Social welfare maximization: social welfare includes the profit of the service provider and passenger surpluses, considering the interests of the flexible bus operator and the passengers simultaneously. The profit of the flexible bus operator should not be less than the threshold θ. The social welfare maximization model of the flexible bus can be expressed as

$$\max Z_1 + Z_2, \tag{33}$$

s.t. (12)-(30)

$$Z_1 \geq \theta, \tag{34}$$

(12), (13), (14), (15), (16), and (17) calculate the probability of passenger j taking the flexible bus. (18), (19), (20), (21), and (22) propose a method of calculating the time distribution of each passenger. (23) calculates the expected fare revenue of passenger j. (24) displays the probability of passenger j taking the flexible bus. (25) calculates the taxi fare. (26) calculates the operating cost. (27) calculates the service provider profit. (28) calculates the fare that a passenger is willing to pay for the flexible bus service. (29) calculates the disutility caused by detours. (30) calculates the passenger surpluses. The objective function aims at maximizing the expected profits of the flexible bus and maximizing the social welfare respectively. Constraint (32) ensures that the fare is within a reasonable range. Constraint (34) ensures that the profit of the flexible bus operator should not be less than a certain threshold.

Combined with the calculation of detour time distribution proposed in Section 2.3, we can find that the passengers in the flexible bus system affect each other. Take the service order as A-B-A-B as an example, the reduction in A's fare will increase A's probability of choosing the flexible bus. Because of the increase in the probability of A choosing the flexible bus, the probability of B's detour also increases, resulting in a decrease in the probability of B choosing to take the flexible bus. Therefore, the fare of each passenger directly affects his probability of taking the flexible bus, and indirectly affects the probability of other people's detours, thereby affecting the probability of other passengers' taking the flexible bus. Figure 6 shows the relationship between A and B. Therefore, a reasonable price should be set for each passenger in the flexible bus system to maximize the expected profit or the social welfare.

The final output of the pricing model is the optimal fare that is shown to each reserved passenger before the flexible bus departs. However, the model is under static demand. The actual operation of a flexible bus is a dynamic process. When a passenger refuses service or completes the service, new passengers will enter the system. This situation will lead to an increase in detour time. It will reduce the level of service for loaded passengers who have already received the initial fare on the bus. Correspondingly, the fare should be reduced. From the perspective of fairness to passengers, the result calculated by the profit maximization model should be the highest price given by the flexible bus in real-world operation.

2.5. Calibration of Parameters. To calibrate the parameters of our model, we conducted an online stated preference survey. The SP survey includes three parts: The first part is the individual socioeconomic attributes of the participants. The second part is regarding people's risk attitudes towards time. We design a set of binary travel plan choices. Each has a different travel time distribution. The average travel time of the choice is equal, but the variance is not. Participants need to choose their preferred travel plan. The third part is related to mode choice between the flexible bus and the taxi. Finally, 246 questionnaires were returned in this SP survey. After sorting the questionnaires, 205 valid questionnaires were obtained, with an effective rate of 83%. Overall, the collected data samples are representative. The model parameters were calibrated using the data of the SP survey.

In many existing studies on travel decision-making based on CPT, researchers use the parameters in the value and weighting functions which are directly estimated by Kahneman and Tversky. However, parameters in Kahneman and Tversky are estimated from the results of gambling in the economic field. In different scenarios, the parameters β of the value function in the loss domain of different travelers should be different [21].

The values of β in our model are estimated based on the experimental results of stated choice questions of the second part and the CPV parameter values (i.e., $\lambda = 2.25$ and $\gamma = 0.69$). A logit model and MLE are used to estimate the value function parameter β. On this basis, the values of β_0, β_t, β_c in our model are estimated by using logistic regression. Our estimation results are $\beta = 0.60$, $\beta_0 = -0.326$, $\beta_t = 0.252$, and $\beta_c = -0.285$. The rate of substitution between CPV(t) and travel cost in our study is $0.252/0.285 = 0.88$ RMB/min. For passengers, the detour is essentially a loss. Compared with taxi fares, the decrease of flexible bus fares is compensation for detours. Therefore, the results mean that travelers' willingness to accept every unit increase in CPV(t) is about 0.88 RMB per minute in our SP survey.

Passengers' willingness to accept CPT(t) includes the willingness to accept detour time and the willingness to accept the uncertainty of travel time. Most studies on the value of time or uncertainty addressed users of cars, buses, and taxis, while few considered ride-sharing or flexible bus passengers [23]. There are mainly three approaches considering reliability in the utility maximization theory. Small [24] proposed a scheduling model to analyze travelers' departure time choices to ensure on-time arrival. Small and Noland [25] improved Small's scheduling model to understand choices under uncertainty by adding the probability distribution of travel time. The mean lateness at departure and/or arrival [26] is another approach to measuring the value of uncertainty. The mean-variance (MV) model [25] estimates the values of travel time and uncertainty within the utility maximization framework.

We choose the mean-variance model to compare with our model based on CPT framework. In the mean-variance model, the utility of mode i is

$$U_i = \beta_t E(t_i) + \beta_{SD} SD(t_i) + \beta_c c_i + \beta_\sigma \sigma_i, \tag{35}$$

where $E(t_i)$ is the expected travel time of mode i, $SD(t_i)$ is the standard deviation of travel time, c_i is the ticket fare, σ_i is

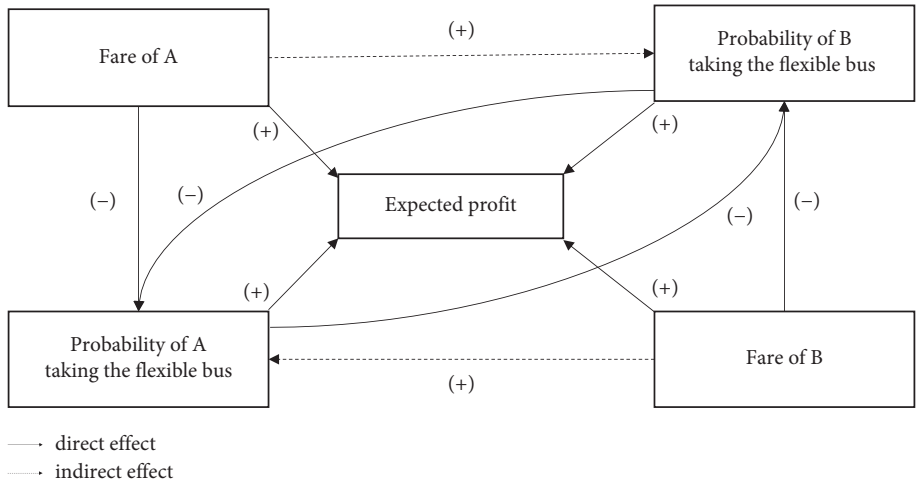


Figure 6: The relationship between A and B.

a dummy variable indicating whether a particular mode-alternative is the flexible bus or not, and $\beta_t, \beta_{SD}, \beta_c$ and β_σ are the parameters. The value of travel time and value of uncertainty can be defined as β_t/β_c and β_{SD}/β_c. We used the data based on the experimental results of stated choice questions to calibrate the parameters in the mean-variance model. The estimation results are $\beta_\sigma = -0.054$, $\beta_t = 0.152$, $\beta_{SD} = 0.018$, $\beta_c = -0.311$. The value of travel time is 0.49 RMB/min and the value of time uncertainty is 0.06 RMB/min. The value of time uncertainty is too small, which is unreasonable. The rate of substitution between CPV(t) and travel cost in our model is 0.88 RMB/min, which is bigger than 0.49 RMB/min. Because our model considers the time uncertainty when calculating the value of travel time are shown in Table 1.

In addition, travel time variability generates utility or disutility depending on the relationship between the reference point. Therefore, the time value and uncertainty perceived by travelers cannot be estimated separately. And according to the above analysis, the subjective perception value of the traveler will affect the utility. Therefore, it is more reasonable to apply CPT to describe the passengers' decision-making behavior in this context.

Table 1: Parameter calibration results of the mean-variance model.

Parameters	β_σ	β_t	β_{SD}	β_c
Coef.	−0.054	0.152	0.018	−0.311
P-value	0.712	0.000	0.615	0.000

Table 2: Passengers' boarding and alighting stops in Case 1 and Case 2.

Stop	Passenger	Stop	Passenger
Case 1			
1	1 (on)	7	1 (off)
2	2 (on), 3 (on)	8	2 (off), 3 (off)
3	4 (on)	9	4 (off)
4	5 (on)	10	5 (off)
5	6 (on), 7 (on)	11	6 (off), 7 (off)
6	8 (on)	12	8 (off)
Case 2			
1	1 (on), 2 (on)	5	3 (off), 6 (off), 7 (off)
2	3 (on)	6	2 (off), 4 (off)
3	4 (on), 5 (on), 6 (on)	7	5 (off)
4	7 (on), 8 (on)	8	1 (off), 8 (off)

Note. 1 (on) means Passenger 1 gets on the flexible bus.

3. Case Study

3.1. Case Description and Parameter Settings. We suppose that there are two small areas, area R is a residential area, and area S is a commercial area. The distance between the two areas is 20 km. There are Q passengers going from R to S, and passengers can choose between flexible bus and taxi. The flexible bus can obtain the demand of all passengers in advance and determines the passengers' boarding and alighting order at the stops. Two cases are set up in this section. Table 2 describes the stops where passengers get on and off the bus. Case 1 adopts the first-on, first-off service principle. In Case 2, there is no specific service principle (random service principle).

The parameters of the two cases are set as follows: s, v, Q, T, μ, φ, and θ are system parameters. We set $s = 20$ km, $v = 30$ km/h, $Q = 8$, $T = 2$ min, $\mu = 3$ RMB/km, $\varphi = 4$ RMB/km, and $\theta = 100$ RMB. λ, γ, β are parameters in CPT. λ, γ are obtained by referring to the research of Tversky and Kahneman. β is calibrated through the SP survey. We set $\lambda = 2.25$, $\gamma = 0.69$, and $\beta = 0.60$. $\beta_0, \beta_t, \beta_c$ are the parameters in the pricing model, which are also calibrated through the SP survey. We set $\beta_0 = -0.326$, $\beta_t = 0.252$, and $\beta_c = -0.285$. In two cases, we omit the average detour time because the average detour time is much smaller than the travel time between regions.

3.2. Results

3.2.1. Case 1. Table 3 shows the fares for each passenger in profit maximization and social welfare maximization

Table 3: Results of case 1.

Passenger		1	2	3	4	5	6	7	8
Detour stops		5	5	5	5	5	5	5	5
Profit maximization model	Fares (RMB)	43.6	43.1	43.1	43.6	43.6	43.1	43.1	43.6
	Maximum profit (RMB)	235.9							
Social welfare maximization model	Fares (RMB)	23.7	21.0	21.3	24.0	23.3	21.3	21.9	23.4
	Maximum social welfare (RMB)	381.9							

models. In addition, the number of stops that affect each passenger's detour is also listed.

In the profit maximization model, the average fare is 43.4, and the std is 0.3. The expected maximum profit is 235.9. The fare of each passenger is almost the same. It is because that the number of detour stops (referring to the detour suffered by the passenger due to serving other stops) is the same. The increase in the fare of passenger 1 will result in a reduction in the probability of Passengers 2–8 taking the flexible bus. It is the same for every passenger. When the individual's fare is set as the average, the expected profit is reduced to 235.8 with negligible differences from the maximum profit.

In the social welfare maximization model, the average fare is 22.5, and the std is 1.2. The maximum social welfare is 381.9 with the constraint that the service provider profit is at least 100. The fare of each passenger is slightly different. When the individual's fare is set as the average, the social welfare is also with negligible differences from the maximum social welfare.

Figure 7 shows that the profit and social welfare (without the constraint that the service provider profit is at least 100 RMB) change with the fare. Here, we set the fare equal for each passenger. As the fare increases, the profit first increases and then decreases, reaching the maximum at around 43 RMB/trip, while the social welfare first remains almost unchanged and then gradually decreased after around 34 RMB/trip.

3.2.2. Case 2. Table 4 shows the fares for each passenger in profit maximization and social welfare maximization models. In addition, the number of stops that affect each passenger's detour is also listed.

In the profit maximization model, the average fare is 45.5, and the std is 2.2. The expected maximum profit is 254.8. In the profit maximization model, the average fare is 22.5, and the std = 2.64. The expected maximum profit is 254.8. Since the number of detour stops for each person is different, the individual fare is also different. Figure 8 shows the relationship between detour stops and fares in Case 2. When the travel distance is certain, the more stops you take, the lower the fare is. It is because the effective distance (the distance from the origin to the destination) is fixed for passengers. Detours caused by serving other stops increase the travel time of the passengers. Therefore, for the flexible bus system, the more detour stops a passenger travels, the lower the fare should be. This is consistent with the calculation results of our model. When an equal fare is used, the expected profit is 249.9, reduced by 1.9% compared to the maximum expected profit, and the social welfare is 388.7. It shows that it is better to take a different fare for each user when aimed at gaining more profits. However, taking a different fare has negligible influence on social welfare.

The travel distance, number of passengers, and other parameters of Case 1 and Case 2 are all the same except for the service order. By comparing Case 1 and Case 2, the maximum profit of Case 2 is 8.0% higher than Case 1, and the maximum social welfare of Case 2 is 1.8% higher than Case 1. This is due to the fact there are more detour stops in Case 1 that affect each passenger than in Case 2. It shows that the detour time has an important influence on the expected profit of the system.

3.3. Effects of the Travel Distance. We set the travel distance to vary from 10 to 40 km while maintaining the other parameters unchanged. Figure 9 shows that as the travel distance increases, the profit and the social welfare per unit distance increase, and the magnitude of the increase decreases. Take the maximum profit model as an example, when the travel distance is 10 km, the maximum profit per unit distance is 7.1 RMB; when the travel distance is 40 km, the profit per unit distance is 15.2 RMB. Compared with the travel distance of 10 km, the profit per unit distance increased by 114.1%, indicating that the flexible bus can obtain more profit in long-distance transportation. The trend of social welfare and profit is consistent. Therefore, the development of long-distance flexible bus services should be a priority strategy for the company and the whole society. Focusing on providing long-distance flexible public transportation services, such as between residential areas, transportation hubs, large shopping malls, and work areas can increase profit and social welfare. With the change of travel distance, the optimal fares per unit distance of the two objectives range from 1.8 to 2.5 and 1.1 to 1.2 respectively. The variation range is small, which indicates that the pricing model has good robustness when travel distance changes.

3.4. Effects of the Average Detour Time. The average detour time is caused by serving other stops for passenger boarding and alighting. We set the average detour time to vary from 1 to 9 minutes for serving one extra stop. Figure 10 shows that as the average detour time increases, the maximum expected profit and social welfare decrease, and the magnitude decreases. Take the maximum profit model as an example, when the average detour time is 1 minute, the maximum profit is 254.2 RMB; when the average detour time is

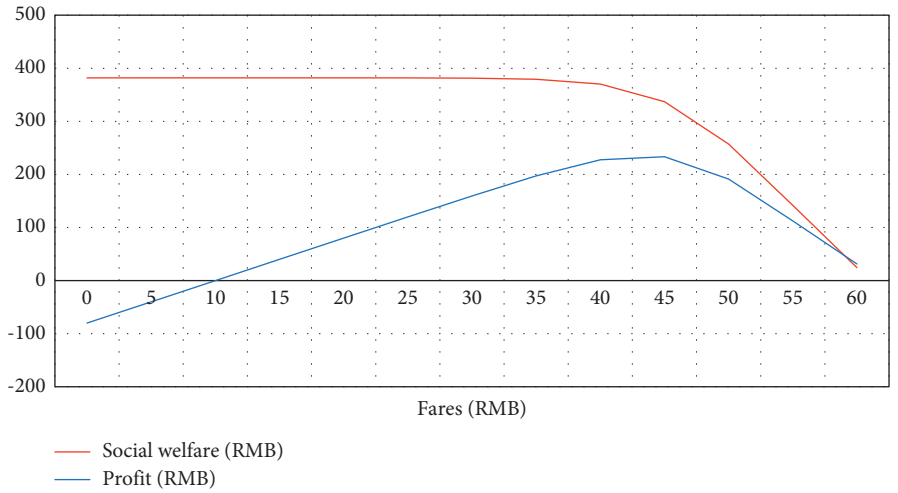


Figure 7: The relationship between fares, profit, and social welfare in Case 1.

Table 4: Results of case 2.

Passenger		1	2	3	4	5	6	7	8
Detour stops		6	4	2	2	3	1	0	3
Profit maximization model	Fares (RMB)	42.0	43.7	45.8	45.7	44.7	47.1	50.0	44.9
	Maximum profit (RMB)	254.8							
Social welfare maximization model	Fares (RMB)	16.6	21.2	24.5	22.4	24.3	23.2	25.9	22.1
	Maximum social welfare (RMB)	388.7							

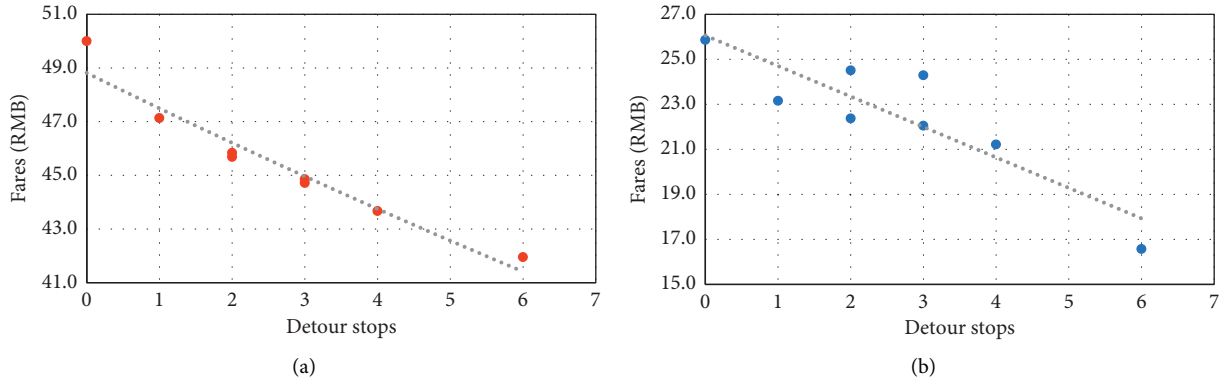


Figure 8: The relationship between detour stops and fares in Case 2. (a) The profit maximization model, (b) The social welfare maximization model.

9 minutes, the maximum profit is 161.1 RMB, substantially reduced by 36.6%. As the average detour time increases, a lower fare is needed to compensate for the delay suffered by passengers, resulting in a decrease in the maximum profit and social welfare. Therefore, the flexible bus company should pay attention to setting an upper limit for the detour time in operation. With the change of average detour time, the optimal fares per unit distance of the two objectives range from 1.7 to 2.3 and 1.1 to 1.2 respectively. This small range indicates that the pricing model has good robustness when average detour time changes.

3.5. Effects of the Value of Time. β_t/β_c measures the value of time, which is related to the individual's attributes. We set β_t/β_c to vary from 0.6 to 1.5 RMB/min while maintaining the other parameters unchanged. Figure 11 shows that as the value of time increases, the maximum expected profit and social welfare decreases, and the magnitude decreases. Take the maximum profit model as an example, when β_t/β_c is 0.4 RMB/minute, the maximum profit is 265.5 RMB; when β_t/β_c is 1.2 RMB/minute, the maximum profit is 217.0 RMB, substantially reduced by 18.3%. As the value of time increases, the operator needs to compensate more for the same

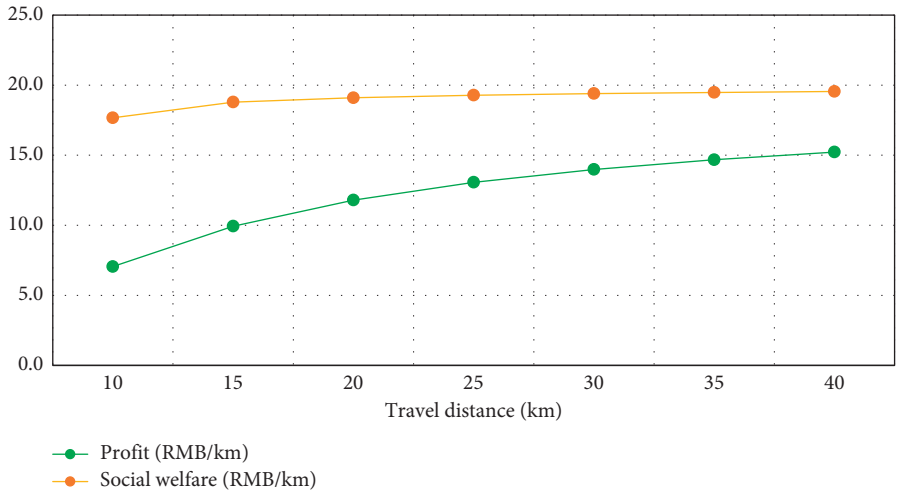


Figure 9: Effects of the travel distance (Case 1).

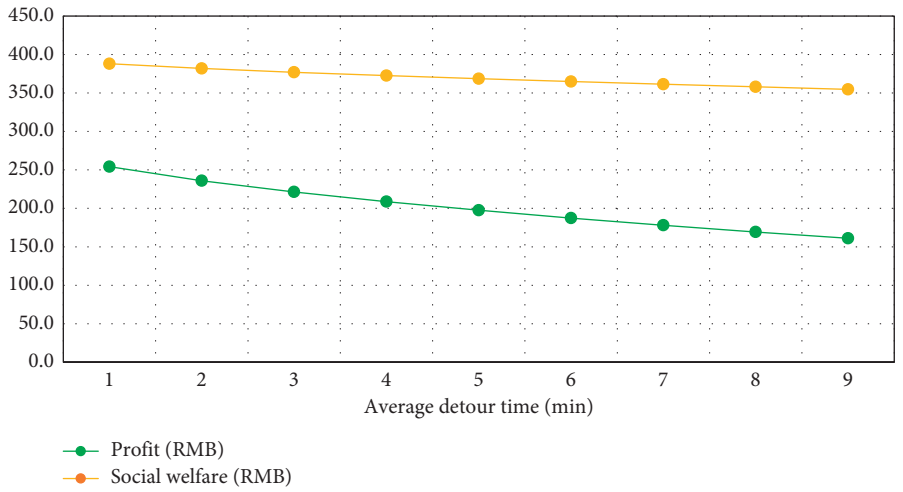


Figure 10: Effects of the average detour time (Case 1).

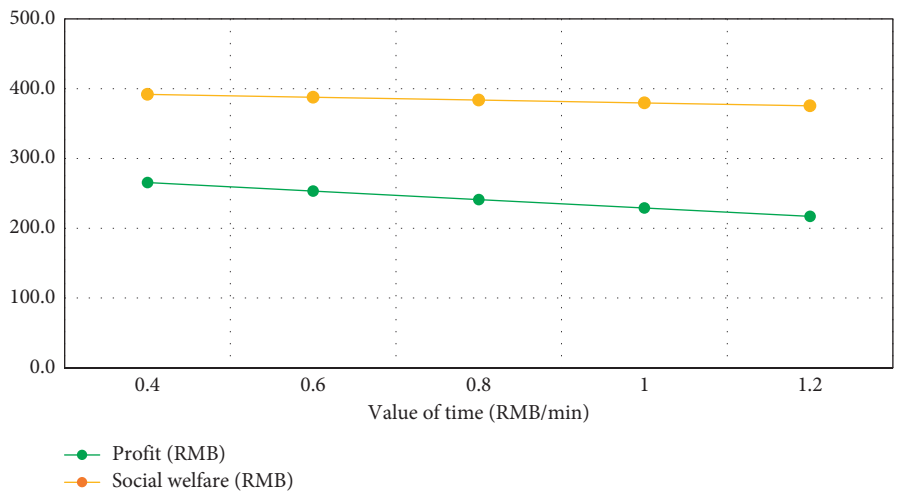


Figure 11: Effects of the value of time.

delay suffered by the passenger whose value of time is higher, resulting in a decrease in the maximum profit and social welfare. Therefore, the flexible bus company should determine the scope of target passengers in combination with the available service level in operation. With the change of the value of time, the optimal fares per unit distance of the two objectives range from 2.0 to 2.3 and 1.1 to 1.5, respectively, indicating good robustness when the value of time changes.

4. Conclusions

With the goals of maximizing the expected profit and maximizing social welfare, the pricing model is constructed based on the cumulative prospect theory for the regional flexible bus service. The parameters of the model were calibrated through a stated preference survey, and two flexible bus cases were designed to analyze the performance of the model. First, we considered the passenger detour in the flexible bus pricing problem and explored the influence of uncertain detour time and fare on passengers' mode choice probability. Second, we modeled the mutual influence among passengers and found that the certain passenger's mode choice probability would affect the detour probability of other passengers. Thus, a calculation method for detour time distribution was proposed.

The results showed that the detour time has a greater impact on the profit and social welfare of the system. More detour stops lead to a lower fare. As the travel distance increases, both the maximum expected profit and social welfare per unit distance increase. As the average detour time increases, the maximum expected profit and social welfare decrease. It is suggested that the flexible bus company may develop long-distance services and set a proper upper limit to the detour time to achieve higher profit. Besides, the passengers' value of time and uncertainty also has a vital influence on the profit and social welfare of the system. The flexible bus company should determine the scope of target passengers in combination with the available service level in operation. With the change of the parameters, the optimal fares per unit distance vary in a small range. The pricing model has good robustness.

There were some limitations to this study. When the area is small and the passengers are densely distributed, a detour time distribution can be calculated. But for few passengers in a large area, there may not be a specific detour time distribution. In addition, the current pricing model can only be used to handle static demand. How to dynamically price the flexible bus needs to be explored further.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The study was supported by the Shanghai Science and Technology Innovation Action Plan Project (20DZ1202805 and 21692110900), National Natural Science Foundation of China (nos. 72101186, 52131204, and 52002280).

References

[1] S. C. Ho, W. Y. Szeto, Y.-H. Kuo, J. M. Y. Leung, M. Petering, and T. W. H. Tou, "A survey of dial-a-ride problems: literature review and recent developments," *Transportation Research Part B: Methodological*, vol. 111, pp. 395–421, 2018.

[2] L. Wang, L. Zeng, W. Ma, and Y. Guo, "Integrating passenger incentives to optimize routing for demand-responsive customized bus systems," *Ieee Access*, vol. 9, pp. 21507–21521, 2021.

[3] Y. Li, X. Li, and S. Zhang, "Optimal pricing of customized bus services and ride-sharing based on a competitive game model," *Omega*, vol. 103, p. 102413, 2021.

[4] H. R. Sayarshad and H. Oliver Gao, "A scalable non-myopic dynamic dial-a-ride and pricing problem for competitive on-demand mobility systems," *Transportation Research Part C: Emerging Technologies*, vol. 91, pp. 192–208, 2018.

[5] I. Kaddoura, G. Leich, and K. Nagel, "The impact of pricing and service area design on the modal shift towards demand responsive transit," *Procedia Computer Science*, vol. 170, pp. 807–812, 2020.

[6] H. Gong, W. Jin, X. Hao, and Q. Luo, "Investigation into pricing game of customized city bus service," *Journal of South China University of Technology*, vol. 45, no. 8, pp. 70–76, 2017.

[7] E. Özkan, "Joint pricing and matching in ride-sharing systems," *European Journal of Operational Research*, vol. 287, no. 3, pp. 1149–1160, 2020.

[8] C. Yan, H. Zhu, N. Korolko, and D. Woodard, "Dynamic pricing and matching in ride hailing platforms," *Naval Research Logistics*, vol. 67, no. 8, pp. 705–724, 2019.

[9] X. Wang and R. Zhang, "Carpool services for ride sharing platforms: price and welfare implications," *Naval Research Logistics*, 2021.

[10] J. Bai, K. C. So, C. S. Tang, X. Chen, and H. Wang, "Coordinating supply and demand on an on-demand service platform with impatient customers," *Manufacturing & Service Operations Management*, vol. 21, no. 3, pp. 556–570, 2019.

[11] Y. Zhong, T. Yang, B. Cao, and T. C. E. Cheng, "On-demand ride-hailing platforms in competition with the taxi industry: pricing strategies and government supervision," *International Journal of Production Economics*, vol. 243, p. 108301, 2022.

[12] J. Ke, H. Yang, X. Li, H. Wang, and J. Ye, "Pricing and equilibrium in on-demand ride-pooling markets," *Transportation Research Part B: Methodological*, vol. 139, pp. 411–431, 2020.

[13] K. Zhang and Y. Nie, "To pool or not to pool: equilibrium, pricing and regulation," *Transportation Research Part B: Methodological*, vol. 151, pp. 59–90, 2021.

[14] T.-M. Choi, S. Guo, N. Liu, and X. Shi, "Optimal pricing in on-demand-service-platform-operations with hired agents and risk-sensitive customers in the blockchain era," *European Journal of Operational Research*, vol. 284, no. 3, pp. 1031–1042, 2020.

[15] C. Wu, S. Le Vine, A. Sivakumar, and J. Polak, "Dynamic pricing of free-floating carsharing networks with sensitivity to travellers' attitudes towards risk," *Transportation*, pp. 1–24, 2021.

[16] D. Kahneman and A. Tversky, "Prospect Theory: an analysis of decision under risk," *Econometrica*, vol. 47, no. 2, pp. 263–291, 1979.

[17] A. Tversky and D. Kahneman, "Advances in prospect theory: cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, vol. 5, no. 4, pp. 297–323, 1992.

[18] K. V. Katsikopoulos, Y. Duse-Anthony, D. L. Fisher, and S. A. Duffy, "Risk attitude reversals in drivers' route choice when range of travel time information is provided," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 44, no. 3, pp. 466–473, 2002.

[19] S. Ghader, A. Darzi, and L. Zhang, "Modeling effects of travel time reliability on mode choice using cumulative prospect theory," *Transportation Research Part C: Emerging Technologies*, vol. 108, pp. 245–254, 2019.

[20] Y. Y. Liu, W. M. Liu, Y. M. Hu, and J. Li, "Pricing model of traffic demand management based on cumulative prospect theory," in *Proceedings of the 2010 8th World Congress on Intelligent Control and Automation*, pp. 5175–5180, Jinan, 7-9 July 2010.

[21] H. Xu, Y. Lou, Y. Yin, and J. Zhou, "A prospect-based user equilibrium model with endogenous reference points and its application in congestion pricing," *Transportation Research Part B: Methodological*, vol. 45, no. 2, pp. 311–328, 2011.

[22] D. Koffman, "Operational experiences with flexible transit services," 2004.

[23] Z. Li, D. A. Hensher, and J. M. Rose, "Willingness to pay for travel time reliability in passenger transport: a review and some new empirical evidence," *Transportation Research Part E: Logistics and Transportation Review*, vol. 46, no. 3, pp. 384–403, 2010.

[24] K. A. Small, "The scheduling of consumer activities: work trips," *The American Economic Review*, vol. 72, no. 3, pp. 467–479, 1982, http://www.scientific.net/SSP.23-24.117.

[25] K. A. Small, R. B. Noland, X. Chu, and D. L. Lewis, *Valuation of Travel-Ime Savings and Predictability in Congested Conditions for Highway User-Cost Estimation*, Transportation Research Board, Washington Dc, 1999.

[26] R. Batley and J. N. Ibáñez, "Randomness in preference orderings, outcomes and attribute tastes: an application to journey time risk," *Journal of Choice Modelling*, vol. 5, no. 3, pp. 157–175, 2012.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 2920532, 18 pages
https://doi.org/10.1155/2022/2920532

Research Article

Pick-Up and Delivery Problem for Sequentially Consolidated Urban Transportation with Mixed and Multi-Puprope Vehicle Fleet

Haoye Chen,[1] **Jonas Hatzenbühler**,[2] **and Erik Jenelius**[2]

[1]*School of Architecture and the Built Environment, KTH Royal Institute of Technology, Teknikringen 74C, Stockholm, 100 44, Sweden*
[2]*Division of Transport Planning, KTH Royal Institute of Technology, Brinellvägen 23, Stockholm, 10044, Sweden*

Correspondence should be addressed to Haoye Chen; haoye@kth.se

Received 7 November 2021; Revised 13 December 2021; Accepted 14 December 2021; Published 11 March 2022

Academic Editor: Lei Wang

Different urban transportation flows (e.g., passenger journeys, freight distribution, and waste management) are conventionally separately handled by corresponding single-purpose vehicles (SVs). The multi-purpose vehicle (MV) is a novel vehicle concept that can enable the sequential sharing of different transportation flows by changing the so-called modules, thus theoretically improving the efficiency of urban transportation through the utilization of higher vehicles. In this study, a variant of the pick-up and delivery problem with time windows is established to describe the sequential sharing problem considering both MVs and SVs with features of multiple depots, partial recharging strategies, and fleet sizing. MVs can change their load modules to carry all item types that can also be carried by SVs. To solve the routing problem, an adaptive large neighborhood search (ALNS) algorithm is developed with new problem-specific heuristics. The proposed ALNS is tested on 15 small-size cases and evaluated using a commercial MIP solver. Results show that the proposed algorithm is time-efficient and able to generate robust and high-quality solutions. We investigate the performance of the ALNS algorithm by analyzing convergence and selection probabilities of the heuristic solution that destroy and repair operators. On 15 large-size instances, we compare results for pure SV, pure MV, and mixed fleets, showing that the introduction of MVs can allow smaller fleet sizes while approximately keeping the same total travel distance as for pure SVs.

1. Introduction

The growing urban transportation brings challenges. Taking Stockholm as an example, it faces the problem of congestion [1] and ensures reliable transport as the number of inhabitants increase [2]. At the same time, Sims et al. [3] pointed out that unless transport emissions can be strongly decoupled from GDP growth, the increasing transport activities could outweigh all mitigation measures to reduce global transport greenhouse gas emissions. As a potential solution, Savelsbergh and Woensel [4] and Los et al. [5] stressed the importance of integrated and/or collaborative transportation.

Transportation demand tends to substantially vary across space and over time. The demand profiles differ depending on the type of transportation. Figure 1 shows an example of unevenly distributed demand of multiple transportation types. Passenger flows peak in the morning and evening, while freight is concentrated during the daytime and recycling demand is solely distributed in the morning.

Passenger, freight, and recycling transportation operations are currently handled by separate fleets of vehicles of distinct types. These single-purpose vehicles (SVs) can transport items of only one type each. When one kind of demand is low, the utilization of corresponding vehicles becomes less efficient as they stand idle or run with lower fill rates, while the vehicle capacity for other demand types may be insufficient. Chen and Li [6] and Liu et al. [7] discussed the similar underutilization of nonmodular

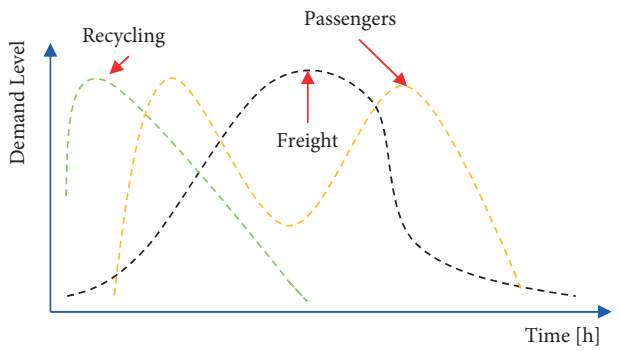


Figure 1: Example demand distribution for passengers, freight, and recycling in an urban context.

vehicles. Each kind of SV is specially designed for a specific demand type and thus cannot efficiently carry items of other types. For example, vehicles for passengers have plenty of seats, which only allow few parcels in the small storage room or on the seats. Some vehicles are not able to carry items of any other types (e.g., using vehicles for waste to carry people).

The concept of multi-purpose vehicles (MVs), which can change their load modules at dedicated change stations to transport items of multiple types while keeping the drive train and other parts, is schematically shown in Figure 2. Given the current rapid development of autonomous vehicles, future MVs may be automated (i.e., no drivers). If not, the vehicle cab for drivers should be located on the fixed component.

The overall cost for operating a fleet including MVs is the fixed cost of vehicles, the variable cost of the total travel distance, and the module change cost at change stations. MVs enable sequential sharing of vehicles, thereby potentially increasing the utilization of vehicles and achieving a lower overall cost by fewer vehicles for the fixed cost and more flexible choices of routes for the variable cost. In the example in Figure 1, parts of the MV fleet may transport passengers and waste, respectively, in the morning peak. Then, MVs can be changed for freight during the daytime. Before the evening peak, the MVs can be changed back to transport people. The MVs utilized for recycling in the morning peak can also transport passengers and freight later. As a result, fewer MVs than SVs are potentially required for urban transportation.

Future urban transportation should be sustainable to cope with the challenges of global warming. Great efforts in developing electric vehicles have been devoted to such ambitions. Therefore, we assume all vehicles are electric in the future. However, even if electric vehicles are environmentally friendly in operation, manufacturing electric vehicles will still cause many emissions. Higher utilization of MVs can theoretically solve this problem to a certain extent by using fewer vehicles.

From the view of MVs, the module change needs extra distance for visiting a change station. Fortunately, if the MVs are electric, this shortcoming will be overcome since increasing the driving distance of electric vehicles does not cause excessive emissions. Therefore, electric MVs can enjoy more flexible route choices and higher utilization without significantly increasing emissions.

Although various forms of shared mobility systems are widely studied (see [8]), little research has been devoted to the sequential sharing problem. In a recently published study by Hatzenbühler et al. [9], the authors propose a pick-up and delivery problem with time windows and sequential sharing. The authors state that the operation of sequentially sharing MVs leads to a reduction in fleet size, while the level of service for passengers and freight requests can be maintained. We will extend it by introducing the mixed fleet including both SVs and MVs to study the transition from the current pure SV fleet to the pure MV fleet or the mixed fleet. The performed numerical experiments will fill the aforementioned research gap by enhancing the understanding of the impacts of MVs in future urban transportation systems.

To understand the potential of MVs, there are two important questions: i) what are the most suitable vehicle fleet characteristics (i.e., a pure MV fleet or a mixed fleet containing both SVs and MVs)? and ii) how to plan routes and fleet configurations (i.e., how many vehicles of each type should be used)? To answer these questions, an adaptive large neighborhood search (ALNS) algorithm is developed, which is able to determine the best routes and fleet configurations. We then compare the best solutions for pure SV operations, pure MV operations, and operations with mixed fleets.

The contributions of this study can be summarized as follows:

(i) To the best of our knowledge, we are the first to consider mixed fleets and MVs in the pick-up and delivery problem and present the mathematical formulation of the problem.

(ii) We integrate the aspects of multiple depots, partial recharging strategies, fleet sizing, and mixed fleets including both MVs and SVs.

(iii) We propose an efficient ALNS algorithm for the proposed problem, which introduces new mechanisms to deal with MVs.

(iv) We validate the performance of the proposed heuristic by comparing its results to the solutions from an exact algorithm.

(v) We show that the mixed fleet may improve the solutions obtained with pure SVs in scenarios with unevenly distributed demand of several types. Both a pure MV fleet and a mixed fleet may lead to fewer vehicles, while the total travel distance is comparable to a SV fleet.

The remainder of the study is organized as follows. Section 2 reviews related work in the literature. Section 3 proposes the problem and gives it the mathematical formulation. Section 4 presents the proposed ALNS algorithm. The results and performances of the ALNS operators are displayed in Section 5 followed by discussions. Section 6

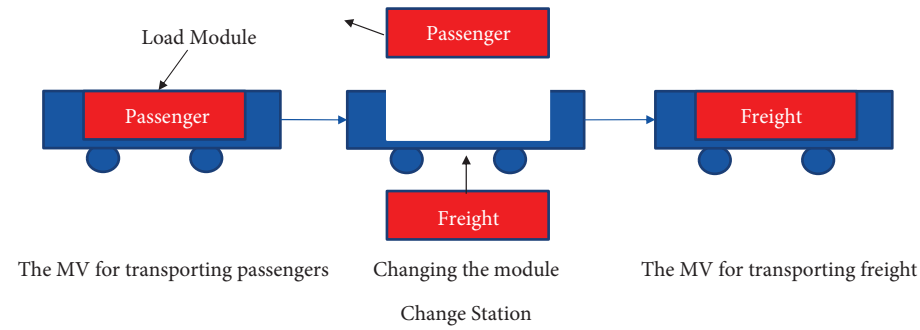


Figure 2: Schematic concept of multi-purpose vehicles.

contains conclusions, future possible applications, and research directions.

2. Literature Review

2.1. Pick-Up and Delivery Problem with Time Windows. One variant of the widely studied vehicle routing problem (VRP) is the PDPTW. In the PDPTW, a set of vehicle routes are determined so that a set of transportation requests are served at a minimum cost with the given vehicle fleet [10]. The PDPTW uses a heterogeneous vehicle fleet stationed at multiple depots to satisfy a set of predetermined transportation requests. Each request consists of a pick-up location, a delivery location, and a certain number of items that should be transported. The pick-up time and delivery time of a request should fall in the time windows of both locations. The objective function may include the fixed cost of vehicles and various variable operational costs (see [11]).

The dial-a-ride problem (DARP) is a variation of the PDPTW in which the transported items are people and thus focuses more on the quality of service and the convenience of passengers [11, 12]. For more comprehensive reviews, we refer the interested reader to articles about PDPTW [13] and DARP [12, 14].

2.2. Combining Different Types of Transportation. In urban transportation, different types of items move in heterogeneous carriers. Most studies on combined transportation discuss the combination of two types of items in the same vehicles, which can be mainly classified into two categories: single-tiered and two-tiered models.

In a single-tiered model, vehicles can simultaneously transport goods and passengers to their delivery points. Li et al. [15] introduced a new kind of model called the share-a-ride problem (SARP) based on the dial-a-ride problem. In the SARP, goods and passengers are transported by taxis in the city.

As for the two-tiered model, the first tier typically describes a scheduled transportation mode such as bus and train, whose spare space can be used to carry goods. In the second tier, smaller vehicles transport goods or passengers to their final destinations. Thus, strict synchronization at the transfer points of each tier is necessary [8].

2.3. Fleet Size and Mixed-Fleet Problem. The fleet size and mixed-fleet (FSM) problem was first proposed by Golden et al. [16] in a VRP study, in which multiple types of vehicles (varying in capacity, range, etc.) are available and the size of the fleet is unlimited. Hiermann et al. [17] introduced the electric fleet size and mix VRP with time windows (E-FSMFTWs) and recharging stations, which considers the choice of recharging times and locations.

2.4. Partial Recharging. Electric vehicles (EVs) have rapidly developed in the last decade. However, EVs suffer problems like long recharging times and insufficient charging station locations. Thus, the optimization of EVs has become an important and timely research topic.

Conrad and Figliozzi [18] considered a recharging vehicle routing problem (GVRP), which allows EVs to be recharged at selected customer locations. Erdoğan and Miller-Hooks [19] proposed a green vehicle routing problem and presented a mathematical formulation with recharging stations. EVs can be fully recharged at recharging stations in a predefined time. However, time window and capacity constraints are not considered in their problem. Schneider et al. [20] introduced the electric vehicle routing problem with time window (E-VRPTW) and recharging stations based on the GVRP, in which the recharging time depends on the remaining electricity when an EV arrives at a recharging station. Time window and capacity constraints are also considered in the proposed E-VRPTW. Grandinetti et al. [21] discussed electric pick-up and delivery problem with time windows, which extends the PDPTW with recharging stations.

2.5. Adaptive Large Neighborhood Search. ALNS, first introduced by Ropke and Pisinger [22], is a heuristic algorithm consisting of competing subheuristics whose selection frequencies are adjusted based on their historical performance. ALNS has been widely used and proven effective for solving the general VRP. Ropke and Pisinger [22] applied ALNS to the PDPTW and improved the best-known solution in benchmark instances. Li et al. [23] designed an ALNS algorithm for the SARP proposed in Li et al. [15] and proved its effectiveness in instances generated from real taxi trails and benchmark instances of DARP. Hiermann et al. [17] solved the E-FSMFTW utilizing branch-and-price and an

adaptive large neighborhood search with an embedded local search and labeling procedure for intensification. The effectiveness of their proposed algorithms is evidenced by a newly created set of benchmark instances for the E-FSMFTW and the existing single vehicle type benchmark using an exact method. Keskin and Çatay [24] proposed an innovative ALNS for the EVRP-PR, which introduces new removal and insertion mechanisms of recharging stations. The computational results of benchmark instances showed that their proposed method is effective and of high quality.

3. Methodology

3.1. Model Description. In this study, we consider a mixed fleet including MVs and several types of SVs, in which the load modules of MVs can be switched at dedicated change stations to transport any item type served by the SVs. The entire vehicle fleet is assumed to be electric. Each vehicle can be either partially or fully recharged at recharging stations. Each capacity-limited vehicle starts from one of the depots and returns to the depot where it originated. Requests, which contain several types of items with pick-up locations, delivery locations, and time windows, are served by the mixed fleet. As an illustration, Figure 3 displays one route by a vehicle. The vehicle serves two passenger requests and recharges the battery at the recharging station. After visiting a change station, it serves one freight request and returns to the same depot where it originated.

We aim to find the best vehicle configuration of the mixed fleet and routes of these vehicles in the network composed of customer points (pick-up and delivery points), recharging stations, change stations, and depots to achieve an overall minimum cost including the fixed cost of vehicles, the variable cost of the total travel distance, and the change cost at change stations. The problem can be regarded as a variant of the pick-up and delivery problem with time windows (PDPTWs) as defined in the literature review. We refer to it as the multi-depot pick-up and delivery problem with time windows, partial recharging strategies, fleet sizing, and mixed fleet of single-purpose vehicles and multi-purpose vehicles.

3.2. Mathematical Problem Formulation. In this section, we present the mathematical model of the proposed problem. The nomenclature and parameters are displayed in Table 1. In this model, all points except depots can be visited only once. Thus, multiple duplicates of change stations and recharging stations are added to the network to allow multiple visits to a station. Each duplicate can be visited once, and the number of duplicates is large enough to cover the maximum number of possible visits at a change or recharging station. To distinguish the arrival time, load, and electricity when a vehicle starts and ends at a depot, we add one copy of depots as the end depots differed from the start depots. Given n_d depots, depot vertices $\{1, \ldots, n_d\}$ denote the start depots and depot vertices $\{n_d + 1, \ldots, 2 * n_d\}$ denote the end depots. MVs must empty their loads before entering a change station.

There are two classes of vehicles. One is MV, and the other one is SV containing three vehicles types. The request of an item type can be handled by an SV or an MV of the corresponding vehicle type. Although MVs can transport items of multiple types at different moments, the type of the MV is defined as the type of item it can carry at a particular moment.

Let N^* denote all points in the network and p denote the class of a vehicle (i.e., 0 for MV and 1 for SV). We define the binary decision variable x^p_{ijkd} as equal to 1 if a vehicle of the class p carries items of type k, travels along the arc (i, j) for $i, j \in N^*$, and originates from the depot d. Otherwise, the variable is 0. Let R denote the request set for all types. We define the binary decision variable z^r_j for tracking the item of request r in the vehicle route. It equals 1 if a vehicle carries the item of request r when arriving at the point j, and 0 otherwise. Given the request $r \in R$, $r^+ \in C$ denotes its pick-up point and $r^- \in C$ denotes its delivery point.

3.3. Objective Formulation. The objective function (1) minimizes the overall costs including the fixed costs of vehicles and variable costs proportional to the distance traveled. The first and second terms denote the fixed costs of SVs and MVs, respectively. The fixed cost of vehicles contains the purchase cost and the labor cost (e.g., driver and packing if any) amortized to each trip. The third and fourth terms represent the variable costs brought by SVs and MVs, respectively. The variable cost contains the cost of energy consumption per distance unit. The fifth term represents change costs, which are proportional to the number of module changes. The meaning of its parameters is displayed in Table 1.

$$\min\Bigg[\sum_{k\in K} f^k \sum_{d\in D}\sum_{i\in D}\sum_{j\in C\cup S'} x^1_{ijkd} + f^0 \sum_{k\in K}\sum_{d\in D}\sum_{i\in D}\sum_{j\in C\cup S'} x^0_{ijkd} + \sum_{k\in K}\alpha^k \sum_{i\in N}\sum_{j\in N'}\sum_{d\in D} d_{ij}x^1_{ijkd} + \alpha^0 \sum_{k\in K}\sum_{i\in N}\sum_{j\in N'}\sum_{d\in D} d_{ij}x^0_{ijkd} + f_c \sum_{i\in F'}\sum_{j\in N'}\sum_{k\in K}\sum_{d\in D} x^0_{ijkd}\Bigg]. \tag{1}$$

3.4. Constraints. The model characteristics are defined in this subsection. The constraints related to depots and the mixed fleet are based on the formulation proposed by Salhi et al. [25], which considers multiple depots and a mixed fleet

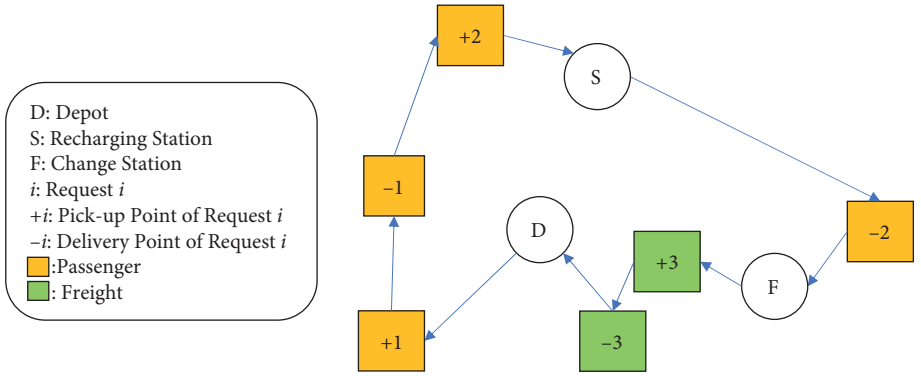


Figure 3: An example of a vehicle route.

Table 1: Sets of points, parameters, and decision variables.

K	Vehicle types (i.e., transported item types)
C_k	Customer points for item type k
C	Customer points for all item types, $C = \cup C_k$
R	Requests, in which a request has r a pick-up point r^+ and a delivery point r^-
C_k^+	Pick-up points for item type k
D	Depots where vehicles start
D'	Depots where vehicles end
F'	Duplicated change stations
S'	Duplicated recharging stations
N	Customer and change points with start depots, $N = C \cup D \cup F' \cup S'$
N'	Customer and change points with end depots, $N' = C \cup D' \cup F' \cup S'$
N^*	All points in the network, $N^* = N \cap N'$
P	Vehicle classes, 1 for SVs, 0 for MVs
f^k	Fixed cost of the SV type k
f^0	Fixed cost of the MV
α^k	Variable cost of the SV type k
α^0	Variable cost of the MV
f_c	Change cost of the MVs at any change station
d_{ij}	The distance/cost for traveling from i to j
$[e_i, l_i]$	Time range for $i \in C \cup D \cup D'$
s_i	Service time for a customer point i.
τ_i	Starting time of a point i when visited by a vehicle at arrival.
E^k	Energy storage capacity for SVs of type k.
E^0	Energy storage capacity for MVs.
g^k	Recharging time per energy unit for vehicle type k.
g^0	Recharging time per energy unit for MVs.
η^k	Energy consumption per unit distance for vehicle type k.
η^0	Energy consumption per unit distance for MVs.
Q^k	Load capacity for SVs of type k.
Q^0	Load capacity for MVs
q_j	Volume of the item at customer point j
M_t	Upper bound of time.
M_y	Upper bound of energy, $M_y = \max\{E^k\}$
M_q	Upper bound of load, $M_q = \max\{Q^k\}$.
x_{ijkd}^p	1, if a vehicle of the class p, carrying items of type k, traveling along arc (i, j) for $i, j \in N^*$, and originating from depot d. 0, otherwise
z_j^r	1, if a vehicle carries the item of request r when arriving at point j. 0, otherwise
y_i	Energy level when a vehicle arrives at point i
Y_i	Energy level when a vehicle leaves the recharging station i
u_i	Remaining capacity at point i when being visited

in a VRP. We extend the formulation by using the variables z_j^c tracking movements of items on vehicle routes and introduce features of MVs, change stations, and recharging stations.

$$\sum_{p\in P}\sum_{i\in N}\sum_{d\in D} x_{ijkd}^{p} = 1, \quad \forall k \in K, \forall j \in C_k^+, \tag{2}$$

$$\sum_{p\in P}\sum_{i\in N}\sum_{d\in D} x_{ijkd}^{p} = 0, \quad \forall k \in K, \forall j \notin C_k^+, \tag{3}$$

$$\sum_{p\in P}\sum_{i\in N}\sum_{k\in K}\sum_{d\in D} x_{ijkd}^{p} \le 1, \quad \forall j \in S', \tag{4}$$

$$\sum_{i\in N}\sum_{k\in K}\sum_{d\in D} x_{ijkd}^{0} \le 1, \quad \forall j \in F', \tag{5}$$

$$\sum_{i\in N}\sum_{k\in K}\sum_{d\in D} x_{ijkd}^{1} = 0, \quad \forall j \in F', \tag{6}$$

$$\sum_{i\in N'}\sum_{k\in K}\sum_{d\in D} x_{jikd}^{1} = 0, \quad \forall j \in F', \tag{7}$$

$$\sum_{i\in D}\sum_{k\in K}\sum_{d\in D} x_{ijkd}^{0} = 0, \quad \forall j \in F', \tag{8}$$

$$\sum_{i\in N} x_{ijkd}^{p} = \sum_{i\in N} x_{jikd}^{p}, \quad \forall p \in P, \forall k \in K, \forall j \in C \cup S', \forall d \in D, \tag{9}$$

$$\sum_{i\in N}\sum_{k\in K} x_{ijkd}^{0} = \sum_{i\in N'}\sum_{k\in K} x_{jikd}^{0}, \quad \forall j \in F', \forall d \in D, \tag{10}$$

$$\sum_{j'\in N'} x_{jj'kd}^{0} - M\left(1 - x_{ijkd}^{0}\right) \le 0, \quad \forall i \in N, \forall j \in F', \forall k \in K, \forall d \in D, \tag{11}$$

$$\sum_{p\in P}\sum_{k\in N}\sum_{d\in D} x_{iikd}^{p} = 0 \quad \forall i \in N^*, \tag{12}$$

$$x_{d_1ikd_2}^{p} = 0, \quad \forall p \in P, \forall i \in C \cup F' \cup S', \forall k \in K, \forall d_1, d_2 \in D, d_1 \ne d_2, \tag{13}$$

$$x_{id_1kd_2}^{p} = 0, \quad \forall p \in P, \forall i \in C \cup F' \cup S', \forall k \in K, \forall d_1 \in D', \forall d_2 \in D, d_1 \ne d_2 + n_d, \tag{14}$$

$$\sum_{p\in P}\sum_{k\in K}\sum_{d\in D} x_{ijkd}^{p} = 0, \quad \forall i, j \in D, \tag{15}$$

$$\sum_{p\in P}\sum_{k\in K}\sum_{d\in D} x_{ijkd}^{p} = 0, \quad \forall i, j \in D', \tag{16}$$

$$\sum_{p\in P}\sum_{k\in K}\sum_{d\in D} x_{ijkd}^{p} = 0, \quad \forall i \in D, \forall j \in D', \tag{17}$$

$$\sum_{p\in P}\sum_{k\in K}\sum_{d\in D} x_{ijkd}^{p} = 0, \quad \forall i \in D', \forall j \in D, \tag{18}$$

$$x_{ijkd}^{p} \in \{0, 1\}, \quad \forall p \in P, \forall i \in N^*, \forall j \in N^*, \forall k \in K, \forall d \in D. \tag{19}$$

Constraints (2) and (3) ensure that each pick-up point is served by a vehicle of the corresponding type. Constraint (4) guarantees that each recharging station may be visited by all vehicles at most once. Constraints (5)-(7) enforce that module change stations can be exclusively visited by MVs. Constraint (8) ensures MVs cannot drive from a depot to a change station. The flow conservation for customer points and recharging stations is handled by Constraint (9). Constraint (10) ensures the conservation of flow for change stations. Constraint (11) guarantees that a vehicle, which visits a change station, must change its vehicle type. M is a sufficiently large parameter for the conditional constraint. Constraint (12) avoids self-circles for all points. Constraints (13) and (14) ensure that vehicles return to the same depot where they depart. Constraints (15)-(18) prevent vehicles directly traveling from depots to depots. Constraint (19) defines the binary decision variable.

The following set of constraints record the carrying status of requests:

$$\sum_{r\in R} z_j^r = 0, \quad \forall j \in D \cup D' \cup F', \tag{20}$$

$$\sum_{p\in P}\sum_{k\in K}\sum_{d\in D} x_{ijkd}^{p} = 1 \longrightarrow z_i^r = z_j^r, \quad \forall i \in N, \forall j \in N, \forall r \in R, i \ne r^+, i \ne r^-, i \ne j, \tag{21}$$

$$\sum_{p\in P}\sum_{k\in K}\sum_{d\in D} x_{r^+jkd}^{p} = 1 \longrightarrow z_j^r = 1, \quad \forall j \in N', \forall r \in R, r^+ \ne j, \tag{22}$$

$$\sum_{p\in P}\sum_{k\in K}\sum_{d\in D} x^{p}_{r^{-}jkd} = 1 \longrightarrow z^{r}_{j} = 0, \quad \forall j \in N', \forall r \in R,\ r^{-} \neq j, \tag{23}$$

$$z^{r}_{r^{-}} = 1, \quad \forall r \in R, \tag{24}$$

$$z^{r}_{j} \in \{0,1\}, \quad \forall j \in N^{*}, \forall r \in R, \tag{25}$$

Constraint (20) defines the initial situation for all start depots and guarantees that vehicles have emptied their loads before visiting end depots and change stations. Constraint (21) passes the carrying status of requests from i to j if they are not picked up or delivered at i. Constraint (22) ensures that vehicles collect items at pick-up points of requests. Constraint (23) guarantees that vehicles drop items at delivery points of requests. Constraint (24) ensures that the delivery point must be served after its corresponding pick-up point. Constraint (25) defines the binary decision variable.

The last set of constraints record the arriving time, energy level, and load. To track the energy usage, the decision variable y_i is defined as the energy level when a vehicle arrives at the point i. The decision variable Y_i is defined as the energy level when a vehicle leaves the recharging station i. The recharged amount of electricity is $(Y_i - y_i)$, which follows the definition first introduced by Keskin and Çatay [24]. The decision variable τ_i records the time when a vehicle visits the point i.

In the proposed model, each type of SV and the MV has a different energy storage capacity, recharging time per energy unit, and energy consumption per unit distance. When expressing parameters of vehicle types, we use the index of $\{1, \ldots, K\}$ representing SV types and 0 representing MVs. The decision variable u_j describes the current remaining capacity at the point j.

$$e_j \le \tau_j \le l_j, \forall j \in C \cup D \cup D', \tag{26}$$

$$\tau_i + \left(t_{ij} + s_i\right)x^{p}_{ijkd} - M_t\left(1 - x^{p}_{ijkd}\right) \le \tau_j, \quad \forall p \in P, \forall i \in C, \forall j \in N', \forall k \in K, \forall d \in D, i \neq j, \tag{27}$$

$$\tau_i + t_{ij}x^{0}_{ijkd} + t_{swap} - M_t\left(1 - x^{0}_{ijkd}\right) \le \tau_j, \quad \forall i \in F', \forall j \in N', \forall k \in K, \forall d \in D, i \neq j, \tag{28}$$

$$\tau_i + t_{ij}x^{1}_{ijkd} + g^{k}\left(Y_i - y_i\right) - M_t\left(1 - x^{1}_{ijkd}\right) \le \tau_j, \quad \forall i \in S', \forall j \in N', \forall k \in K, \forall d \in D, i \neq j, \tag{29}$$

$$\tau_i + t_{ij}x^{0}_{ijkd} + g^{0}\left(Y_i - y_i\right) - M_t\left(1 - x^{0}_{ijkd}\right) \le \tau_j, \quad \forall i \in S', \forall j \in N', \forall k \in K, \forall d \in D, i \neq j, \tag{30}$$

$$0 \le y_j \le Y_j, \quad \forall j \in S', \tag{31}$$

$$0 \le Y_i \le E^{0} + M_y\left(1 - x^{0}_{ijkd}\right), \quad \forall i \in S', \forall j \in N', k \in K, d \in D, i \neq j, \tag{32}$$

$$0 \le Y_i \le E^{k} + M_y\left(1 - x^{1}_{ijkd}\right), \quad \forall i \in S', \forall j \in N', k \in K, d \in D, i \neq j, \tag{33}$$

$$0 \le y_j \le y_i - \left(\eta^{0}d_{ij}\right)x^{0}_{ijkd} + M_y\left(1 - x^{0}_{ijkd}\right), \quad \forall i \in C \cup F', \forall j \in N', k \in K, d \in D, i \neq j, \tag{34}$$

$$0 \le y_j \le y_i - \left(\eta^{k}d_{ij}\right)x^{1}_{ijkd} + M_y\left(1 - x^{1}_{ijkd}\right), \quad \forall i \in C \cup F', \forall j \in N', k \in K, d \in D, i \neq j, \tag{35}$$

$$0 \le y_j \le E^{0} - \left(\eta^{0}d_{ij}\right)x^{0}_{ijkd} + M_y\left(1 - x^{0}_{ijkd}\right), \quad \forall i \in D, \forall j \in N', k \in K, d \in D, i \neq j, \tag{36}$$

$$0 \le y_j \le E^{k} - \left(\eta^{k}d_{ij}\right)x^{1}_{ijkd} + M_y\left(1 - x^{1}_{ijkd}\right), \quad \forall i \in D, \forall j \in N', k \in K, d \in D, i \neq j, \tag{37}$$

$$0 \le u_j \le u_i - q_i x^{p}_{ijkd} + M_q\left(1 - x^{p}_{ijkd}\right), \quad \forall p \in P, \forall k \in K, \forall i \in C \cup F' \cup S', \forall j \in N', d \in D, i \neq j, \tag{38}$$

$$0 \le u_j \le Q^0 + M_q\left(1 - x^0_{ijkd}\right), \quad \forall k \in K, \forall i \in D, \forall j \in N', d \in D, i \ne j, \tag{39}$$

$$0 \le u_j \le Q^k + M_q\left(1 - x^1_{ijkd}\right), \quad \forall k \in K, \forall i \in D, \forall j \in N', d \in D, i \ne j. \tag{40}$$

Constraint (26) defines the time windows of depots and customer points for all types. Constraints (27) and (28) describe the time changes at customer points and change stations, respectively. Constraints (29) and (30) consider the time changes at recharging stations for MVs and SVs, respectively. Constraint (31) ensures that the energy level of a vehicle leaving a recharging station is higher than the energy level of the vehicle when arriving. Constraints (32) and (33) guarantee that the energy levels are no more than the maximum energy capacity of the corresponding vehicle types when leaving the recharging station. Constraints (34) and (35) describe the energy change at recharging stations for MVs and SVs, respectively. Constraints (36) and (37) describe the energy consumption from depots to other points for MVs and SVs, respectively. Constraint (38) describes the load changes at customer points, change stations, and recharging stations. Constraints (39) and (40) describe the load changes when leaving start depots for MVs and SVs, respectively.

4. Adaptive Large Neighbourhood Search

This section details the ALNS algorithm for the proposed problem. Section 4.1 gives an overview of the proposed algorithm. Section 4.2 defines the process to generate initial solutions. Section 4.3 discusses the strategy for sequential sharing. Section 4.4 discusses the strategy for partial recharging. Section 4.5 describes the evaluation of solutions. Sections 4.6, 4.7, 4.8, and 4.9 introduce the customer destroy (CD) operators, customer repair (CR) operators, station destroy (SD) operators, and station repair (SR) operators, respectively.

4.1. Solution Generation. The pseudocode of the proposed ALNS is displayed in Algorithm 1. In the first step, an initial solution is created (see Section 4.2). Then, the algorithm iteratively destroys and repairs the current solution. For each iteration, if the new solution is not energy feasible (i.e., nonfulfillment of energy constraints), the station insertion algorithm Greed (see section 4.9) is applied to get an energy feasible solution.

The new solution is accepted or rejected based on a simulated annealing (SA) approach: Let X_{New} and X_{Current} be the new solution and the current solution, respectively. If $f(X_{\text{New}})$ is less than $f(X_{\text{Current}})$, where f is the objective function of the model plus load penalty E_l and time penalty E_t (defined in Algorithm 2), the new solution is accepted as the current solution. Otherwise, the new solution may still be accepted with a probability of $e^{-(f(X_{\text{New}}) - f(X_{\text{Current}})/T)}$. After each iteration, the temperature T is multiplied by the cooling rate α and the best solution will be updated if the new solution is best so far. The initial temperature guarantees that the first new solution worse than the current solution will be accepted with a probability of 0.5. If $f(X_{\text{New}})$ is less than $f(X_{\text{Current}})$ and if E_l and E_t are zero, the new solution will be updated as the best solution. Infeasible solutions with better f will be accepted as the current solution, which increases the flexibility of ALNS.

The analysis function f contains the objective of the model given the solution plus a load penalty E_l and a time penalty E_t scaled by parameters β_l and β_t, respectively, as shown in Algorithm 2. $E_l = 0$ guarantees the fulfillment of load constraints. $E_t = 0$ ensures the fulfillment of the time windows. The new solution is always energy feasible according to Algorithm 1. Therefore, a penalty for electricity is not needed. The outputs f, E_t, and E_l are utilized in the SA approach.

The ALNS algorithm's structure is based on the algorithm proposed by Keskin and Çatay [24], which is proven useful for recharging stations in the vehicle routing problem. An iteration of operators for charging stations is followed after every $N_{\text{SO}} - 1$ iterations of operators for destroying and repairing requests.

When choosing the destroy operator and repair operator during each iteration, we use a roulette wheel selection. This mechanism is proposed by Ropke and Pisinger [22]. As shown in Algorithm 1, iterations are divided into segments with the lengths of N_c and N_e for customers and recharging, respectively. The probability P_i of selecting an operator i during a segment s depends on its weight in the previous segment W_l^{s-1} (i.e., $P_i^s = W_i^{s-1} / \sum_{l=1}^{m} W_l^{s-1}$). In particular, the weight of the operator i in the segment s is $W_i^s = W_i^{s-1}(1 - r_p) + r_p \pi_i^{s-1} / \vartheta_i^{s-1}$, where r_p is the roulette wheel parameter, ϑ_i^{s-1} represents the usage of the operator i during the segment $s - 1$, and π_i^{s-1} is the score of the operator i during the segment $s - 1$. For each iteration, the score of the chosen operator π_i^s is increased by σ_1, σ_2, or σ_3, which are scores if the iteration generates the best solution so far, improves the current solution, or accepts a worse solution according to the SA criterion.

After each customer/recharging segment, the scores and weights of the customer/recharging operators are reset. In the reset of the segment s, if the operator i is not selected within the segment (i.e., $\pi_i^s = 0$), the weight of this operator in the next segment $s + 1$ will be $W_i^{s+1} = W_i^s(1 - r_p) + r_p \sigma_3 / 3$. The small value $\sigma_3/3$ guarantees that W_i^s does not approach zero so that every operator can be selected even if they have bad performance for a period.

4.2. Initial Solution. In the generation of the initial solution, requests are handled according to their type. Out of all requests from the same type, one request is randomly selected as the base request to generate a vehicle. The start depot is chosen so that f increases the least for this vehicle carrying only the base request. The request is randomly

(1) Generate an initial solution
(2) $j \leftarrow 1$
(3) **while** $j \leq Maximum$ **do**
(4) **if** $j \equiv 0 \pmod{N_{SO}}$ **then**
(5) Select SD operator and remove stations on the current solution
(6) Select SR operator and repair solution to get a new solution
(7) **else**
(8) Select CD operator and remove customer on the current solution
(9) Select CR operator and repair solution to get a new solution
(10) Updating change stations of the new solution
(11) **if** energy infeasible solution **then**
(12) Perform the station insertion algorithm *Greed* on the new solution
(13) **end**
(14) **end**
(15) Apply the sequential sharing strategy
(16) Apply the partial recharging strategy
(17) Algorithm 2: Assignment of vehicles' class and solution evaluation
(18) Using the SA criterion to accept/reject the new solution as the current solution and updating the best solution if it is feasible and best so far
(19) $T \leftarrow \alpha * T$
(20) **if** $j \equiv 0 \pmod{N_c}$ **then**
(21) reset scores and update weights of CD and CR operators
(22) **end**
(23) **if** $j \equiv 0 \pmod{N_e}$ **then**
(24) reset scores and update weigths of SD and SR operators
(25) **end**
(26) $j \leftarrow j + 1$
(27) **end**

Algorithm 1: ALNS.

(1) **Input** new solution X
(2) **Output** $f(s)$, E_t and E_t
(3) **for each** *vehicle in the solution* **do**
(4) **if** requests of different types in the vehicle **then**
(5) Assign the vehicle as a MV
(6) **else**
(7) calculate the objective value Obj_0 contributed by the vehicle if it is multi-purpose
(8) calculate the objective value Obj_1 contributed by the vehicle if it is single-purpose
(9) Assign the vehicle with the class p of lower Obj_p
(10) **end**
(11) If the arriving time of point i: τ_i is earlier than its time-window, the vehicle will wait
(12) If the vehicle load at point i is beyond its load capacity, record the excess load as v_i
(13) **end**
(14) Calculate the load penalty $E_l = \sum_{i \in N^*} v_i$
(15) Calculate the time penalty $E_t = \sum_{i \in C \cup D \cup D'} \max(\tau_i - l_i, 0)$
(16) Sum up the objective value contributed by each vehicle as *Objective*
(17) $f(s) \leftarrow$ Objective $+ \beta_l * E_l + \beta_t * E_t$

Algorithm 2: Assignment of vehicles' class and solution evaluation.

assigned to an MV or SV. Then, the remaining requests will be inserted into the best position while considering the feasibility of time and load. Requests with earlier possible visiting times have priorities in the insertion order. If it is impossible to insert any remaining request into this vehicle, a new vehicle is generated based on one request randomly chosen from the remaining requests. The request insertion and vehicle generation processes are repeated until all remaining requests of this type are inserted. The process is then for all types. Finally, the station insertion algorithm (see Section 4.9) is performed to make the initial solution feasible in energy constraints.

4.3. Strategy for Sequential Sharing. Sequential sharing does not allow vehicles to simultaneously carry different types of items. New change stations may be in need when inserting a request into a vehicle with at least one different type of request. Conversely, existing change stations may be removed when deleting a request whose neighboring points on both sides belong to the same type. After applying destroy and repair operators, the proposed algorithm checks every two successive customer points. Then, if the two customer points on both sides are not of the same type, the change station adding the lowest distance increment is inserted. If the two customer points are of the same type, the change station is removed from the solution.

However, the insertion of new change stations into a vehicle route is performed after the destroy and repair operators have been applied. The insertion of a change station adds additional travel distance to the route, which is initially not considered when applying the destroy and repair operators. Therefore, the distance matrix equivalent as the direct distance plus the possible additional travel distance (see Equation (41)) is used in order to integrate the additional travel distance into the route, hence correctly updating the cost values.

$$d'_{ij} = \begin{cases} \min_{f \in F'} \left[d_{if} + d_{fj} + f_c \right] & \forall i \in C_{k_1}, \forall j \in C_{k_2}, k_1 \neq k_2 \\ d_{ij} & \text{Otherwise} \end{cases} \tag{41}$$

If two points are not of the same type, their equivalent distance is the distance same as the two points as endpoints to pass the nearest change station. Otherwise, their equivalent distance is the same as the direct distance.

4.4. Partial Recharging Strategy. Keskin and Çatay [24] showed that: "if an optimal solution exists such that an EV leaves the depot with its battery partially charged, then the same EV departing from the depot fully charged is also optimal since fully recharging the battery at the depot does not delay the departure time of the EV." Therefore, vehicles start with a fully recharged battery in our proposed algorithm.

We propose a simple and optimal partial recharging strategy that avoids any time window violations. Since the recharging time increases with the energy that is recharged, after obtaining an energy feasible solution, our proposed algorithm sets the recharged amount of energy at each station to a minimum value, which ensures that the EV arrives at the next station/depot with an empty battery.

4.5. Assignment of Vehicles' Class and Solution Evaluation. In the proposed problem, there are K kinds of SVs and one kind of MV. The purpose of a vehicle is important since each kind of vehicle has a different capacity, maximum energy capacity, recharging speed, and energy-consuming speed. In Algorithm 2, the objective function is computed and the purpose of vehicles is changed based on the evaluation result.

4.6. Customer Destroy Operators. The customer destroy operators delete or move customer requests.

(i) Random location: in the current solution, a number γ_c of points are selected including customer points, change stations, and recharging stations. Then, requests containing the selected customer points are removed.

(ii) Path: inspired by Demir et al. [26], a request is randomly selected. Then, stations and requests with at least one point within the sequence between the pick-up point and the delivery point of the preselected request are removed.

(iii) Worst distance: inspired by Keskin and Çatay [24], the request with the $\lceil |R|\mu^{\kappa} \rceil^{\text{th}}$ largest distance increment, calculated as the total distance of the vehicle minus the total distance after the removal, is removed. Here, $|R|$ means the total number of requests, μ is a random number from 0 to 1, and κ is a parameter influencing the randomness.

(iv) Shaw: inspired by Ropke and Pisinger [22], the idea of this operator is to remove similar requests. For two requests $r_1, r_2 \in R$, the similarity is defined as follows:

$$R(r_1, r_2) = \frac{1}{d_{r_1^+ r_2^+} + d_{r_1^- r_2^-} + \lambda\left(\left| \tau_{r_1^+} - \tau_{r_2^+} \right| + \left| \tau_{r_1^-} - \tau_{r_2^-} \right| \right)}, \tag{42}$$

where τ_i is the arriving time at the point i. Equation (42) considers the differences in distances and arriving time between the two pick-up points and the two delivery points. The two requests with maximum similarity are removed.

(v) Random vehicle: a random vehicle is selected and deleted.

(vi) Pair: two requests are selected and removed. The information of their original vehicles (i.e., the vehicle from which each request is removed) is swapped between those vehicles.

(vii) Cross-swap: two vehicles are randomly selected. All points with empty loads, when being visited, are recorded. For each vehicle, the route is broken into two parts based on their one random recorded point. The back half of each vehicle route is connected to the front half of the other vehicle route. Finally, a time feasibility check is conducted from the breaking point to the end depots for the two vehicles, which deletes the requests containing infeasible points until all remaining requests are time feasible.

Table 2: Results of parameter tuning (Obj $*10^4$).

	Default	Tested							
β_t	Penalty for the time infeasibility								
	6	1	5	5.5	6.5	7	**8**	10	20
Obj	2.240 9	2.417 4	2.395 2	2.221 9	2.213 5	2.238 5	**2.195 4**	2.336 5	2.362 4
β_l	Penalty for the load infeasibility								
	100	10	50	**80**	90	120	150	200	300
Obj	2.227	2.217 7	2.204 5	**2.195 2**	2.216 3	2.230 6	2.223	2.211 5	2.199 4
N_{so}	Period to destroy and repair recharging stations								
	10	3	4	5	6	7	**8**	9	11
Obj	2.208 6	2.361 7	2.222 2	2.252 6	2.201 5	2.235 7	**2.192 3**	2.195 1	2.199 8
N_c	Length of customer segment for resetting scores								
	25	10	20	22	27	**30**	32	35	40
Obj	2.243 3	2.301	2.192 2	2.206 4	2.188 3	**2.187 3**	2.208 1	2.224 5	2.204 2
N_e	Length of station segment for resetting scores								
	200	100	150	175	190	210	**220**	250	300
Obj	2.191	2.239 7	2.208 6	2.209 5	2.252 2	2.217 1	**2.175 5**	2.180 6	2.197 7
r_p	Roulette wheel parameter								
	0.25	0.15	**0.2**	0.3					
Obj	2.185 1	2.267 2	**2.181 6**	2.191 1					
κ	Randomness in the *Worst Distance*								
	5	2	3	**4**	6	7	8		
Obj	2.198 0	2.181 8	2.276 8	**2.180 0**	2.182 3	2.245 1	2.206 2		
λ	Parameter for time differences in the Shaw removal								
	2	0.1	1	1.25	**1.5**	2.5	5	10	100
Obj	2.204 8	2.283 8	2.197 4	2.195 7	**2.179 0**	2.189 2	2.211 2	2.201 6	2.248 4
γ_c	Proportion for deleting customers in destroy operators								
	0.2	0.1	0.125	0.15	0.175	0.225	0.25	0.3	
Obj	**2.193 8**	2.23	2.205	2.242 4	2.197 6	2.243 5	2.230 6	2.237 4	
γ_s	Proportion for deleting stations in destroy operators								
	0.2	0.1	**0.15**	0.25					
Obj	2.201 8	2.197 7	**2.193 9**	2.213 4					
$\sigma_{1,2,3}$	Scores in the simulated annealing								
	15,10,5	15,5,10	15,10,13	15,10,8					
Obj	**2.180 7**	2.210 3	2.235 4	2.193 4					

(viii) Swap: two requests are randomly selected, and their pick-up points and delivery points are swapped. Therefore, no requests are deleted.

4.7. Customer Repair Operators. The customer repair operators insert requests deleted in destroy operators. The best position for inserting a request i is defined as the position leading to the lowest distance increment compared to the ones brought by other positions. The distance increment brought by the insertion of request i is defined as follows:

$$\Delta d_i = \begin{cases} d_{\text{prev}(i^+),i^+} + d_{i^+,\text{next}(i^+)} - d_{\text{prev}(i^+),\text{next}(i^+)} + d_{\text{prev}(i^-),i^-} + d_{i^-,\text{next}(i^-)} - d_{\text{prev}(i^-),\text{next}(i^-)}, & i^+ \text{and } i^- \text{are adjacent,} \\ d_{\text{prev}(i^+),i^+} + d_{i^+,i^-} + d_{i^-,\text{next}(i^-)} - d_{\text{prev}(i^+),\text{next}(i^-)}, & \text{Otherwise,} \end{cases} \tag{43}$$

where prev and next mean the previous point and the next point, respectively.

(i) Original vehicle: the insertion order of requests is first randomly shuffled. Then, each request is inserted into its original vehicle where it is removed at the best position. If the request's original vehicle does not exist after destroy operators, it will be inserted into a random vehicle.

(ii) Inter vehicle: the process is similar to the first repair operator. The only difference is that the insertion considers all positions among all vehicles.

(iii) New vehicle: the operator generates vehicles and inserts requests following the same method in the generation of the initial solution.

(iv) Random insertion: the order is randomly disrupted in which vehicles are inserted. Requests are sorted

by their earliest possible visiting time for insertion. Then, each request in the list is inserted by order into the best position within the vehicle while considering the time and load feasibility. If all remaining requests cannot be inserted into the vehicle, the next vehicle will be considered. If there are still remaining requests and all vehicles have been considered, operator *New Vehicle* will be executed for these requests.

4.8. Station Destroy Operators. The station destroy operators are used to eliminate or move recharging stations in the solution, which may improve the overall performance by removing inefficient ones or adjusting their positions. Four operators are implemented. A fixed number γ_s of recharging stations are selected for destroying operators, except for the inefficient visit operator.

(i) Worst usage: following the operator proposed by Keskin and Çatay [24], this operator aims to make use of the battery as much as possible. Thus, recharging stations are sorted by the energy levels when they are visited. γ_s recharging stations are removed with the highest energy level.

(ii) Inefficient visit: we propose a new destroy operator for recharging stations. To be recharged, vehicles need to visit recharging stations, which increases the total travel distance. Thus, avoiding a large increment with a short recharging stop is reasonable. The proposed operator sorts the recharging stations by the ratio of the recharged amount to distance increment by station and counts the number of stations with zero recharging amount as γ_0. Then, the first $\max(\gamma_0, \gamma_s)$ stations are removed.

(iii) Random station: γ_s recharging stations are randomly removed in order to diversify the solution.

(iv) Random walk: γ_s recharging stations are selected and then moved one position forward or backward in their vehicles' routes.

4.9. Station Repair Operators. The station repair operators are utilized to make the solution electric feasible by inserting recharging stations. There are three operators implemented, and all of them are based on Keskin and Çatay [24].

(i) Greed: this operator checks the route from the start depot. Once a point is visited with a negative battery state, a recharging station is inserted into the arc between this point and its previous point, which brings the least distance increment. If the insertion at this arc is not feasible, the previous arc is considered.

(ii) Comparison: this operator checks the first point visited with a negative battery state and looks at arcs before this point. If the nearest feasible arc for insertion has an infeasible previous arc or a feasible previous arc for insertion of more distance increments, a recharging station with the least distance increment is inserted into the arc. Otherwise, the previous arc is considered in the same way.

(iii) Best arc: once a point is visited with a negative battery state, all arcs between the point and the previous station or depot are considered. The least distance increment brought by the insertions of a recharging station at all feasible arcs is calculated. The station at the arc generating the shortest distance increment is inserted.

To avoid the special case that there is no feasible arc between the last station/depot to the point visited with a negative battery state, we assume a relatively large amount of vehicle energy capacity in the experiments.

5. Numerical Experiments and Analysis

In order to validate our proposed ALNS algorithm, 15 small instances are solved using ALNS and the solver Gurobi 7.5.2. The results of both methods for each instance are presented. The problems are solved using an 8-core i7-8700 CPU of a 3.20 GHz computer. In addition to the numerical results, we record and analyze the selection probabilities of operators and the objective values of one optimization run. Finally, 15 large-size cases are tested to compare the results by pure SV, pure MV, or mixed fleets.

5.1. Parameter Tuning. Inspired by the tuning approach of Ropke and Pisinger [22], Demir et al. [26], and Keskin and Çatay [24] for testing each setting of a parameter, the proposed ALNS is executed 20 times for only 1,500 iterations. Then, the average objective value of 20 trials is recorded. Finally, the setting with the best average objective value is chosen as the parameter for the numerical experiments in the remainder of this study. Only one parameter is varied at a time while keeping the others as default values if they are not tuned or tuned values otherwise. We tuned parameters according to the order in Table 2. An artificial instance of nine requests with three types is used in tuning. Table 2 displays the details of the parameter tuning results. The best parameter values are indicated in bold. We set a high cooling rate α of 0.995 to make the first 1,500 iterations converge fast while manually tuning the cooling rate as 0.999 5 for experiments to make the program result in a good performance in the entire iterations.

5.2. Small-Size Instances. We have manually created 15 instances. Table 3 displays general information about the number of requests, the number of recharging stations, depots, and change stations for each instance. As for the request vector, the number in position i of the vector represents how many requests are there for each type i. The instances have different numbers of recharging stations, change stations, depots, requests, and request types to ensure that the instances represent a varied set of possible scenarios. The fixed costs of SVs and the MV are randomly chosen from the range [2000, 4000], which enables our program to tradeoff between shorten routes and fewer vehicles. The fixed

Table 3: Results for small instances.

Instance	Requests	$[n_S, n_d, n_F]$	Avg.t_l (s)	Avg.t_a (s)	t_e (s)	Avg.dev	Optimum
1	[3,1]	[2,1,1]	0.316 807	2.112 883	105.22	0.00%	5
2	[3,1]	[2,1,1]	0.194 065	2.003 296	269.6	0.00%	5
3	[3,1]	[1,2,1]	0.322 63	1.768 072	5.16	0.00%	5
4	[3,2]	[1,2,1]	0.178 373	2.017 891	30.42	0.00%	5
5	[3,2]	[1,1,2]	0.513 943	1.901 372	4.09	0.00%	5
6	[3,1]	[2,2,1]	0.254 505	2.347 111	2787.43	0.00%	5
7	[3,1]	[3,1,1]	0.517 603	2.162 484	2887.55	0.71%	3
8	[2,2]	[1,2,1]	0.104 138	1.809 524	2.94	0.00%	5
9	[3,3]	[1,1,2]	0.572 434	2.186 973	21.37	0.00%	5
10	[3,1]	[1,3,1]	0.968 397	1.630 676	180.1	0.57%	0
11	[3,2]	[1,2,2]	0.708 596	1.738 11	341.12	0.00%	5
12	[2,2,2]	[1,2,1]	0.549 16	1.611 34	19.06	0.00%	5
13	[2,1,3]	[1,2,2]	0.659 06	1.597 949	267.50	0.00%	5
14	[2,2,3]	[1,2,2]	1.496 54	2.440 7	7200+[1]	0.01%	4
15	[2,2,3]	[1,2,2]	2.158 38	3.102 671	2481.43	0.34%	2

[1]The objective bound limit was not reached within 2 h. The objective value reported is the last value computed by the MIP solver.

Table 4: Time window settings of base instances (pick-up ⟶ delivery, hour).

Type 1	Type 2	Type 3
[0,10] ⟶ [1,12]	Half requests: [5,14] ⟶ [6,15] Half requests: [12,24] ⟶ [12,24]	[6,18] ⟶ [6,18]

cost of MV can be higher than all, some, or none of the modeled SV types. The positions, time windows, loads of customer points, and capacity of vehicles are also varied from instance to instance to increase their diversity.

Each instance is solved by the proposed ALNS algorithm five times and the MIP solver one time. The maximum number of iterations of ALNS is set to 10,000 in Case 1-13, 15,000 in Case 14, and 20,000 in Case 15. In Table 3, n_S, n_d, n_F are the numbers of recharging stations, depots, and change stations, respectively. Avg.t_l means the time to find the best solution. Avg.t_a means the time to finish all iterations. t_e denotes the time obtaining the optimal solution by the exact algorithm. Avg.Dev denotes the average percentage of the result obtained by the ALNS algorithm that deviates from the optimal solution obtained by the MIP solver. Optimum records how many times the ALNS reaches the optimal solution within 5 trials.

As Table 3 shows, the Avg.t_l of cases 1-13 is less than 1 second, which indicates that the proposed algorithm can reach a good solution fast. The Avg.t_a of cases 1-15 except the last case is less than 2.5 seconds, which suggests that the proposed ALNS normally has stable and fast computation time under the same stopping conditions. Comparing the Avg.t_l and the Avg.t_a with the t_e for all cases, hence, we could show that the proposed ALNS is efficient to solve the problems.

At the same time, there is at most a 0.71% average deviation for all the 15 cases. The ALNS reaches the optimum solution for all 5 trials in 11 out of 15 cases. Although the ALNS is stuck in a locally optimal solution in Case 10, the deviation is small. Therefore, the quality of solutions generated by the ALNS is high.

5.3. Operator Analysis. To analyze the operators implemented in this study, we run the proposed algorithm on instance L90-3 (see detailed instance description in 5.4). As mentioned in Section 4, iterations are divided into fixed-length segments and the probability for choosing one operator at one segment depends on its performance on the previous segment. Figures 4 and 5 display the objective curves and selection probabilities at all segments for customer and station operators, respectively. Components of the stacked bar at a segment represent the probabilities of operators at this segment.

Looking at the customer destroy operators in Figure 4(a), before the objective starts to decrease, the probabilities of the operators Cross-swap and Random Vehicle increase, while the probabilities of operators Swap, Pair, and Shaw decrease. At the same time, the operators Worst Distance and Path keep their probabilities. The probability of the operator Random Location increases first, then decreases, and later increases again. With the decrease in objective value, the probabilities of the operators Worst Distance, Path, and Cross-swap dominate while the probability of the operator Random Vehicle continuously decreases. Additionally, the operators Random Location, Shaw, Pair, and Swap remain at low probabilities throughout the entire optimization, indicating poor optimization performance.

Considering the customer repair operators in Figure 4(b), Original Vehicle and Inter Vehicle have low probabilities throughout the entire optimization, indicating poor possibilities to improve the objective value. Before the drop of the objective, the algorithm tries to generate lots of new vehicles proven by the increasing probability of the operator New Vehicle while the operator Random Insertion keeps a high probability at the same time. After the objective value reduces, the operator Random Insertion dominates and fewer new vehicles are generated.

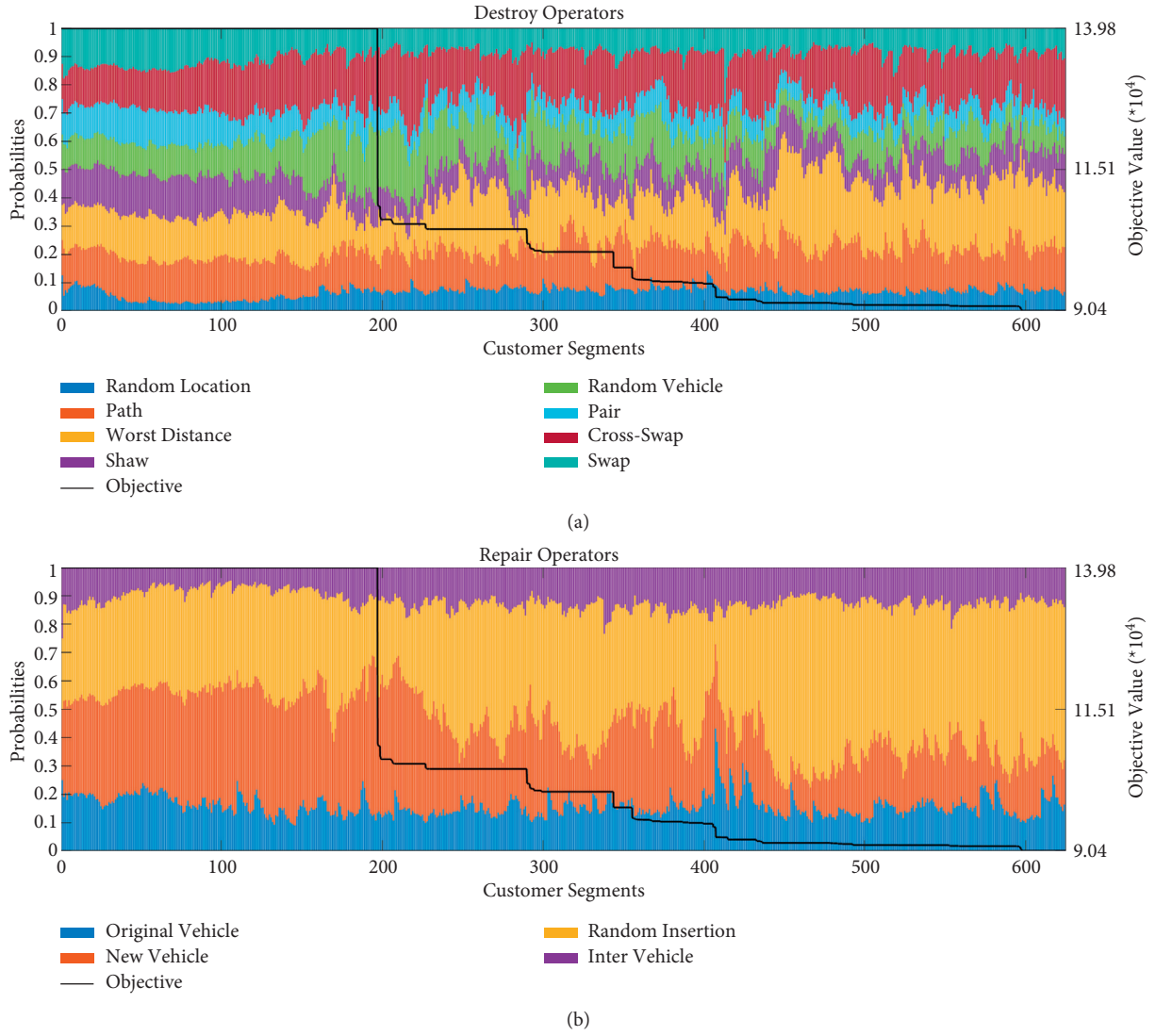


Figure 4: The probabilities of customer operators at each segment. (a) Customer destroy operators. (b) Customer repair operators.

Before the decrease in the objective in Figure 5(a), the probability of the station destroy operator Random Walk continuously decreases while the other three operators share approximately equal probabilities. After the objective starts to decline, the operator Inefficient Visit dominates while the other three operators fluctuate at relatively lower probabilities. Toward the end of the optimization, all four operators share roughly equal probabilities.

In Figure 5(b), the three implemented station repair operators have approximately equal probabilities before the objective starts to decrease. Subsequently, the operators oscillate for the majority of the optimization until returning to an equal share.

In ALNS, we apply both conventional operators and problem-specific operators. Among the conventional operators, Worst Distance and Path are frequently chosen during the optimization, suggesting superior performance. The problem-specific operators Cross-swap, Random Insertion, and Inefficient Visit show good performance, while Pair performs worst according to the reported selection probabilities.

5.4. Large-Size Cases. To evaluate the effectiveness of the consolidated urban transport by MVs given requests with uneven time distribution, we create three base instances. For requests of each type in all base instances, their time windows are concentrated in specific time periods during the day to simulate demands of different types, which are unevenly distributed in time. These three base instances have a total number of requests of 30, 50, and 90 each. The total number of requests is divided into three item types. Base instance L30 has 9, 14, and 7 requests, respectively, for three item types, while base instance L50 has 17, 16, and 17 requests, respectively. Base instance L90 contains 30 requests for each type. Three base instances have the same distribution of time windows as shown in Table 4.

In Table 4, the left side of an arrow denotes a time window of the pick-up points. The right side of an arrow denotes a time window of the delivery points. The time windows for type 1 are concentrated around the first half of a day. Time windows for type 2 have two kinds representing

Figure 5: The probabilities of station operators at each segment. (a) Station destroy operators. (b) Station repair operators.

morning and evening peaks of passengers. The time windows for type 3 cover the whole daytime. Requests of different types need to be served in different periods (i.e., unevenly distributed time windows for different demand types).

To study the performance of pure SVs, pure MVs, and the mixed vehicle fleet, the proposed program can be slightly modified to obtain solutions for pure MVs and pure SVs besides the mixed vehicle fleet. To consider pure SV fleets, a large enough value may be assigned as the fixed cost of MVs so that no MV will be utilized. For pure MV fleets, all vehicles may be assigned as MVs in Algorithm 2.

At the same time, the MV's fixed cost f^0 is an important factor for decision-makers to consider the fleet configuration. For simplicity, we assume that all SV types have the same fixed cost $f^* = 2000$. Furthermore, we create two different situations for MVs based on their fixed costs. One situation assumes that MVs have equal fixed cost f^* as SVs and the other situation assumes that MVs have a slightly higher fixed cost of 1.2 $f^* = 2400$. To control other factors, all SVs and MVs have the same parameters of recharging performance, energy capacity, and volume capacity.

As Table 5 and Figure 6 show, the three base instances are computed with different vehicle fleet characteristics and MV cost structures. Each instance is computed 3 times using 50,000 iterations, and the solution with the overall best objective is displayed. The program uses the same parameter settings as shown in Table 2 except larger N_c of 80, which extends the length of resetting customer operator scores to accommodate longer maximum iteration and more requests. We record the total travel distance, objective value, and vehicle configuration for each instance.

In instances with the same fixed cost structure (i.e., L30-1 to L30-3, L50-1 to L50-3, and L90-1 to L90-3), MVs and mixed vehicle fleets can improve the objective value by 5-15% compared to SVs. At the same time, fewer vehicles are utilized in pure MV and mixed vehicle fleet solutions. The introduction of MVs (i.e., both mixed fleet and pure MVs) implies fewer vehicles and lower objectives if the same fixed costs are assumed.

In instances L30-4 to L30-5, L50-4 to L50-5, and L90-4 to L90-5, the improvement of the objective value in L30-4 is much smaller than those in L30-2, L30-3, and L30-5. The pure MV fleet even has a higher objective value than the SV

Table 5: Results of large-scale instances.

Instance	Fleet	f^0	Distance		Objective	
			Value	Percentage	Value	Percentage
L30-1	SV	Inf	11 498.47	100.00%	27 498.47	100.00%
L30-2	MV	f^*	11 405.70	99.19%	23 705.70	86.21%
L30-3	Mixed	f^*	11 247.01	97.81%	23 547.01	85.63%
L30-4	MV	1.2 f^*	11 247.01	97.81%	25 947.01	94.36%
L30-5	Mixed	1.2 f^*	11 070.83	96.28%	24 070.83	87.54%
L50-1	SV	Inf	21 975.75	100.00%	51 975.75	100.00%
L50-2	MV	f^*	21 640.93	98.48%	47 840.93	92.04%
L50-3	Mixed	f^*	21 606.89	98.32%	47 906.89	92.17%
L50-4	MV	1.2 f^*	22 985.32	104.59%	52 285.32	100.60%
L50-5	Mixed	1.2 f^*	21 872.38	99.53%	48 872.38	94.03%
L90-1	SV	Inf	41 421.75	100.00%	95 421.75	100.00%
L90-2	MV	f^*	42 029.33	101.47%	90 629.33	94.98%
L90-3	Mixed	f^*	41 057.86	99.12%	89 557.86	93.85%
L90-4	MV	1.2 f^*	42 169.33	101.80%	100 669.33	105.50%
L90-5	Mixed	1.2 f^*	42 598.40	102.84%	93 598.40	98.09%

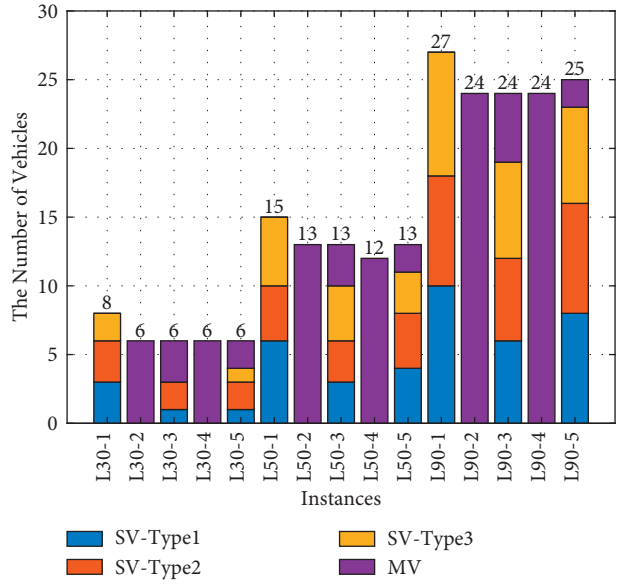


Figure 6: Vehicle configurations of large-scale instances.

fleet in L50 and L90. However, the mixed vehicle fleet can still improve the objective value compared to the SV fleet when the MV has a slightly higher fixed cost than SVs. This is surprising, but a potential explanation for the increased objective values for large-scale instances with pure MV fleets and increased fixed costs (i.e., L50-4 and L90-4) can be the underlying demand pattern. The more the customer requests are in a certain area, the less the time and space gaps of the customer points are. Hence, the more evenly distributed the requests are within that area. Since the operational benefit of MVs mainly stems from the efficient consolidation of unevenly distributed demand in time and/or space, the higher demand density potentially neutralizes this operational advantage of MVs. However, the total number of vehicles for both fleets, mixed and MV, can be reduced compared to SVs.

As for the total travel distance, there are no significant trends regarding the number of requests, fleet characteristics, and MV cost structures given the results. As a potential explanation, fewer vehicles require more travel distance although MVs enable more flexible routes. Therefore, fewer or more total distance depends on the adversarial relationship between more flexible routes and fewer vehicles. However, compared with pure SVs (i.e., L30-1, L50-1, and L90-1), the total distances of other instances have not changed more than 5%.

In the analyzed instances, it can be seen that MV and mixed vehicle fleets utilize fewer vehicles to serve the requests. This is true for scenarios with an equal or a higher fixed cost for MVs than for SVs. The introduction of MVs (i.e., both mixed fleet and pure MVs) brings improvements in terms of the objective value and fleet

size, if MVs have the equivalent fixed cost as SVs. When increasing the fixed costs of MV by 20% compared to SV costs, the objective value of pure MVs rises or decreases less than the situation of equal fixed costs. However, the mixed vehicle fleet can still reduce the objective value in such a situation. At the same time, MVs do not significantly change the total distance. Namely, the MVs improve the objective mainly by fewer vehicles.

6. Conclusions and Future Works

This study investigates MVs in the pick-up and delivery problem with features of multiple depots, mixed fleet, and partial recharging strategies. The mathematical formulations of the problem and an efficient ALNS algorithm are proposed. Additionally, we propose new heuristic operators and strategies in order to accommodate the problem-specific characteristics induced by change stations and partial recharging.

In several numerical experiments, we show that the proposed ALNS can find solutions for small-scale cases with high qualities in an efficient time. We display probabilities' variation of proposed operators over entire iterations in one run and analyze it combining the objective curve. Through large-scale instances, the mixed fleet can reduce the total cost compared to SVs in scenarios where MVs have the same or higher fixed cost compared to SVs. MVs can lead to smaller fleet sizes of pure MVs or mixed fleets but have little positive or negative influence on the total travel distance. The improvement of the objective brought by MVs mainly derives from fewer vehicles.

Solution algorithms to the problem are not limited to our proposed ALNS. Some of the most representative computational intelligence algorithms are promising, such as ant colony optimization [27, 28], artificial bee colony algorithm [29, 30], monarch butterfly optimization [31, 32], and Harris Hawks optimization [33].

The MV is a promising direction for a more efficient urban transport system by enabling fewer fleet size while approximately keeping the total travel distance. When transitioning from the current fleet of SVs to a fleet of MVs, there will be a long period that SVs and MVs coexist. At the same time, SVs may have comparative advantages in some aspects. For example, some SVs are efficient in transporting items of a certain type. Therefore, the coexistence of SVs and MVs is still possible from a long-term perspective. The proposed model can help decision-makers in the transition phase and/or the coexistence of SVs and MVs. The mixed vehicle fleet analysis has shown that the transition toward MV is beneficial since the coexistence of SVs and MVs brings a reduction in the objective value and fleet size. Hence, the transportation system could be already improved in the transition phase. Pure MVs may also be the future trend if they have an even lower fixed cost compared to SVs.

To increase the practical applicability and generality of the model, a future research direction is to further investigate extra large-scale situations using real demand datasets. Another future research direction is the addition of dynamic demand considerations to the proposed problem so that decision-makers can dynamically accept or delete requests.

As for the aspect of MVs, different options of MV operations can be considered in future studies. Namely, there may be MVs with heterogeneous capacities, speed, or other vehicle-related parameters. Besides, MVs do not have to be able to transport items of all types. For example, some MVs may only be switched between passengers and freight, while some vehicles may only be switched between freights and recycling. Flexible changes between two types may be more feasible in a shorter time period and may already lead to benefits.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] J. Eliasson, *The stockholm Congestion Charges: An Overview*, KTH Royal Institute of Technology location, Stockholm, Sweden, 2014.

[2] D. Firth, *City of stockholm, the City of stockholm Traffic Administration*, KTH Royal Institute of Technology location, Stockholm, Sweden, 2012.

[3] R. Sims, R. Schaeffer, F. Creutzig et al., "Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change," Cambridge University Press, Cambridge, UK, IPCC 2014, 2014.

[4] M. Savelsbergh and T. Van Woensel, "Fiftyth anniversary invited article-city logistics: challenges and opportunities," *Transportation Science*, vol. 50, no. 2, pp. 579–590, 2016.

[5] J. Los, F. Schulte, M. T. J. Spaan, and R. R. Negenborn, "The value of information sharing for platform-based collaborative vehicle routing," *Transportation Research Part E: Logistics and Transportation Review*, vol. 141, Article ID 102011, 2020.

[6] Z. Chen and X. Li, "Designing corridor systems with modular autonomous vehicles enabling station-wise docking: discrete modeling method," *Transportation Research Part E: Logistics and Transportation Review*, vol. 152, Article ID 102388, 2021.

[7] T. Liu, A. Ceder, and A. Rau, "Using deficit function to determine the minimum fleet size of an autonomous modular public transit system," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2674, no. 11, pp. 532–541, 2020.

[8] A. Mourad, J. Puchinger, and C. Chu, "A survey of models and algorithms for optimizing shared mobility," *Transportation Research Part B: Methodological*, vol. 123, pp. 323–346, 2019.

[9] J. Hatzenbühler, E. Jenelius, G. Gidófalvi, and O. Cats, "Multi-purpose vehicle assignment for combined passenger and freight transport," Accepted for Presentation: Transportation Research Board (TRB) 101, Washington, DC, USA, 2022.

[10] P. Toth and D. Vigo, "MOS-SIAM Series on Optimization," *Vehicle Routing: Problems, Methods, and Applications*, Society for Industrial and Applied Mathematics : Mathematical Optimization Society, Philadelphia, PA, USA, 2nd edition, 2014.

[11] J. Desrosiers, Y. Dumas, M. M. Solomon, and F. Soumis, "Chapter 2 Time constrained routing and scheduling," in *Handbooks in Operations Research and Management Science*vol. 8, , pp. 35–139, Elsevier, 1995.

[12] K. F. Doerner and J.-J. Salazar-González, "Chapter 7: pickup-and-delivery problems for people transportation," *Vehicle Routing*, Society for Industrial and Applied Mathematics : Mathematical Optimization Society, Philadelphia, PA, USA, pp. 193–212, 2014.

[13] M. Battarra, J.-F. Cordeau, and M. Iori, "Chapter 6: pickup-and-delivery problems for goods transportation," *Vehicle Routing*, Society for Industrial and Applied Mathematics : Mathematical Optimization Society, Philadelphia, Pennsylvania, PA, USA, pp. 161–191, 2014.

[14] J.-F. Cordeau and G. Laporte, "The dial-a-ride problem: models and algorithms," *Annals of Operations Research*, vol. 153, no. 1, pp. 29–46, 2007.

[15] B. Li, D. Krushinsky, H. A. Reijers, and T. van Woensel, "The Share-a-Ride Problem: people and parcels sharing taxis," *European Journal of Operational Research*, vol. 238, no. 1, pp. 31–40, 2014.

[16] B. Golden, A. Assad, L. Levy, and F. Gheysens, "The fleet size and mix vehicle routing problem," *Computers & Operations Research*, vol. 11, no. 1, pp. 49–66, 1984.

[17] G. Hiermann, J. Puchinger, S. Ropke, and R. F. Hartl, "The electric fleet size and mix vehicle routing problem with time windows and recharging stations," *European Journal of Operational Research*, vol. 252, no. 3, pp. 995–1018, 2016.

[18] R. G. Conrad and M. A. Figliozzi, "The recharging vehicle routing problem," in *Proceedings of the Sixty first Annual IIE Conference*, vol. 8, Nevada, NV, USA, May 2011.

[19] S. Erdoğan and E. Miller-Hooks, "A green vehicle routing problem," *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 1, pp. 100–114, 2012.

[20] M. Schneider, A. Stenger, and D. Goeke, "The electric vehicle-routing problem with time windows and recharging stations," *Transportation Science*, vol. 48, no. 4, pp. 500–520, 2014.

[21] L. Grandinetti, F. Guerriero, F. Pezzella, and O. Pisacane, "A pick-up and delivery problem with time windows by electric vehicles," *International Journal of Productivity and Quality Management*, vol. 18, no. 2/3, p. 403, 2016.

[22] S. Ropke and D. Pisinger, "An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows," *Transportation Science*, vol. 40, no. 4, pp. 455–472, 2006.

[23] B. Li, D. Krushinsky, T. Van Woensel, and H. A. Reijers, "An adaptive large neighborhood search heuristic for the share-a-ride problem," *Computers & Operations Research*, vol. 66, pp. 170–180, 2016.

[24] M. Keskin and B. Çatay, "Partial recharge strategies for the electric vehicle routing problem with time windows," *Transportation Research Part C: Emerging Technologies*, vol. 65, pp. 111–127, 2016.

[25] S. Salhi, A. Imran, and N. A. Wassan, "The multi-depot vehicle routing problem with heterogeneous vehicle fleet: formulation and a variable neighborhood search implementation," *Computers & Operations Research*, vol. 52, pp. 315–325, 2014.

[26] E. Demir, T. Bektaş, and G. Laporte, "An adaptive large neighborhood search heuristic for the Pollution-Routing Problem," *European Journal of Operational Research*, vol. 223, no. 2, pp. 346–359, 2012.

[27] M. Dorigo, V. Maniezzo, and A. Colorni, "Ant system: optimization by a colony of cooperating agents," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 26, no. 1, pp. 29–41, 1996.

[28] B. Yu and Z. Z. Yang, "An ant colony optimization model: the period vehicle routing problem with time windows," *Transportation Research Part E: Logistics and Transportation Review*, vol. 47, no. 2, pp. 166–181, 2011.

[29] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Technical report-tr06, Erciyes university, engineering faculty, computer engineering department, Kayseri, Turkey, 2005.

[30] D. Karaboga and B. Basturk, "On the performance of artificial bee colony (abc) algorithm," *Applied Soft Computing*, vol. 8, no. 1, pp. 687–697, 2008.

[31] G.-G. Wang, S. Deb, and Z. Cui, "Monarch butterfly optimization," *Neural Computing & Applications*, vol. 31, no. 7, pp. 1995–2014, 2015.

[32] Y. Feng, S. Deb, G.-G. Wang, and A. H. Alavi, "Monarch butterfly optimization: a comprehensive review," *Expert Systems with Applications*, vol. 168, Article ID 114418, 2021.

[33] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: algorithm and applications," *Future Generation Computer Systems*, vol. 97, pp. 849–872, 2019.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 3815306, 12 pages
https://doi.org/10.1155/2022/3815306

Research Article

A Practical and Economical Ultra-wideband Base Station Placement Approach for Indoor Autonomous Driving Systems

Shengchuan Jiang,[1] **Cong Zhao**,[1] **Yifan Zhu**,[1] **Chenwei Wang**,[1] **and Yuchuan Du**[1,2]

[1]*Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 201804, China*
[2]*Shanghai Engineering Research Center of Urban Infrastructure Renewal, Shanghai 200032, China*

Correspondence should be addressed to Cong Zhao; zhc@tongji.edu.cn

Received 17 December 2021; Revised 18 January 2022; Accepted 31 January 2022; Published 3 March 2022

Academic Editor: Lei Wang

Automated valet parking (AVP) has attracted much attention as the entry point to autonomous driving. In an indoor environment, high-precision positioning systems are essential for AVP. Ultra-wideband (UWB) is one of the most widely adopted techniques. However, the base station placement significantly influences the system's positioning accuracy, especially for the irregular architecture of underground parking lots. This article proposes a three-stage practical and economical layout planning approach for UWB base stations, including determining the deployment strategy and layout parameters and comprehensive adjustment and scheme verification. The approach considers regional differentiation accuracy requirements for AVP, such as ramp area, surface fluctuation area, and narrow area. The adopted positioning method of a UWB system is the time difference of arrival (TDOA), and the evaluation index of positioning accuracy is the horizontal dilution of precision (HDOP). Through experimental tests in an actual parking lot, the proposed approach is confirmed to ensure stability and economy with fewer UWB base stations and can meet the positioning accuracy requirements of AVP.

1. Introduction

With the rapid development of the Internet of Things (IoT) [1], advanced sensors [2], and connected and automated vehicles [3], the demand for high-accuracy location-based services has exploded. Among these, smart indoor parking lots (mostly indoor) are typical integrated applications of the IoT, sensors, and location services, improving the use of parking resources and user satisfaction, especially with the advent of automated valet parking (AVP) [4]. AVP allows human drivers to leave their vehicles in a drop-off zone (e.g., the elevator entrance of a parking lot), and vehicles execute the parking tasks independently [5]. During AVP, vehicles must obtain their location for path planning, decision-making, and control. The quality of task execution strongly depends on positioning accuracy. However, for an indoor parking lot, infrastructure-enabled high-precision positioning is one of the most economical and reliable schemes in existing AVP technologies [6].

Many indoor positioning systems are suitable for smart parking lots, such as Wi-Fi [7], lidar [6], LED light communication [8], and ultra-wideband (UWB) [9]. Among these, UWB has been considered one of the key technologies of next-generation wireless communication because of its low power, anti-interference, strong penetration ability, and high positioning accuracy [9]. However, a UWB positioning system also depends on the beacon positioning method, which must use the known location information of several UWB base stations in the positioning area to determine the mobile station's location. The layout of UWB base stations strongly impacts the positioning accuracy and stability and affects the overall construction cost. Existing studies have confirmed that the layout of base stations influences positioning accuracy more than the measurement noise and fast fading effect [10, 11]. Therefore, optimizing the base station layout for different physical environments and

accuracy requirements are highly significant to realizing the extensive application of UWB positioning systems.

In an existing study on base station layout planning, Monica and Ferrari [12] proposed an optimization method for a UWB base station layout for large indoor scenarios, which deduced the optimal layout spacing concerning corridor width and base station height and verified that its mean squared error (MSE) could reach the Cramer–Rao lower bound (CRLB) through simulation. Long et al. [13] proposed a new base station layout design method to minimize the MSE in circular, square, and hexagonal location areas. Qiao et al. [14] proposed a maximum distance measurement variation criterion-based method to optimize the layout of base stations when adding new base stations to the network. Redondi and Amaldi [15] studied an indoor positioning system's base station layout optimization problem based on wireless sensor signal strength, formulated as mixed integer nonlinear programming to minimize CRLB. A tabu search algorithm was applied to solve the problem. Zhou et al. [16] proposed four CRLB-based metrics to achieve the optimal base station layout and evaluation for indoor rectangular industrial positioning facilities. Sharma and Badarla [17] proposed a scheme to place the positioning base stations on the wall and ceiling, greatly reducing the occlusion during signal propagation and minimizing the positioning error caused by the geometric shape. Meanwhile, they proposed a multiobjective optimization approach to analyze the beacon layout problem (BLP) for the 3D coordinate point cloud representation of indoor environments. The BLP was solved using a nondominating sorting genetic algorithm (NSGA)-II [18]. Kim and Choi [19] introduced a local optimal layout planning method for UWB base stations, formulated as binary integer linear programming with heuristics-based constraints. The experimental results showed that the method could guarantee high-precision and reliable positioning performance for navigating autonomous systems. Wu et al. [20] regarded multiple UWB base stations as a cluster and applied dynamic particle swarm optimization and a genetic algorithm to optimize the layout of base station clusters. Pan et al. [21] proposed a UWB base station layout design method based on a genetic heuristic differential evolution algorithm, which could reduce the average localization error by 28.2% and 12.5% compared to the random and default schemes.

However, some studies aim at reducing the number of base stations. Santoro et al. [22] provided a detailed analysis of the uncertainty of the positioning system while the UWB infrastructure grows, and developed a genetic algorithm that minimizes the deployment of new base stations. Wang et al. [23] aimed at achieving full coverage of indoor scenarios with the least number of base stations deployed. The solution reduced the number of required base stations by 6% to 23% based on a greedy and random sampling algorithm. Balac et al. [24] studied the positioning problem based on fault-tolerant triangulation and proposed several algorithms to optimize the number of base stations. However, the optimization results reduced the number of anchor nodes by less than 0.5%. Leune et al. [25] placed base stations by combining genetic optimization with radio impulse propagation simulation and equipped fewer base stations to provide lower root mean square deviation positioning results in complex environments. In [26], genetic algorithm was applied to determine the location of sensors in intelligent buildings combined with the design drawings of buildings, which effectively reduces the number of sensors. A new three-base station UWB positioning algorithm based on TOF was proposed in [27], which can get rid of the strict placement restrictions of the three base stations without sacrificing the positioning speed excessively.

Although much progress has been made in base station layout planning of indoor positioning systems, room remains for improvement. Most studies do not consider economic issues in base station layout planning. Meanwhile, when confronted with a real situation, especially for the actual complex environment, some algorithms converge slowly, and many restrictions exist. Moreover, most studies have not considered the spatial features of different positioning areas in the indoor parking lot, the different error sources, and differences in accuracy requirements, which bring additional challenges. This article proposes and verifies an AVP application-oriented optimal and economical layout planning method for UWB base stations. The contributions of this article can be summarized as follows:

(i) We analyze several different standard polygon coverage schemes of UWB base stations and propose the most economical coverage scheme

(ii) We develop a three-stage UWB base station placement planning approach, which is applicable to all kinds of indoor parking lots with high practicability and economy

(iii) We design the layout of UWB base stations and conduct experiments in a real underground parking lot to verify the effectiveness of the three-stage method

The organization of this article is as follows: firstly, a two-dimensional UWB positioning system based on TDOA is chosen as the research object. The plane geometric accuracy factor HDOP is selected as the evaluation index of positioning accuracy. Based on this, the accuracy measurement experiments of the UWB positioning system are designed and performed. The different base station layout styles are studied via theoretical analysis and MATLAB simulation. The base station layout in various real scenarios is then simulated, and a method of adjustment and optimization is proposed. Meanwhile, various practical scenario layout methods are integrated to form a practical application-oriented one. Finally, the method is discussed and summarized based on the test results.

2. Assumptions and Accuracy Criteria

We use the following assumptions throughout the article to simplify the model while keeping it realistic.

2.1. Environmental Conditions. The actual environmental scenarios considered in this article include areas where base

stations cannot be evenly laid out, ramp areas, undulating surface areas, narrow and long areas, and subareas with different accuracy requirements.

2.2. Simplification of UWB Positioning System. In this article, we mainly study the impact of UWB base station layout on positioning accuracy. Therefore, the influence of clock synchronization error between base stations under the TDOA algorithm is not considered. Simultaneously, this article does not consider the aging or damage of UWB base stations over time. In AVP scenarios, the vehicle height is not critical because the vehicle must be moving close to the ground. Therefore, the positioning in AVP scenarios can be simplified to two-dimensional plane positioning.

2.3. Accuracy Evaluation Criteria. It is necessary to define and summarize the evaluation indexes of positioning accuracy to evaluate the performance of a positioning system correctly and effectively under different base station layouts.

The geometric dilution of precision (GDOP) is one of the most important indexes to measure the positioning accuracy of a UWB indoor positioning system. This is defined as the ratio between the system positioning error and ranging error, which indicates the geometric spatial layout between a mobile station (MS) and base station (BS) and the magnification of ranging error by the degree of clock synchronization between devices. The GDOP can be calculated as follows:

$$GDOP = \frac{\sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \sigma_{c\cdot\delta t}^2}}{\sigma}, \tag{1}$$

where σ_x^2, σ_y^2, and σ_z^2 represent the error variances in the x, y, and z directions, respectively; $\sigma_{c\cdot\delta t}^2$ represents the error variance due to clock synchronization error; and σ represents the standard deviation of the ranging error. Based on the above assumptions, the horizontal dilution of precision (HDOP) of a two-dimensional plane can be obtained as follows:

$$HDOP = \frac{\sqrt{\sigma_x^2 + \sigma_y^2}}{\sigma}. \tag{2}$$

In this article, the HDOP is applied as the accuracy evaluation index to analyze the two-dimensional positioning accuracy of the UWB positioning system under different BS layouts. Assuming the same ranging error, the smaller the HDOP, the higher the plane positioning accuracy obtained.

2.4. Accuracy Verification. We apply the positioning values of the total station coordinate measurement as true values and use the 1,000 times UWB positioning values as measured values. The Euclidean distances of the measured and true values in the x and y directions are then calculated. The total Euclidean distance and its variance for the measured and true values are also obtained. Finally, we take the four values as the judging index of the positioning performance.

3. Methodology

3.1. Geometry and Coverage Scheme of UWB BS Layout. The number of BSs directly impacts the positioning accuracy. Therefore, we explore the numerical relationship between the number of BSs and the positioning accuracy.

3.1.1. The Relationship between the Number of BSs and the Theoretical Minimum Value of the HDOP. According to the definition of the HDOP [28], the HDOP of UWB positioning based on the TDOA principle meets

$$HDOP \geq \sqrt{\frac{1}{\sum_{i=1}^{n} a_{xi}^2} + \frac{1}{\sum_{i=1}^{n} a_{yi}^2}}, \tag{3}$$

where n is the number of BSs participating in the UWB positioning system; and a_{yi} are the cosine and sine of the angle between the connection of the i-th BS and the positioning UWB label and horizontal direction. Let $a_{xi} = \cos\alpha_i$ and $a_{yi} = \sin\alpha_i$. (3) can be written as

$$HDOP \geq \sqrt{\frac{n}{\left(\sum_{i=1}^{n} \cos^2\alpha_i\right) \times \left(\sum_{i=1}^{n} \sin^2\alpha_i\right)}} = \sqrt{\frac{n}{f(n)}}. \tag{4}$$

Equation (5) can be proved by mathematical induction:

$$f(n) = \frac{n^2}{4} - \frac{1}{4}\left(\sum_{i=1}^{n} \cos 2\alpha_i\right)^2. \tag{5}$$

First, when $n = 2$,

$$f(2) = 1 - \frac{1}{4}\left(\cos 2\alpha_1 + \cos 2\alpha_2\right). \tag{6}$$

When $n = N\,(N \geq 3)$ and (5) is satisfied for $n = N - 1$,

$$\begin{aligned} f(N) &= \left(\sum_{i=1}^{N} \cos^2\alpha_i\right) \times \left(\sum_{i=1}^{N} \sin^2\alpha_i\right) \\ &= f(N-1) + \sin^2\alpha_N \sum_{i=1}^{N-1} \cos^2\alpha_i + \cos^2\alpha_N \sum_{i=1}^{N-1} \sin^2\alpha_i + \cos^2\alpha_N \cdot \sin^2\alpha_N, \\ &= \frac{(N-1)^2}{4} - \frac{1}{4}\left(\sum_{i=1}^{N-1} \cos^2\alpha_i\right)^2 - \frac{1}{2}\sum_{i=1}^{N-1}\left(1 - \cos^2\alpha_i\cos^2\alpha_N\right) + \frac{1}{4}\left(1 - \cos^2 2\alpha_N\right), \\ &= \frac{N^2}{4} - \frac{1}{4}\left(\sum_{i=1}^{N} \cos^2\alpha_i\right)^2. \end{aligned} \tag{7}$$

According to the above, (5) is proved. According to (4), the theoretical lower bound of the HDOP can be calculated as follows:

$$HDOP \geq \sqrt{\frac{4}{n}}. \tag{8}$$

The theoretical lower bound of the HDOP decreases as the number of BSs involved in the system increases. However, when the number of BSs exceeds five, the decrease slows significantly.

3.1.2. Simulation and Analysis of the Impact of the Number of BSs on the HDOP. Considering the installation convenience, UWB BSs are generally arranged in triangles or quadrangles. By simulating the HDOP of each point in the positioning area in MATLAB, the contour map of the location accuracy characteristics in the area when a different number of BSs participate in the calculation is obtained, as shown in Figure 1. Based on the simulation results, in ensuring the accuracy and stability of positioning, four to five times the coverage ensures sufficient positioning accuracy and high economy.

3.1.3. BS Coverage Scheme considering Economy. Based on the construction cost, the premise is to ensure that any point in the location area meets K times the coverage, and K has been proved to be 4–5 as discussed above. The fewer the BSs used to cover the same area, the more economical is the layout scheme.

The BS layout of the UWB positioning system relates to the problem of area coverage, which can be divided into regular triangle and quadrilateral structures. As stated in [29], a square structure has a higher efficiency of four times the coverage layout, and a triangle structure has a higher efficiency of three times the coverage layout. If the effective coverage radius of the UWB tag's pulse signal is 30 m, the number of BSs needed to achieve three times the coverage of the triangle structure, four times the coverage of the square structure, and five times the coverage of the square structure in different areas is shown in Figure 2. Four times the coverage of the square structure is the best coverage and the best economical solution. Minimal difference exists between the number of BSs required to achieve three and four times the coverage in the same area. Conversely, the number of BSs required to achieve five times the coverage is almost twice the number of BSs required to achieve four times the coverage. However, as mentioned earlier, the positioning accuracy achieved by five and four times the coverage is almost the same. Therefore, the UWB BS layouts for the actual environment are all optimized based on four times the coverage of the square structure.

3.1.4. Layout Scheme Adjustment. In practice, the scenario space is often irregular. Thus, it is impossible to realize the uniform layout of BSs. The BSs mostly adopt a rectangular or diamond layout to ensure the coverage of four signals. The HDOP of the positioning area with different side-length ratios is simulated using MATLAB; the results are shown in Table 1.

We draw the following conclusions according to the above results. Based on the positioning accuracy requirements in [30], 90% of the positioning error standard deviation should be within 10 cm. We deduce that the 90% grading value of the HDOP must be less than 2.0, and the ratio of the long side to the short side should not exceed 2.6 for the rectangular layout. For the diamond layout, the ratio of the long axis to the short axis should not exceed 2.6.

3.1.5. UWB BS Layout in Ramp Area. When the tag to be measured moves on the ramp area, its height is constantly changing. The BS horizontal spacing must then be adjusted according to the ramp gradient. That is, the effective coverage radius of the BS must be adjusted during the calculation.

Assuming that θ is the angle between the slope and the horizontal plane, the slope direction is that in which the

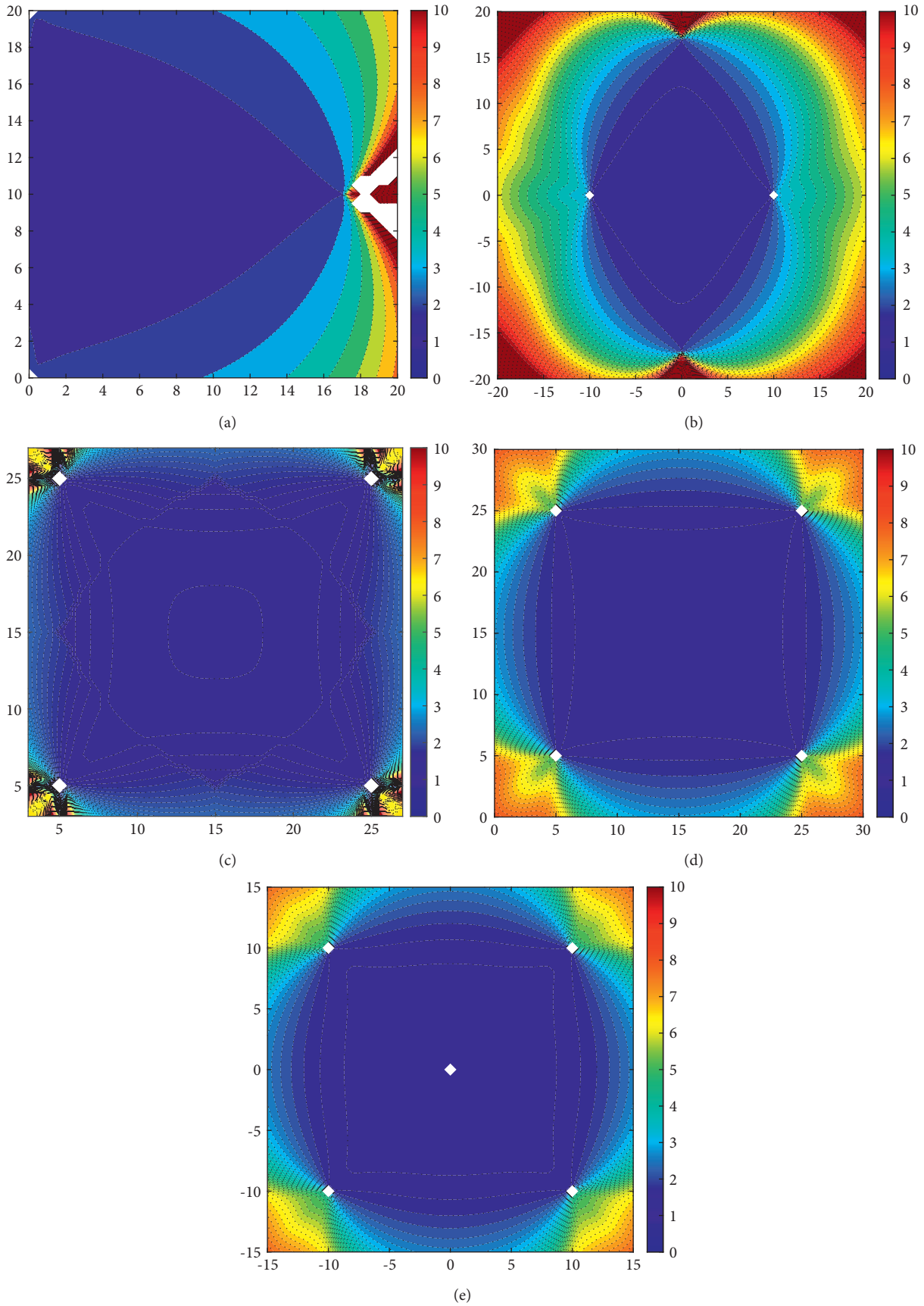


Figure 1: HDOP distribution map of positioning area for different numbers of BSs in a uniform layout; (a) triangle three times coverage; (b) triangle four times coverage; (c) square three times coverage; (d) square four times coverage; (e) square five times coverage.

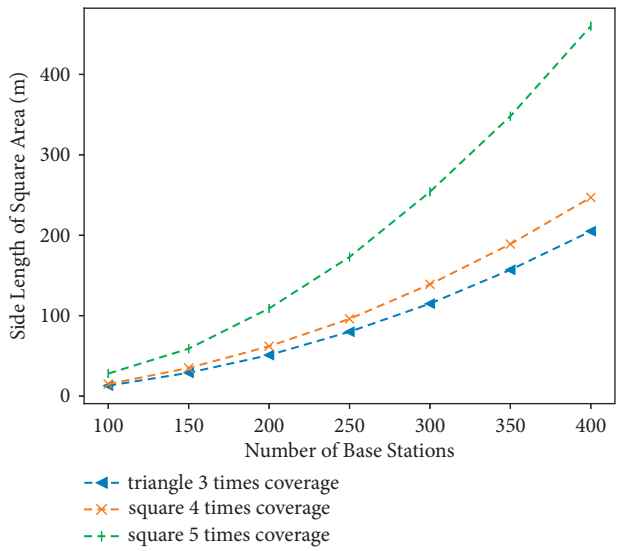


Figure 2: Number of BSs required for different coverages in the same area.

elevation on the slope decreases fastest. The slope area studied in this article is the ideal slope. That is, the slope and aspect of each point on the slope are the same. When the orientation of interest is consistent with the gradient direction, the effective radius r' is

$$r' = r\cos\theta. \tag{9}$$

3.1.6. UWB BS Layout in Surface Fluctuation Area. In the actual positioning scenario, the height of the tag to be tested will inevitably change. After deduction, the positioning error caused by the change in tag height is as follows:

$$G_h = \frac{1}{\sin\alpha} \cdot \sqrt{\|\Delta n_1\|^2 + \|\Delta n_2\|^2 - 2\cos\alpha\|\Delta n_1\| \cdot \|\Delta n_2\|}, \tag{10}$$

$$\Delta r_{io} = \left(\sqrt{r_i^2 + h^2} - r_i\right) - \left(\sqrt{r_o^2 + h^2} - r_o\right), \quad i = 1, 2, \tag{11}$$

$$\nabla \mathbf{r}_{i0} = \begin{bmatrix} \dfrac{\partial r_{i0}}{\partial x} \\ \dfrac{\partial r_{i0}}{\partial y} \end{bmatrix}, \mathrm{u}_i = \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}, \tag{12}$$

where Δn_i denotes the projection of u_i on ∇r_{i0}; (x, y) denotes the measured coordinate value; ∇r_{i0} denotes the gradient vector of r_{i0} at the point to be measured; r_{i0} denotes the distance difference between BSs; and α denotes the angle between Δn_1 and Δn_2. Taking the tag height estimation error h as a variable, the positioning error G_h caused by h is simulated and analyzed. The results are shown in Figure 3.

As h increases, the average positioning error caused by the error in tag height estimation gradually increases, and the maximum error will increase rapidly. When the surface fluctuation is less than 0.3 m, the influence will be less for the BS layout covered by four times the square structural signals (Max $G_h < 10$ cm). Therefore, the influence of the elevation difference can be neglected when it is less than 0.3 m. However, when the elevation difference exceeds 0.3 m, the different elevation areas should be divided into subareas.

3.1.7. UWB BS Layout in Narrow Area. For narrow areas, such as lanes and roadways, the length of one direction is much smaller than that of the other. This can be formulated as

$$HDOP(x, y) = \sqrt{G_{11} + G_{22}}, \tag{13}$$

where G_{11} and G_{22} denote the direction cosines of the tags in the x and y directions, namely, the precision factors of the x and y directions, respectively. Considering the difference in location accuracy between two directions in the narrow region, we evaluate the location accuracy of the BS layout in the narrow region using the direction-weighted plane accuracy factor, that is, through pairing and giving different weights in different directions. Meanwhile, the error variance of the positioning accuracy in two directions is applied to determine the corresponding weight coefficients:

$$w_x = 1 - \frac{\sigma_x}{\sigma_x + \sigma_y}, \tag{14}$$

$$w_y = 1 - \frac{\sigma_y}{\sigma_x + \sigma_y}, \tag{15}$$

$$HDOP_w(x, y) = \sqrt{w_x \cdot G_{11} + w_y \cdot G_{22}}. \tag{16}$$

Assuming that the width of a narrow area is 3 m and the length is 30 m, 90% of the area to be measured should satisfy the positioning error variance along the width direction (x direction) and the length direction (y direction) when the tag moves in the narrow area. The surface of the narrow area is planar, and the plane coverage radius of the BS is 50 m. The genetic algorithm toolbox of MATLAB is applied to solve the optimization problem of weighted geometric factors. The simulation results are shown in Figure 4.

The distance between adjacent BSs in the y direction should not exceed 10.5 m to ensure that 90% of the positioning area meets the direction accuracy requirement. The rectangular aspect ratio can then reach $10.5/3 = 3.5$ m, exceeding the previous requirement of 2.6 m. When the absolute size of the σ_x and σ_y changes, or the width of the narrow area changes, the distance will also change.

3.1.8. UWB BS Layout considering the Difference in Regional Precision Requirements for AVP. Under actual working conditions, different location areas may have different location accuracy requirements for AVP. The weighted region HDOP model is then formulated to obtain an economical

Table 1: Distribution characteristics of the HDOP in rectangular and diamond layouts with different axial ratios.

Layout	Axial ratio	Minimum HDOP	Maximum HDOP	Average HDOP	80% grading value HDOP	90% grading value HDOP
Rectangular	1.0	1.285	1.816	1.455	1.426	1.582
	1.4	1.289	1.860	1.485	1.451	1.584
	1.8	1.293	1.976	1.551	1.537	1.688
	2.2	1.296	2.125	1.636	1.601	1.792
	2.6	1.296	2.299	1.735	1.729	1.946
	3.0	1.297	2.492	1.844	1.844	2.136
Diamond	1.0	1.285	1.651	1.410	1.510	1.547
	1.4	1.285	1.761	1.450	1.577	1.640
	1.8	1.285	1.981	1.494	1.647	1.725
	2.2	1.285	2.244	1.549	1.723	1.871
	2.6	1.285	2.555	1.617	1.839	2.058
	3.0	1.285	2.878	1.680	1.960	2.226

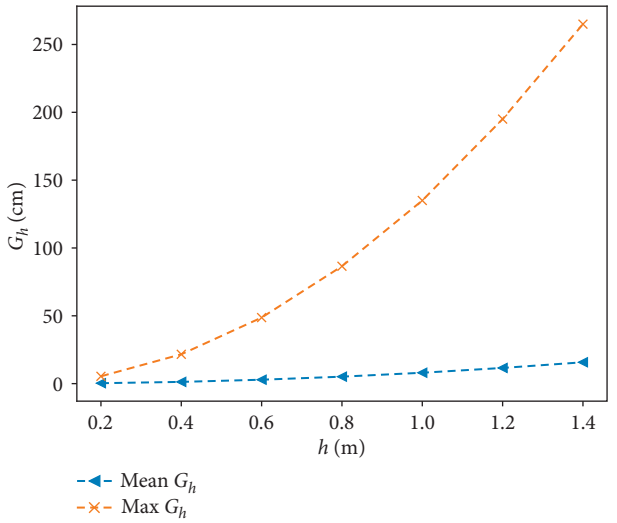


Figure 3: Graph of G_h change with tag height.

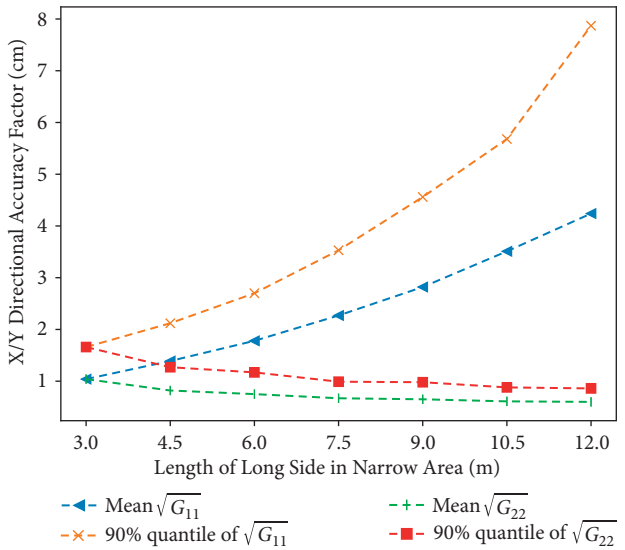


Figure 4: Directional factor changes with rectangular length.

and efficient UWB BS layout scheme. The theoretical optimization model is transformed into a multiobjective optimization solution model using the weight of the judgment matrix. Furthermore, the model considers the different positioning accuracy requirements for different areas. The scale of 1–9 is introduced, where a_{ij} denotes the importance of area s_i relative to area s_j, and the corresponding meanings of different values are shown in Table 2.

After obtaining the weight coefficients, based on the linear weighting method, the optimization problem of the BS layout considering the difference in location accuracy requirements is transformed into the optimal solution of equation (17). The solution is then the optimal layout of the UWB BSs:

$$\begin{cases} \min \sum_{i=1}^{m} w_i \cdot HDOP_i(x, y), xy \in D, \\ HDOP_i(x, y) \le \dfrac{\sigma_i}{\sigma_{\Delta r}}, \end{cases} \tag{17}$$

where $H\ DO\ P_i(x, y)$ denotes the HDOP in positioning area i; σ_i denotes the standard deviation of the maximum positioning error required by the area; and $\sigma_{\Delta r}$ denotes the standard deviation of the difference error.

Using MATLAB to simulate this method, take the positioning area shown in Figure 5(a) as an example, in which the standard deviation of the positioning error required by Area 1 is $\sigma_1 \le 10$ cm; Area 2: $\sigma_2 \le 15$ cm; Area 3: $\sigma_3 \le 20$ cm; and Area 4: $\sigma_4 \le 30$ cm. According to the judgment matrix, the weights of each region are obtained as follows:

$$\begin{aligned} w_1 &= \frac{2}{5}; \\ w_2 &= \frac{4}{15}; \\ w_3 &= \frac{1}{5}; \\ w_4 &= \frac{2}{15}. \end{aligned} \tag{18}$$

The most suitable layout scheme is obtained by a MATLAB genetic algorithm simulation, as shown in Figure 5(b).

Table 2: Corresponding meanings of different relative importance values.

a_{ij}	1	3	5	7	9
Meaning of a_{ij}	S_i equals S_j	S_i slightly more important than S_j	S_i evidently more important than S_j	S_i extraordinarily more important than S_j	S_i superficially more important than S_j

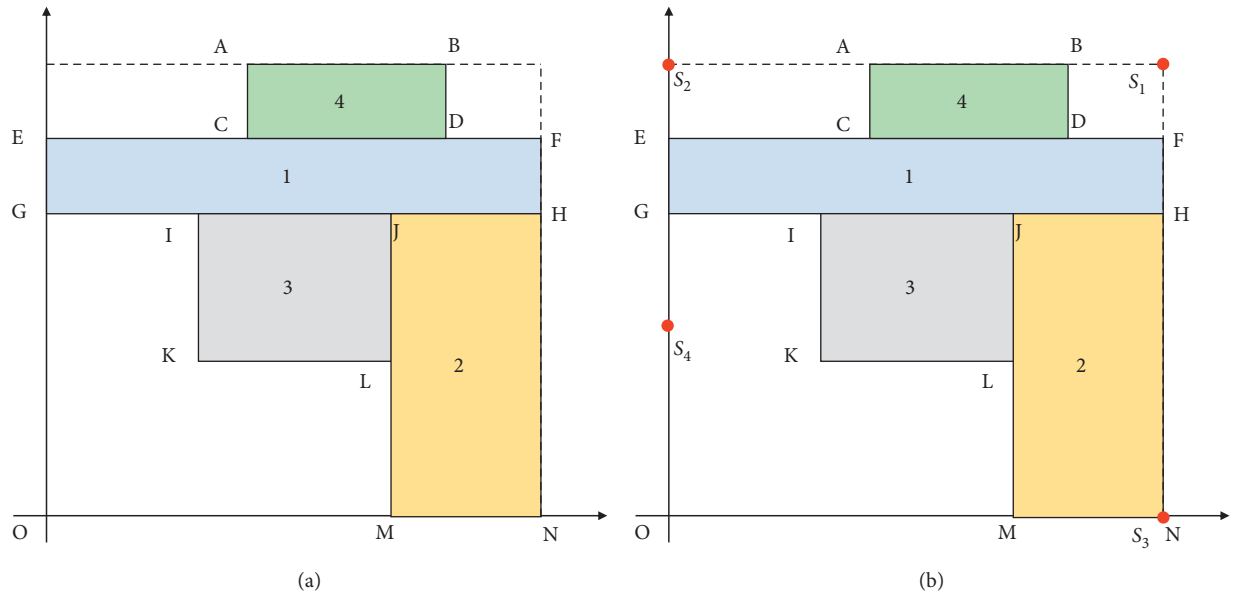


Figure 5: Base station layout considering different positioning accuracy requirements.

If the standard deviation of the time difference error measurement, $\sigma_{\Delta r}$, is 5 cm, the layout of the BS is checked by inverse calculation. The HDOP of the positioning error of each location area is shown in Table 3 and meets the location accuracy requirement. Therefore, the layout scheme is feasible.

3.1.9. Three-Stage UWB Station Layout Planning Method. According to the previous analysis, the layout design methods for various practical scenarios are integrated, and a practical application-oriented regional UWB station layout optimization design method is proposed. The method comprises three stages:

Step 1: determine the deployment strategy. The environmental characteristics of each subregion in the region to be located are extracted, and the BS layout strategy of each subregion is determined according to the characteristic environmental parameters.

Step 2: determine the layout parameters. Based on the positioning accuracy requirement of each subregion, the layout strategy, and equipment performance parameters, the specific layout spacing, layout number, and other parameters are obtained by the corresponding methods.

Step 3: comprehensive adjustment and scheme verification. The layout scheme of each subregion is integrated, and the adjusted scheme is validated by simulating the HDOP. If the location accuracy requirement is satisfied, the scheme of parameters, symbols, and graphics can be further generated. Otherwise, the layout strategy must be determined again until the scheme is validated.

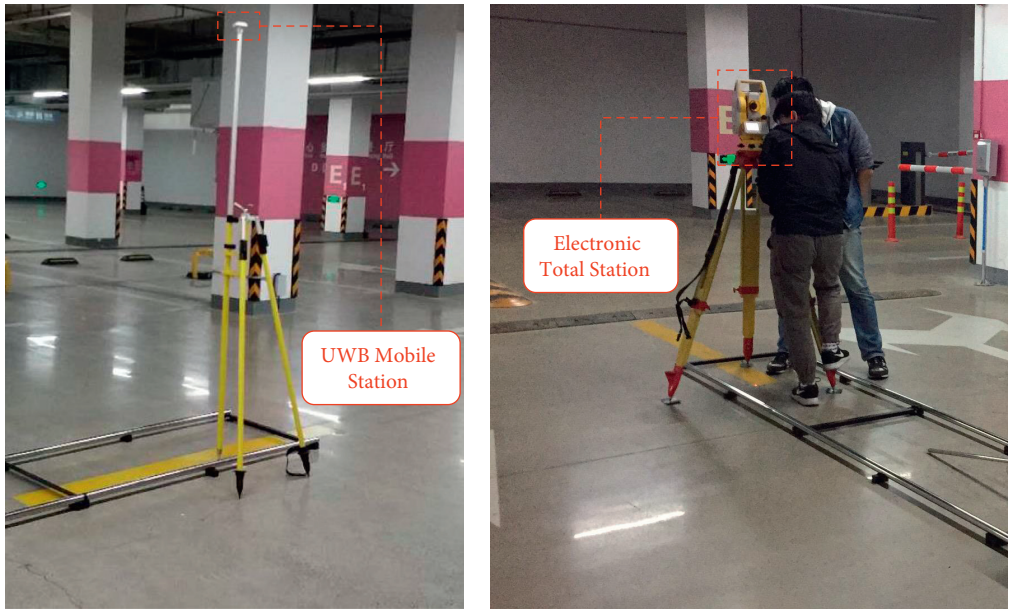
Table 3: Simulation results of plane accuracy factors for each localization area.

Location area	HDOP			
	Average	Maximum	90% indexing value	Required value
Area 1	1.4	1.82	1.42	2
Area 2	1.56	2.36	1.69	3
Area 3	1.33	1.55	1.36	4
Area 4	1.54	1.78	1.58	6

In the whole design process of the BS layout scheme, the main parameters involved include (1) characteristic environmental parameters, such as slope, undulating area of surface, and narrow area; (2) positioning accuracy parameters, including the size and direction of positioning errors required by each region; and (3) equipment performance parameters, including the BS signal coverage distance and antenna type.

4. Case Study

We verify the performance of the proposed optimization design method of the UWB BS layout using the underground parking lot of Tongji University, Shanghai. As shown in Figure 6, tripod with a MS and an electronic total station is



(a) (b)

Figure 6: A tripod with MS and an electronic total station.

applied to obtain the high-precision positioning data as the ground truth. According to the spatial characteristics of different areas in the parking lot and the requirements for ensuring the safety of AVP in different areas [27], the positioning accuracy requirements of each area are analyzed and summarized in Table 4.

According to the three-stage method proposed above, the layout optimization design of BSs is conducted as follows.

4.1. Determining Layout Strategy. According to the analysis of the layout in the underground parking lot, the whole area can be divided into four subareas, as shown in Figure 7(a). Among them, Area 1 is the entrance and exit of the parking lot, and a rectangular layout scheme can be adopted. Area 2 is a ramp with coupling characteristics between the ramp area and narrow area. After calculating the coverage radius of the ramp according to the slope gradient, the BS layout method based on direction-weighted HDOP can be adopted. Area 3 is a driving lane belonging to a typical narrow region. The BS layout design method based on direction-weighted HDOP should be adopted. Area 4 is the parking spots for which a rectangular layout scheme can be adopted.

4.2. Determining Layout Parameters. The installation height of the BSs in the UWB positioning system is 2.6 m, the height of the tag is approximately 2.0 m, and the effective coverage distance of the tag signal is 30.0 m. Meanwhile, the standard deviation of the station positioning error and time measurement error in the system are both 5.0 cm. Then, the plane coverage radius of the BS is $r = \sqrt{R^2 - (H-h)^2} = \sqrt{30^2 - (2.6-2)^2} = 29.9$ m.

According to the above equipment parameters and the site size parameters of the parking lot, and the layout strategy of each subarea determined in the first stage, the specific layout parameters of each subarea are calculated.

(1) Area 1 is the rectangular area with four stations. The simulation results show that the 90% quantile of the HDOP in this area is 1.6081, meeting the positioning accuracy requirement for AVP.

(2) Area 2 is the ramp area. The maximum slope of the ramp is 15%, and the ramp belongs to a narrow area. The aspect ratio should not exceed 2.6. Six BSs should be set up with the distance between adjacent BSs on the same side being $22.3/2 = 11.15$. Simultaneously, the optimal layout method based on direction-weighted HDOP is used to check out. In the direction-weighted HDOP model, the weight along the slope direction is 2/3, and the weight along the vertical slope direction is 1/3. The simulation results with the genetic search algorithm show that the 90% indexing value of the positioning accuracy factor along the slope direction is $1.2641 \leq 2$; in the vertical slope direction, this is $2.2665 \leq 4$. Both of these meet the positioning accuracy requirements.

(3) Area 3 is the driving lane. According to AVP requirements, the aspect ratio of the rectangular scheme does not exceed 2.6. Six pairs of BSs should be set up with an interval between adjacent BSs of 13.4 m on one side. These should then be checked using the direction-weighted HDOP model. The accuracy of each direction meets the requirements.

(4) Area 4 is the parking spots. If the standard deviation of 90% regional positioning error is less than 20 cm, the

Table 4: Accuracy requirements of location in parking lot area.

Areas	Features	Permissible standard deviation of 90% regional positioning error
Vehicle lane	High speed and simple situation	Driving direction 10 cm, vertical driving direction 20 cm
Plane intersection	Low speed and complex situation	10 cm
Ramp	Low speed and complex situation	Driving direction 10 cm, vertical driving direction 20 cm
Entrance and exit	Low speed and complex situation	Standard deviation 10 cm
Parking space	Low speed and simple situation	Driving direction 20 cm, vertical driving direction 20 cm

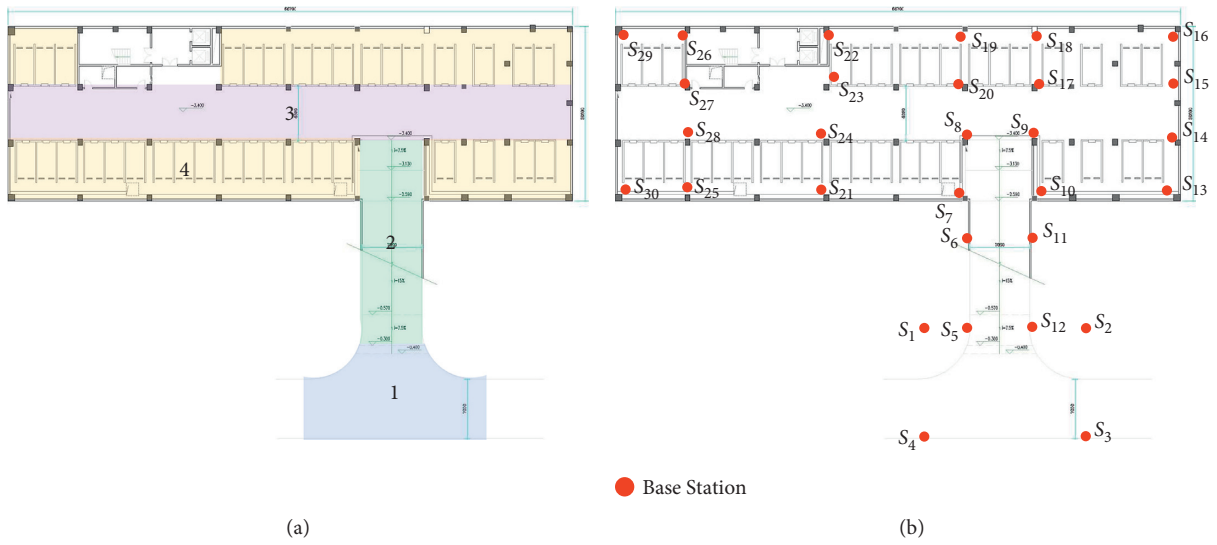


Figure 7: Schematic map of areas and adjusted UWB BS layout scheme.

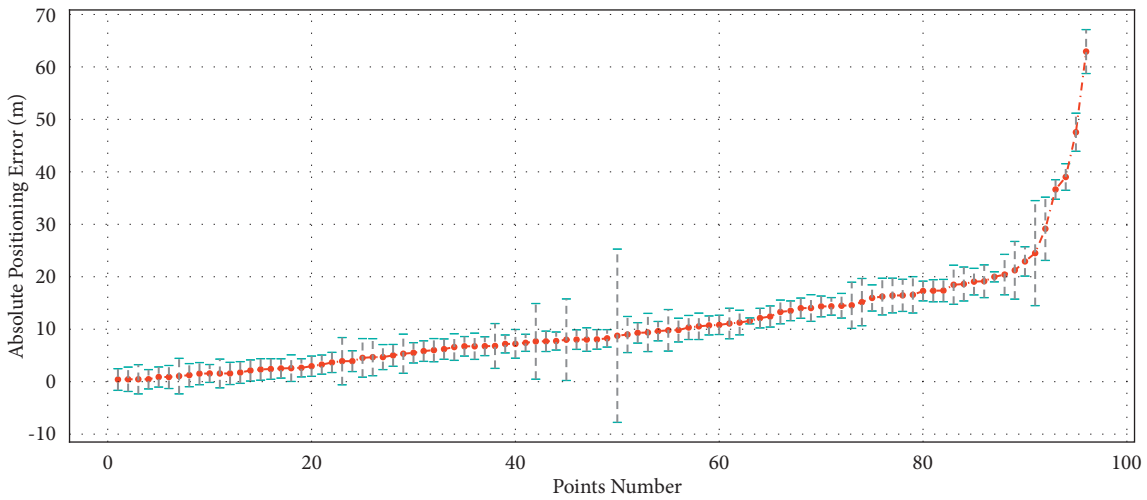


Figure 8: Absolute positioning error at each verification point arranged from small to large. The error bars represent the standard deviation of measurements.

length-width ratio of the rectangular layout can exceed 3. The sub-blocks, which are physically separated in this area, can be arranged according to this aspect ratio.

4.3. Comprehensive Adjustment and Scheme Verification. The BS layout schemes of each subregion obtained in the second stage are aggregated, and the BSs with overlapping or extremely similar locations are merged. Finally, the layout scheme shown in Figure 7(b) is obtained. After adjustment, the total number of BSs decreases to 30 compared with the original 41 BSs, with a reduction of 27%.

The actual positioning accuracy of the underground parking lot is verified by selecting the grid points with a spacing of 2.5 m. The results are shown in Figure 8. The range of the grey dotted line represents plus or minus the standard deviation of the positioning error. As shown in Figure 9, there are 96 verification points (sorted by positioning error). Among them, 90 points have standard

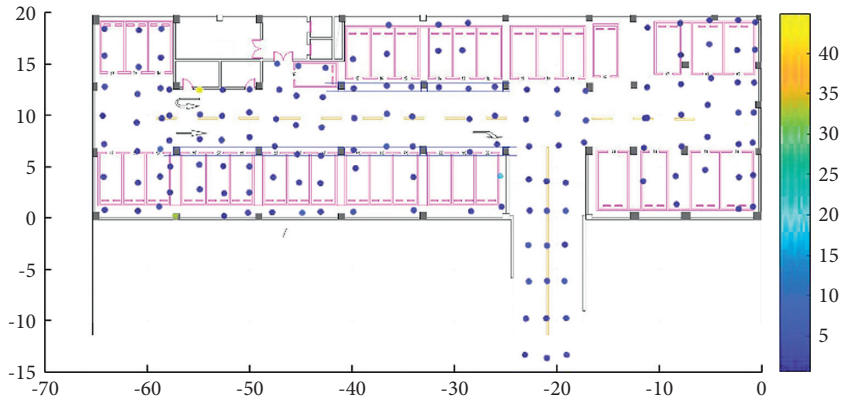


Figure 9: Variance distribution of measured positioning errors of the underground garage.

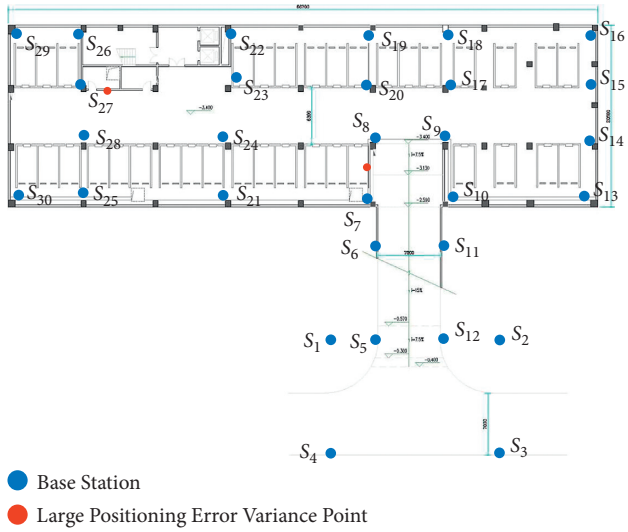


Figure 10: Point distribution with large positioning error variances.

deviations of less than 5 cm, accounting for 93.75%; four points have standard deviations between 5 and 10 cm, accounting for 4.17%; and only two points have standard deviations exceeding 10 cm, accounting for 2.08%. The positioning accuracy meets the AVP requirements. In Figure 10, we indicate the positions of two points with significantly larger positioning error variance, which can be seen that they are both close to the wall. Therefore, our method has high performance for the actual required area (a reasonable distance from the wall).

5. Conclusion

This article mainly focuses on the layout planning of UWB BSs for indoor localization of AVP. The TDOA is applied as the positioning approach of the UWB system, and the HDOP is chosen as the accuracy evaluation index. After conducting theoretical analysis and simulation for ideal and actual conditions, we determine the targeted layout analysis method under each condition of the underground parking lot, such as ramp area, surface fluctuation area, and narrow area. A three-stage practical and economical layout design method of UWB BSs is then introduced, including determining the deployment strategy and layout parameters, making comprehensive adjustments, and verifying the scheme. Taking the actual layout planning of a UWB system in an underground parking lot of Tongji University, Shanghai, China, the experimental results show that the proposed three-stage layout design method of UWB BSs for regional differential positioning ensures the accuracy requirements for AVP and greatly reduces the number of BSs. This indicates that the proposed layout planning method of UWB BSs for AVP has high efficiency and good economy.

Data Availability

The data used to support the findings of this study are included within the article and are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was jointly sponsored by the National Key R&D Program of China under Grant 2021YFB1600403, the Innovation Program of Shanghai Municipal Education Commission under Grant 2021-01-07-00-07-E00092, and the Scientific Research Program of Shanghai Municipal Science and Technology Commission under Grants 19DZ1208700 and 21DZ1205100. The work of Cong Zhao was supported by the Shanghai Sailing Program under Grant 21YF1449400.

References

[1] L. Chettri and R. Bera, "A comprehensive survey on internet of things (IoT) toward 5G wireless systems," *IEEE Internet of Things Journal*, vol. 7, no. 1, 2020.

[2] Y. Du, B. Qin, C. Zhao, Y. Zhu, J. Cao, and Y. Ji, "A Novel s-temporal Synchronization method of roadside

asynchronous MMW radar-camera for Sensor fusion," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2021.

[3] Y. Du, J. Chen, C. Zhao, C. Liu, F. Liao, and C. Y. Chan, "Comfortable and energy-efficient speed control of autonomous vehicles on rough pavements using deep reinforcement learning," *Transportation Research Part C: Emerging Technologies*, vol. 134, p. 134, Article ID 103489, 2022.

[4] C. Zhao, J. Cao, X. Zhang, and Y. Du, "From search-for-parking to dispatch-for-parking in an era of connected and automated vehicles: a macroscopic approach," *Journal of Transportation Engineering, Part A: Systems, Vol*, vol. 2, Article ID 04021112, 2022.

[5] C. Zhao, F. Liao, X. Li, and Y. Du, "Macroscopic modeling and dynamic control of on-street cruising-for-parking of autonomous vehicles in a multi-region urban road network," *Transportation Research Part C: Emerging Technologies*, vol. 128, p. 103176, 2021.

[6] A. Ibisch, S. Stümper, H. Altinger et al., "Towards autonomous driving in a parking garage: vehicle localization and tracking using environment-embedded LIDAR sensors," in *Proceedings of the 2013 IEEE Intelligent Vehicles Symposium (IV)*, pp. 829–834, IEEE, Gold Coast, QLD, Australia, June, 2013.

[7] C. Yuan, L. Fei, C. Jianxin, and J. Wei, "A smart parking system using WiFi and wireless sensor network," in *Proceedings Of the 2016 IEEE International Conference On Consumer Electronics-Taiwan (ICCE-TW)*, pp. 1-2, IEEE, Nantou, Taiwan, May, 2016.

[8] M. Qiang, L. Jiaming, Z. Qinfeng, and H. S. U. Chunliang, "Indoor parking navigation system using visible LED light communication," in *Proceedings of the 2019 IEEE International Conference on Computation, Communication and Engineering (ICCCE)*, pp. 157–159, IEEE, Fujian, China, November, 2019.

[9] A. Alarifi, A. Al-Salman, M. Alsaleh et al., "Ultra wideband indoor positioning technologies: analysis and recent advances," *Sensors*, vol. 16, no. 5, p. 707, 2016.

[10] W. J. Wei-Jen Chen and R. M. Narayanan, "Antenna placement for minimizing target localization error in UWB MIMO Noise radar," *IEEE Antennas and Wireless Propagation Letters*, vol. 10, pp. 135–138, 2011.

[11] D. Jourdan, D. Dardari, and M. Win, "Position error bound for UWB localization in dense cluttered environments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 2, pp. 613–628, 2008.

[12] S. Monica and G. Ferrari, "UWB-based localization in large indoor scenarios: optimized placement of anchor nodes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 2, pp. 987–999, 2015.

[13] F. Long, A. Behnad, and X. Wang, "Optimum reference node deployment for indoor localization based on the average Mean Square Error minimization," in *Proceedings of the 2015 IEEE 34th International Performance Computing and Communications Conference (IPCCC)*, pp. 1–6, IEEE, Nanjing, China, December 2016.

[14] T. Qiao and H. Liu, "Incremental anchor layout for indoor positioning," in *Proceedings of the 2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, Paris, France, May 2017.

[15] A. E. C. Redondi and E. Amaldi, "Optimizing the placement of anchor nodes in RSS-based indoor localization systems," in *Proceedings of the 2013 12th Annual Mediterranean Ad Hoc Networking Workshop (MED-HOC-NET)*, pp. 8–13, IEEE, Ajaccio, France, June 2013.

[16] J. Junyi Zhou, J. Jing Shi, and X. Xiuli Qu, "Landmark placement for wireless localization in rectangular-shaped industrial facilities," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 6, pp. 3081–3090, 2010.

[17] R. Sharma and V. Badarla, "Analysis of a novel beacon placement strategy 3D localization in indoor spaces," in *Proceedings of the 2019 11th International Conference on Communication Systems & Networks (COMSNETS)*, pp. 320–327, IEEE, Bengaluru, India, 2019.

[18] R. Sharma and V. Badarla, "A m optimization tool chain for 3-D indoor beacon placement problem," *IEEE Internet of Things Journal*, vol. 8, no. 17, pp. 13439–13448, 2021.

[19] E. Kim and D. Choi, "Planning of UWB indoor positioning network using binary integer linear programming," *International Journal of Ultra Wideband Communications and Systems*, vol. 3, no. 3, pp. 166–176, 2016.

[20] Y. Wu, S. Ding, Y. Ding, and M. Li, "UWB base station cluster localization for unmanned ground vehicle guidance," *Mathematical Problems in Engineering*, vol. 2021, 2021.

[21] H. Pan, X. Qi, M. Liu, and L. Liu, "Map-aided and UWB-based anchor placement method in indoor localization," *Neural Computing & Applications*, pp. 1–15, 2021.

[22] L. Santoro, D. Brunelli, and D. Fontanelli, "On-line optimal ranging sensor deployment for robotic exploration," *IEEE Sensors Journal*, 2021.

[23] G. Wang, W. Zhu, and N. Ansari, "Robust TDOA-based localization for IoT via joint Source position and NLOS error estimation," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8529–8541, 2019.

[24] K. Balac, M. Prevostini, and M. Malek, "Optimizing sensor nodes placement for fault-tolerant trilateration-based localization," in *Proceedings of the 2015 IEEE 21st Pacific Rim International Symposium on Dependable Computing (PRDC)*, pp. 109–118, IEEE, Zhangjiajie, China, November 2016.

[25] T. Leune, C. Koch, and G. Von Cölln, "Anchor node placement in complex environments with physical raytracing and genetic algorithms," in *Proceedings of the 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1–6, IEEE, Alcala de Henares, Spain, October 2016.

[26] R. Adasiewicz, M. Ganzha, M. Paprzycki et al., "Optimal placement of internet of things infrastructure in a smart building," in *Proceedings of the Second International Conference on Information Management and Machine Intelligence. Lecture Notes in Networks and Systems*, pp. 661–669, Singapore, January 2021.

[27] Z. Xu, "A precise 3D positioning approach based on UWB with reduced base stations," in *Proceedings of the 2021 IEEE 15th International Conference on Anti-counterfeiting, Security, and Identification (ASID)*, pp. 145–149, IEEE, Xiamen, China, October 2021.

[28] R. Yarlagadda, I. Ali, N. Al-Dhahir, and J. Hershey, "GPS GDOP metric," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 147, no. 5, pp. 259–264, 2000.

[29] Z. Miao, L. Cui, B. Zhang, and J. Li, "Deployment patterns for k-coverage and l-connectivity in wireless sensor networks," in *Proceedings of the IET International Conference on Wireless Sensor Network 2010 (IET-WSN 2010)*, pp. 73–77, Beijing, China, November 2010.

[30] China-SAE and China Communications Industry Association, *General Technical Requirements of Automated Valet Parking Systems*, Beijing, China, 2020.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 5918128, 13 pages
https://doi.org/10.1155/2022/5918128

Research Article

A Software-Defined Networking Roadside Unit Cloud Resource Management Framework for Vehicle Ad Hoc Networks

Hongming Li,[1,2] **Dongxiu Ou**,[1,2] **Iftikhar Rasheed**,[3] **and Meiting Tu**[1,4]

[1]*The Key Laboratory of Road and Traffic Engineering, Ministry of Education, School of Transportation Engineering, Tongji University, Shanghai 201804, China*
[2]*Key Laboratory of Railway Industry of Proactive Safety and Risk Control, School of Transportation Engineering, Tongji University, Shanghai 201804, China*
[3]*Department of Information and Communication Engineering, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan*
[4]*College of Automatic, University Paris-Saclay, Saint Aubin, Paris 91190, France*

Correspondence should be addressed to Dongxiu Ou; ou.dongxiu@tongji.edu.cn

Received 19 November 2021; Accepted 11 January 2022; Published 8 February 2022

Academic Editor: Alain Lambert

Roadside unit (RSU) cloud and its vehicle-to-infrastructure (V2I) connectivity can enable various security, entertainment, and shared mobility applications for vehicles in intelligent transportation systems (ITS) through wireless communications. In this article, the deep programmability of software-defined networking (SDN) is employed to dynamically reconfigure network hosting services and their data forwarding information for effectively meeting the basic shared mobility applications' needs in vehicle ad hoc networks (VANETs). Multipath is also enabled to forward data flow for balancing network links utilization rate and SDN is thus utilized to achieve the minimum cloud delay with the least number of hosts, which can be summarized as a mixed-integer linear programming (MILP) problem. The joint optimization (JO) algorithm is proposed and in contrast to the two single-objective algorithms which are the delay optimization (DO) algorithm and host optimization (HO) algorithm, respectively. Results show that, for the single-threading instance, the JO and DO algorithms are the same in essence. For the multithreading instance, the JO algorithm generally outperforms the two single-objective optimization algorithms, respectively, under given demands. Furthermore, results also demonstrate that the services should be deployed globally in a distributed manner rather than in the centralized manner for achieving the minimized cloud delay in designing an RSU cloud.

1. Introduction

Transportation is leading an important service industry in the national economy while traffic congestion and safety are the major problem faced by almost all core cities in the world [1]. The traditional method of controlling traffic flow based on traffic lights and stop signs cannot dynamically adjust the interval time in real-time according to the actual traffic flows. Variable speed limit is a major ITS technology for controlling freeway mainline traffic, which have been increasingly used to improve traffic safety and operations efficiency of freeway traffic management [2]. What is more, electric vehicles and its related applications are the main trend in recent years, and its configuration optimization to prolong driving mileage has become an essential issue in affecting the future prospect of electric vehicles [3]. Though there are some shared mobility applications, such as ride splitting, enabling riders with similar routes to share a ride sourcing trip, which is a promising transportation technology to reduce traffic congestions and air pollution, the traffic management can be enhanced through effective communications among the vehicles [4, 5]. However, the pure vehicle-to-vehicle (v2v) communications are unstable, due to the occlusion of towering buildings on the roadside or the other obstacles resisting the communication link, in which case the links are likely to be unstable and decrease the communication quality. Besides, the network topologies are dynamic, because of the high-speed movements of the

vehicles. Therefore, the established communication links among vehicles may be disconnected, which affects the success rate of data transmission. In V2V, the communications among the vehicles are not stable enough and the quality of communication is determined by many uncontrollable factors, such as vehicle speeds, geographical locations, and weather conditions. Considering the problems that existed in the V2V communications, the RSUs are utilized for assisting the efficient communications to form the VANETs environment.

The RSUs play a vital role in the VANETs system and are usually deployed on both sides or the middle of the road, which can be equivalent to an AP (access point) in a wireless network. Vehicles driving on the roads can use RSUs to realize interconnection and intercommunication. At the same time, the existence of RSUs makes the network more stable and reliable when uploading and downloading data, and the bandwidth will increase accordingly. To improve the network performance in the VANETs, Yang et al. propose a connectivity-oriented maximum coverage RSU (CMCS) deployment scheme which aims to maximize the performance of V2I communication in urban environments [6]. Another role of the RSUs is to provide computing services and serve as a temporary storage point for data. Some data content such as video data with high throughput rates from vehicles and road condition information that are commonly requested can be temporarily stored in RSUs, without repeated requests to the Internet. VANETs play a vital role in assisting driving, traffic accident warning, traffic management, and Internet services. To better adjust the traffic congestion situation, intelligent transportation system (ITS) applications would be of great importance with the help of VANETs [7]. Under this circumstance, smart cars can automatically adjust the interval between traffic lights through interaction with the local vehicle-road collaboration system under VANETs environment, dynamically optimize the traffic flow at road intersections, improve vehicle traffic efficiency, and alleviate congestion.

The main communication ways in VANETs include the communications among vehicles named V2V (vehicle-to-vehicle) and the communications between the vehicles and RSUs called V2I (vehicle-to-infrastructure), where the typical features of V2V are described above. As far as V2I communication is concerned, due to the fixed geographic location of the RSUs, the relatively stable network, and relatively few interference factors, V2I has been studied in-depth. At present, there are still many problems in V2I communication, such as access selection, data distribution, and frequent switching of RSUs, which restrict the further development of VANETs and cannot provide users with more real-time and high-quality communication services. In terms of RSUs service, due to the high dynamics of VANETs and the limited coverage areas of RSUs, multiple vehicles enter the areas covered by RSUs at the same time. Delay-sensitive safety information would probably result in not being delivered in time, and nonsafety application entertainment information might occupy most of it. It is related to whether the vehicles can get timely processing information and satisfactory services. It will also affect the operating status of the RSUs in the entire region, including throughput, bandwidth utilization, and load balance. Network congestion sometimes occurs in some RSUs, while other RSUs have idle bandwidth and waste. The unbalanced load of RSUs will directly lead to the degradation of network performance [8]. Therefore, how the vehicles choose the appropriate RSUs to get served will be of great importance. In the RSU cloud where there are multiple RSUs, a more scientific algorithm is needed to obtain the best solution for vehicles to get served so that the performance of all aspects of the network is maintained stable. In the same way, in the current VANET data distribution system, due to the limitation of the coverage of RSU and the solidification of the Internet of vehicles architecture and communication methods, the service instances of the vehicles need to be transferred or copied to other RSUs when transmitting large-volume data.

Various methods have been studied with different ways to achieve efficient communications among the vehicles in the ITS systems under VANETs environment. As a matter of fact, the RSU clouds are proposed to help build the VANETs environment, for example, to increase the number of communication links for sending data packets. When the service of the current vehicles cannot be met or need to be transferred, deleted, or copied to other RSUs, how to manage the computing and communication resources in the RSU cloud is an essential problem. Therefore, to overcome this problem, the SDN-based approach with an RSU-based resources management framework for VANETs is proposed. The problems are firstly focused on how to manage the RSUs' characters in the process of setting data dissemination policies and then optimize the RSUs' resources deployment with the dynamic existing communication demands. In this case, a mixed-integer linear programming (MILP) model is developed to optimize the allocation of the RSUs' resources in the designed SDN RSU cloud architecture while minimizing the communication delays and number of working hosts. Furthermore, the RSUs' resources redundancy is also considered upon the basis of ensuring the requirements of communication delay and working hosts' number.

For the rest part of the paper, the related work is firstly reviewed and then the proposed SDN-based RSU cloud framework is introduced. Next, the SDN-based RSU cloud resource management problem is formulated into the mixed-integer linear programming (MILP) problem with some restrictions, which is solved by building a simulation framework using Python that is mixed programmed with Gurobi. Meanwhile, the JO algorithm is compared with the single-objective optimization algorithms which are HO and DO algorithms, respectively. Finally, the paper is concluded, and the future work is summarized.

2. Literature Review

To ensure the efficient communication in the VANETs, constructing the cloud of OBUs and (or) RSUs and (or) the hybrid cloud are the typical ways for achieving the generality of the entire network. In this section, lots of existing work on the various clouds in the VANETs shall be briefly reviewed.

On the one hand, a cloud of on-board units (OBUs) are playing vital roles in the VANETs, where some RSUs still act as the gateways to the traditional communication networks [9]. Behbehani et al. [10] introduce a vehicular cloud architecture supported by fixed edge computing nodes and the central cloud; mixed-integer linear programming (MLP) model is developed to optimize the allocation of the computing demands in the distributed architecture while minimizing power consumption. Nodes that are consisted of "SmartCloud" vehicles in the minimum connected dominating set form a virtual backbone in the VANETs [11], where a "SmartCloud" vehicle sends periodic beacon messages to the nearest RSU. For the above method, the common communication evaluation metrics such as the transmission range, processing power, and delay reduction are greatly improved. Meanwhile, the methods mentioned above mainly devote their efforts into the utilization of the vehicles' OBUs, rather than taking advantages of the RSUs.

On the other side, Wang et al. [12] construct a scene where information centres (or base stations) are distributed along the road so that the information centres can broadcast messages in time, aiming to work out the information dissemination problem in a joint vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) communication system. Though it is possible to interact between the vehicles and roads in the local area, the decision-making is based on the individuals or the local groups, they do not consider the problem from the global situation in a road network, which cannot make full use of the RSUs' extra communication and computing resources and it is impossible to achieve the overall situation from the perspective of safety or efficiency optimization. Meanwhile, cloud computing is a new and promising paradigm delivering IT services as computing utilities, which is designed to provide services to external users, providers for sharing their resources, and capabilities [13]. For a cloud which is composed of vehicles, RSUs, and passengers, the traffic flows are usually distributed in an unbalanced way in the region, and it remains dynamic as time goes by, the types of the ITS applications and the bandwidth demands of the vehicles are different for the various cloudlets of the entire network. Therefore, it is time to make full use of the RSUs' extra communication and computing resources with less demands, letting the RSUs with larger demands alleviated in serving the cars. Tang et al. have proposed an RSU empowered vehicular network that consists of three hierarchical layers-vehicular cloud, RSU-enabled cloudlet, and central cloud, respectively, where RSUs are enhanced with edge servers so that the entire system can intelligently respond to the real-time resource requests [14]. Zheng et al. employ the knapsack problem to formulate the problem of cooperative scheduling and radio resource management in vehicular networks, but their research only orients to the radio resources for nonreal-time services where they assume the problem in the 2-hop vehicular network, greatly restricting the application scenario [15]. Their proposed methods are not efficient enough considering the high mobility of vehicles in VANETs with strict delay requirements. In fact, VANET evolves with two emerging paradigms: vehicle cloud computing and information centric networks, where vehicle cloud computing brings the mobile cloud model to vehicle networks and information centric networks change the notion of data routing and dissemination [16].

Besides, to date, there is no set architecture for cloud computing from academic point. Olariu et al. envisioned the paradigm shift from conventional VANET to vehicular cloud computing by merging VANET with cloud computing [17], whereas they did not carry out their thoughts in a specific scenario. With the development of computer networks, McKeown et al. have proposed OpenFlow which is the de facto standard of software-defined networks, where it can provide an excellent opportunity to improve network performance through flexible centralized control, decoupling control plane (network control functions) from the data plane (forwarding functions) [18]. The control plane communicates with the data plane using OpenFlow protocol, where the controller enables the ITS operators to set data dissemination policies and integrate services. Salahuddin et al. have proposed a unique SDN Cloud RSU architecture and its SDN advances from the previous in two parts: firstly, it consists of traditional RSUs and small data centres; secondly, it can be dynamically reconfigured to meet changes in service demands ensuring QoS [19]. However, they do not take the RSUs' resources redundancy into consideration in case that the networks communications demands are too fluctuant to meet. In this research, a mixed linear integer (MILP) programming model is developed to optimize the RSU resource allocation in the designed SDN RSU cloud architecture. While minimizing communication delays and the number of hosts in the cloud, the redundancy of RSU resources is also considered.

3. Materials and Methods

Since that SDN was proposed as a new-generation network architecture, its unique concept of separation for data and control brings us convenience and flexibility in designing a cloud. The SDN architecture based on the OpenFlow protocol is divided into three planes: application plane, control plane, and infrastructure plane. Control plane mainly includes OpenFlow controller and network operation system. The OpenFlow controller encapsulates the northbound interface driver that connects to the required plane; the southbound interface driver connects to the switch, ensuring the internal intelligent logic of SDN control. The control plane realizes data transmission with OpenFlow switches via the southbound interface, obtaining the status information of the underlying infrastructure and performing reasonable scheduling of the underlying network resources. Obviously, the controller is the key to the centralized control function of the entire SDN network. To achieve the best performance of the RSU cloud infrastructure, it is also necessary to implement multipaths in an SDN that supports OpenFlow. In the proposed SDN RSU cloud framework, the RSUs are composed of the physical servers which can store and compute data and switches that can exchange nonsafety apps information from the vehicles on the road sections and nearby RSUs. The physical machine of the RSU enables

virtualization so that the virtual machines can run multiple replications to host services on a single physical machine. Meanwhile, the services can be deleted, transferred, or copied to other RSUs when the vehicles are travelling to the next RSU. The switches on the RSUs are all OpenFlow switches which are the de facto standards of SDN, performing as the corresponding functions of the data forwarding layer. However, though there is no stringent QoS requirement for ITS apps, it is highly desirable and usually crucial for its users [20]. The proposed SDN RSU cloud framework is shown in Figure 1.

Figure 1 depicts the SDN RSU cloud framework and there are n $(0 \leq n < N)$ RSUs which are connected to each other through a high-speed backbone. The development of the vehicle-road collaborative technology via cloud computing aims to achieve vehicle-road information sharing system based on advanced communication technologies by deploying roadside intelligent sensing computing equipment (collectively referred to as roadside infrastructure) and complementing the car-side through roadside perception. The RSUs are deployed in the centre of the roads or intersections for maximizing the coverage area of the vehicles to be served. For the entire region, since there are more vehicles on the arterial roads and less vehicles on the secondary roads, the demands are unbalanced distributed. The unevenly distributed network demands would greatly fluctuate as the traffic density changes. As the density of traffic flows is changing with time, network demands are varying to a large extent. Due to the varying network demands, it is necessary to dynamically reconfigure the SDN RSU cloud for optimizing the communication efficiency.

There are three RSU function characters which are the OpenFlow controllers, RSU cloud resource managers (CRMs), and normal RSUs. OpenFlow controllers are responsible for formulating and managing the forwarding strategy of the flow table which can update the OpenFlow switch flow rules via the control plane. Meanwhile, the CRMs would disseminate information regarding service hosting, service migration, data flow changes, and virtual machines (VMs) instantiating and (or) eliminating via the data plane; the designed characters are simplified compared with the research by Salahuddin et al. [21] where there are three characters. It is worth mentioning that the RSU cloud resource managers communicate with the OpenFlow controllers via the data plane.

The top layer of the architecture in Figure 1 is the Internet, which contains a large quantity of Internet servers to provide data upload and download services. The RSUs in the cloud not only transmit information of ITS services but also serve as the different characters, respectively, ordered by the controller information that the network manager has given. The fixed SDN RSU cloud upper layer connections in Figure 1 are all wired broadband connections, including the connections between normal character RSUs and other normal character RSUs, the normal character RSUs and OpenFlow controllers, normal character RSUs and CRMs, and OpenFlow controllers and CRMs, ensuring the reliability of the data transmission. The communications between mobile vehicles and the RSUs adopt IEEE 802.11p wireless communication. There are mostly RSUs with normal characters which are given nor the function of OpenFlow controllers neither cloud resource managers as shown in Figure 1.

4. SDN RSU Cloud Resource Management

Cloud resource management pays attention on the optimal allocation of computing and network resources while meeting the fluctuating demands required by the ITS apps. Normally, optimizing the placement of virtual machines on physical host machines is determined by various capacity planning tools such as VMware Capacity Planner, IBM WebSphere CloudBurst, Micro Focus, and OneIQ; these cloud resource management tools seek to consolidate physical memory, VMs for CPU, and power consumption savings, etc., yet without considering consumption of network resources [22]. Consequently, some RSUs will inevitably cost lots of network resources, while others are almost idle, causing the unbalanced load of the entire network, which would directly lead to the degradation of network performance. Defining the configuration as the number of service hosts and effective data forwarding rules, the mission is to find the best configuration in order that the whole SDN RSU cloud will achieve the minimum delay and minimum number of hosts when the network demands change.

4.1. Problem Formulation. Considering a commonly seen traffic network which is fixed with a bunch of RSUs facing the fluctuating communication demands, the mathematical problem is then abstracted as follows. Assume that a network graph $G = (Vertices, Edges)$ is given with fixed number of RSUs, whose *vertices and edges* are representing the RSUs and the wired connections among the RSUs, respectively.

From what has been shown in Figure 2, given a set of traffic information services S, a set of average demands $D = \{D_{t_0}, D_{t_1}, D_{t_2}, \ldots, D_{t_n}\}$ that change over time periods $T = \{t_0, t_1, t_2, \ldots, t_n\}$. The vehicles are constantly sending their demands at each RSU and requesting for services from the RSU cloud. There is an initial network configuration Z_{t_0} during period t_0. For any edge, there is a bandwidth capacity C_e. During the time period, it is assumed that there is a demand $b_{n,k}$ at RSU n for service k with the average demand D_{t_i}. The mathematical model is designed to compute the optimal network configuration while minimizing the host numbers and RSU cloud delay during period t_i, under the premise of ensuring a certain amount of service resources redundancy. In this case, when some over-burdened RSUs send part of demands to other RSUs with less demands, the demands in the cloud are all met with a load-balanced network.

Defining the $delay$ as the sum of the delays in all edges that a single routing path traverses from RSU m to n, which includes processing time T_P, queuing time T_q, transmission time T_{tr}, and propagation time T_{pg}, processing time T_P is a constant, which is usually 10 μs. For the queuing process of a data packet, the Poisson distribution combined with the traffic background is used for modelling. For the Poisson

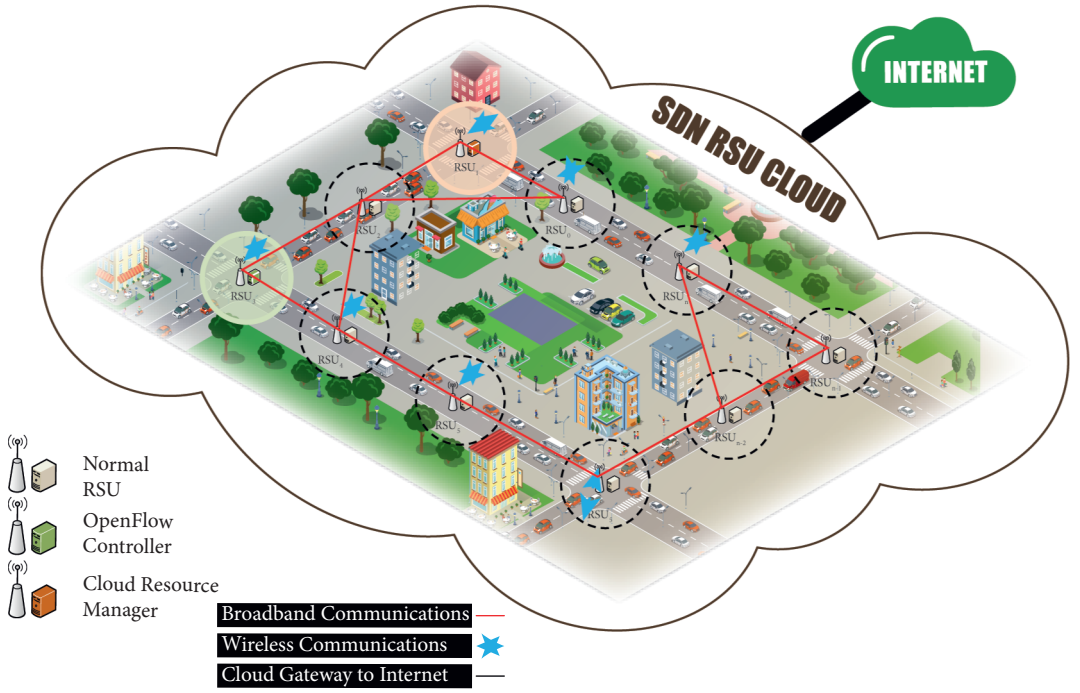


Figure 1: System architecture of the proposed SDN RSU cloud.

process with the packet interarrival times λ and packet processing times μ, whose mean and standard deviation are λ_a, σ_a and λ_p, σ_p, respectively, the coefficients $Cof_a = \sigma_a/\lambda_a$, $Cof_p = \sigma_p/\lambda_p$ are defined by the Kingsman formula to describe queuing delay T_q[23]:

$$T_q = \frac{Cof_a + Cof_p}{2} \cdot \frac{\lambda/\mu}{\mu - \lambda}. \tag{1}$$

According to the basic knowledge of physics, the transmission speed of signal in the wire equals about 2/3 light speed. Propagation time T_{pg} is determined by the size of packet divided by the capacity C_e:

$$T_{tr} = \frac{\text{length}}{(2/3) \cdot \text{light speed}}, \tag{2}$$

$$T_{pg} = \frac{\text{size of packet}}{C_e}. \tag{3}$$

As for the convenience of describing the parameters in the model, the notations of the input parameters are listed in Table 1, and output parameters and constants are depicted in Table 2, respectively.

As what can be seen in Table 1, the essential variables are all displayed with the detailed explanations and there are several binary variables which are of great significance in determining the output results. In the output parameters, there are also several variables which are corresponding to the ones in Table 1, which are listed in Table 2.

Table 2 lists the output parameters and constants, parts of which are related to some variables in Table 1. Since all the parameters needed for this paper are depicted in detail above, it is time to design a joint optimization model to find the best configuration that can satisfy the delay requirements and minimize the number of hosts simultaneously. For the implementation of the SDN RSU cloud problem, two objective functions are involved: the minimization of the delay problem, which is designed to be at a given bandwidth to optimize the delay between the RSUs and users, and the minimization of service hosts number problem, which aims to deploy the minimum number of service hosts to maximize the entire network's service capabilities. What is more, all the things would be done under the premise of ensuring a certain amount of service resources redundancy so that the network instability caused by sudden service demand increase would be prevented. The object function of the problem is formulated in equation (4) and the formula X rounds the number X to the next largest integer.

$$\min \begin{cases} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \sum_{j=0}^{\Omega_{m,n}-1} d_{m,n,j}, \\ \sum_{m=0}^{N-1} \sum_{k=0}^{S-1} \lceil h_{m,k} \cdot R \rceil. \end{cases} \tag{4}$$

In the process of solving this multiobjective function, due to the reason that the priority of the two goals is hard to be determined, two strategies are tried to control the priority for the object function. That is, the strategy that the delay takes priority of the host numbers, its worked-out results are compared to the results of the strategy that the host numbers take priority of the delay. After trying many times, finally, it is found that the priority of the two goals has almost no effect on the working-out results, and the priority of the two goals is set to equal with each other. The goal of achieving the minimum service hosts number is contradictory to the minimization of delay. By transforming the proposed multiobjective problem into a single-objective optimization problem, linear scaling technology is used to solve the proposed multiobjective problem [24]. The delay

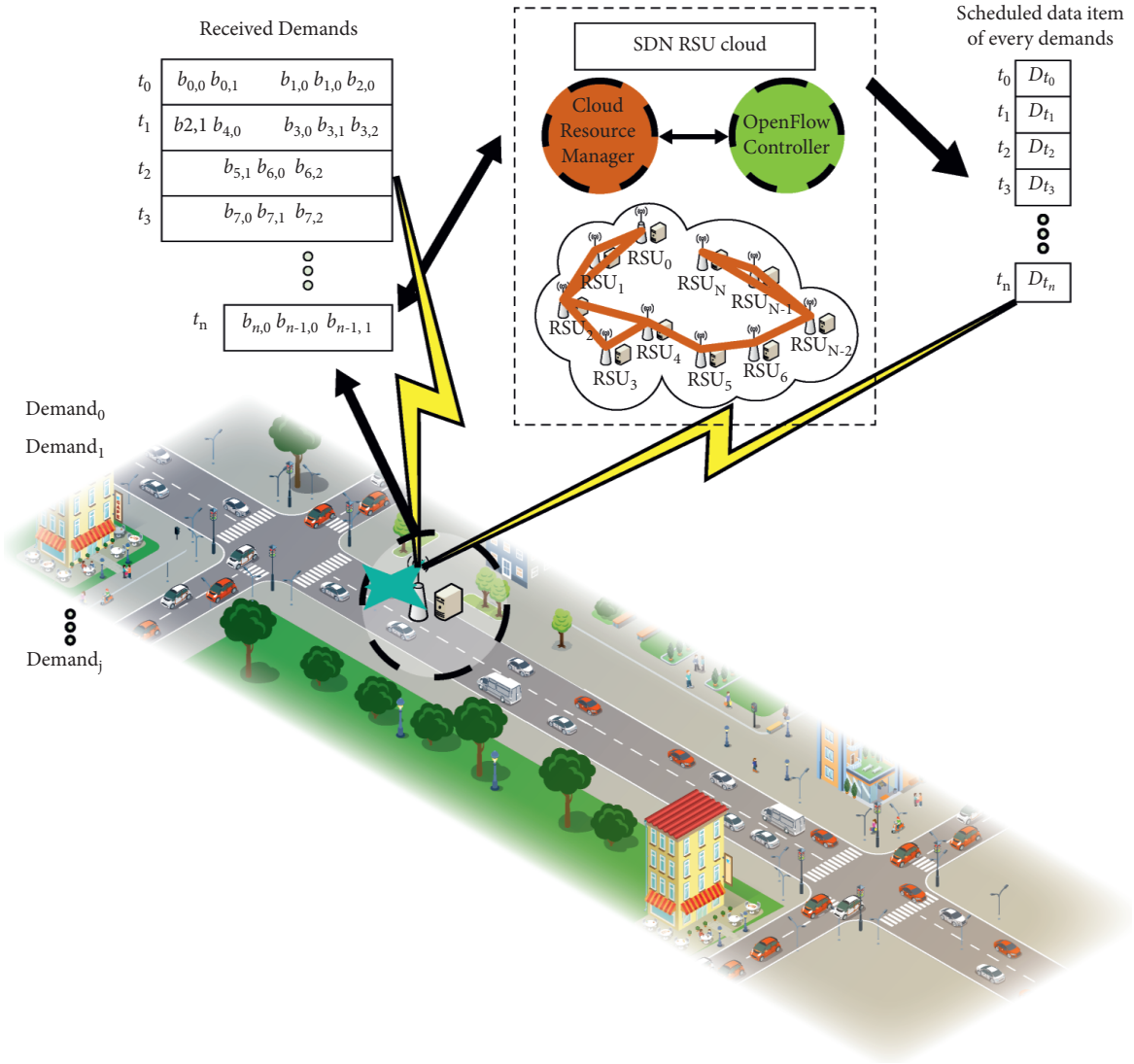


Figure 2: The mechanism of the proposed SDN RSU cloud.

requirements are firstly ensured, and then the number of hosts is minimized. Besides, the constraints of the goal (4) are listed from (5) to (22) as follows:

$$B \cdot \chi_{m,n,j} \geq h_{m,k} - \eta_{m,k}, \ \forall 0 \leq k < S, 0 \leq m < N, \tag{5}$$

$$h_{m,k} - \eta_{m,k} + \left(1 - \chi_{m,n,j}\right) \cdot B \geq 0, \ \forall 0 \leq k < S, 0 \leq m < N. \tag{6}$$

The constraints (5) and (6) restrict the definition of the variable $\chi_{m,n,j}$, where if $h_{m,k} - \eta_{m,k} > 0$, $\chi_{m,n,j}$ equals 1, otherwise 0.

$$\sum_{m,j} \beta^{k}_{m,n,j} \geq b_{n,k}, \ \forall 0 \leq k < S, 0 \leq m, n < N, \tag{7}$$

$$B \cdot h_{m,k} \cdot \zeta_{n,k} \geq \sum_{j} \beta^{k}_{m,n,j}, \ \forall 0 \leq k < S, 0 \leq m, n < N, \tag{8}$$

$$h_{m,k} \cdot \zeta_{n,k} \leq \sum_{j} \beta^{k}_{m,n,j}, \ \forall 0 \leq k < S, 0 \leq m, n < N, \tag{9}$$

$$\beta^{k}_{m,n,j} \leq B \cdot p^{k}_{m,n,j}, \ \forall 0 \leq k < S, 0 \leq m, n < N, 0 \leq j < \Omega_{m,n}, \tag{10}$$

$$\beta^{k}_{m,n,j} \geq p^{k}_{m,n,j}, \ \forall 0 \leq k < S, 0 \leq m, n < N, 0 \leq j < \Omega_{m,n}. \tag{11}$$

The general communication processes can be concluded as if demands exist, there must be loads. As a matter of fact, the RSU servers must meet or exceed the demand of the vehicles, listed as (7). Constraints (8) and (9) ensure that the routes carrying the network loads are only between the RSU servers and the users. To step further, (10) and (11) ensure that there are enough loads for carrying the rule of the control plane which is the natural cost of the system itself.

$$l_e = \sum_{m,n,j,k} \alpha^{e}_{m,n,j} \cdot \beta^{k}_{m,n,j}, \ \forall 0 \leq e < |\text{Edges}|, \tag{12}$$

$$l_e = \xi \cdot \tau_e, \ \forall 0 \leq e < |\text{Edges}|, \tag{13}$$

Table 1: Input parameters.

Input	Description
S	Number of services.
N	Number of RSUs.
$b_{n,k}$	The demand at RSU n for service k, $\forall 0 \le n < N, 0 \le k < S$.
$\zeta_{n,k}$	$\zeta_{n,k} = 1$ if a demand happens on RSU n for service k; otherwise, $\zeta_{n,k} = 0$, $\forall 0 \le n < N, 0 \le k < S$.
$\Omega_{m,n}$	Number of paths from RSU m to RSU n, $\forall 0 \le m, n < N$.
$\alpha^e_{m,n,j}$	1, if RSU m to RSU n uses path j with edge e; 0, otherwise. $\forall 0 \le m, n < N$, $0 \le j < \Omega_{m,n}$, $0 \le e < \lvert Edges \rvert$.
C_e	Bandwidth capacity of edge e, 100 Mbps is given here, $0 \le e < \lvert Edges \rvert$.
ϖ_m	RSU m's threshold on number of service hosts, $\forall 0 \le m < N$.
ϖ_k	Service k's threshold on number of hosts, $\forall 0 \le k < S$.
φ_k	Delay threshold for service k, $\forall 0 \le k < S$.
ξ	For controlling the granularity of the lookup table, $\forall 0 \le \xi < C_e$.
$\theta_{i,e}$	Load i's delay on edge e, $\forall 0 \le i < C_e$, $0 \le e < \lvert Edges \rvert$.
$\eta_{m,k}$	1, if service k was hosted on RSU m in the past; 0, otherwise. $\forall 0 \le k < S$, $0 \le m < N$.
$\pi^k_{m,n,j}$	1, if a control plane existed for service k on path j between RSUs m and n; 0 otherwise. $\forall\ 0 \le k < S$, $0 \le j < \Omega_{m,n}$, $0 \le m, n < N$.

Table 2: Output parameters and constants.

Output and constants	Description
$h_{m,k}$	1, if service k is hosted on RSU m; 0, otherwise. $\forall\ 0 \le k < S, 0 \le m < N$.
$\beta^k_{m,n,j}$	The load for service k on path j from RSU m to RSU n, $\forall 0 \le k < S, 0 \le m, n < N$, $0 \le j < \Omega_{m,n}$.
$p^k_{m,n,j}$	1, if a control plane now exists for service k on path j between RSUs m and n; 0, otherwise. $\forall 0 \le k < S$, $0 \le j < \Omega_{m,n}$, $0 \le m, n < N$.
l_e	The load on edge e, $0 \le e < \lvert Edges \rvert$.
d_e	The delay on edge e, $0 \le e < \lvert Edges \rvert$.
τ_e	The load on edge e is mapped to a bunch of ξ.
$F_{e,i}$	1, if the load on edge e equals to $i \cdot \xi$; 0, otherwise. $\forall 0 \le i < C_e, 0 \le e < \lvert Edges \rvert$.
$d_{m,n,j}$	The delay on path j from RSU m to n, $0 \le j < \Omega_{m,n}, 0 \le m, n < N$.
$\chi_{m,n,j}$	1, if $h_{m,k} - \eta_{m,k} > 0$; 0, otherwise. $\forall 0 \le m, n < N, 0 \le k < S$.
B	A large constant.
R	A constant, usually greater than or equal to 1. The parameter size can be adjusted according to different scenarios to ensure a certain amount of service resource redundancy.

$$\tau_e = \sum_{i=0}^{(C_e-1)/\xi} i \cdot F_{e,i},\ \forall 0 \le e < |\text{Edges}|, \tag{14}$$

$$\sum_{i=0}^{(C_e-1)/\xi} \mathscr{F}_{e,i} = 1,\ \forall 0 \le e < |\text{Edges}|, \tag{15}$$

$$d_e = \sum_{i=0}^{(C_e-1)/\xi} \mathscr{F}_{e,i} \cdot \theta_{i,e},\ \forall 0 \le e < |\text{Edges}|. \tag{16}$$

Equation (12) describes the loads on the *edges* of the flow between RSU m and RSU n, which contains the generating and forwarding traffic on the routes. The delay on the edge matches with the load on it, where the index is used to find it in the lookup table (LUT). Equation (13) represents that the load on an edge e is mapped to a bunch of ξ. Constraint (14) ensures that the traffic load on an edge cannot exceed its capacity, where (15) and (16) state that there is only one specific value for the traffic load on edge e and the delay corresponding to it.

$$d_{m,n,j} = \sum_{e=0}^{|\text{Edges}|-1} \alpha^e_{m,n,j} \cdot d_e,\ \forall 0 \le j < \Omega_{m,n}, 0 \le m, n, m \ne n < N, \tag{17}$$

$$d_{m,n,j} \cdot h_{m,k} \cdot \zeta_{n,k} \le \varphi_k,\ \forall 0 \le j < \Omega_{m,n}, \quad 0 \le m, n, m \ne n < N, 0 \le k < S, \tag{18}$$

$$d_{m,n,j} \cdot h_{m,k} \le d_{m,n,j},\ \forall 0 \le j < \Omega_{m,n}, 0 \le m, n, m \ne n < N, 0 \le k < S. \tag{19}$$

Moreover, the delay of the specific path between the RSU m and n equals the summation of all the delays on its edges, which is shown in (17). Constraint (18) represents the service k's delay threshold, while constraint (19) describes whether a routing path is being used or not.

$$\sum_{m=0}^{N-1} h_{m,k} \ge 1,\ \forall 0 \le m < N, 0 \le k < S, \tag{20}$$

$$\sum_{m=0}^{N-1} h_{m,k} \le \varpi_k,\ \forall 0 \le m < N, 0 \le k < S. \tag{21}$$

On the other side, constraint (20) ensures that there is at least one host working on the RSU m. What is more, constraint (21) represents service k's threshold on number of service hosts. Salahuddin et al. [19] assume that all RSUs have unlimited resources to meet any service demand for simplicity. However, it would be hard and nearly impossible to make the RSUs working without resource restriction.

To further figure out this problem, it is necessary to take the number of working hosts for a single RSU into consideration, which is usually limited.

$$\sum_{k=0}^{S-1} h_{m,k} \leq \varpi_m, \ \forall 0 \leq m < N. \tag{22}$$

The maximum number of hosts for a single RSU is limited, which is reflected in constraint (22). Besides, formula (22) is used to present the resource constraints in sharp contrast to the work of Salahuddin et al. [19], where they assume that the RSU's resource is infinite to hold the services.

4.2. Hardness Analysis. In this section, the complexity of the proposed model is analysed via testing the complexities of the two terms of it. For the minimization of the delay problem in the implementation of SDN RSU cloud, all OpenFlow switches utilized to route the traffic loads in the network are multipath enabled.

Theorem 1. *The minimization of the delay problem is NP-hard.*

Proof 1. The minimization of the delay problem can be converted into the minimum multicommodity flow (MCF) problem [25]. For an example of the MCF, where a set of nodes Nd exists, which contains k elements $\{Nd_0, Nd_1, Nd_2, \ldots, Nd_k\}$ and an edge set $edges$ with m elements $\{e_0, e_1, e_2, \ldots, e_m\}$, there is a total of K traffic demand, and each traffic demand contains a source node m and a destination node n. A route exists from source m to the destination n whichever paths it chooses when multipath is enabled. The delay on each edge in the network topology is d_e. A graph is therefore formed by all the nodes and edges, where each node Nd in the network is considered as a point in the graph. The delay of a path corresponds to the sum of the delay on every edge it traverses.

It is easy to prove that there is a minimum path delay if and only if there is an optimal solution to the MCF problem via mathematic construction. The construction can be done in polynomial time and the MCF problem is an NP problem [25]. Therefore, minimization of the delay problem is also an NP-hard problem.

To ensure the best reachability and reliability of the entire network, it is best to enable all the RSUs to work on with the threshold number of hosts simultaneously. However, it is wasteful and not acceptable for the users nor for the providers from a pure cost of money perspective, since high operating cost would cost a lot of money. The minimization of service hosts number problem, which is aiming at deploying the minimum number of service hosts without exceeding the threshold while serving all flows in the network. □

Theorem 2. *The minimization of service hosts number problem is NP-hard.*

Proof 2. The minimization of service hosts number problem can be converted into the minimum weighted vertex coverage (MWVC) problem. Here it shows an example of the MWVC problem, containing a set of nodes Nd, which has k elements $\{Nd_0, Nd_1, Nd_2, \ldots, Nd_k\}$ and an edge set $edges$ with m elements $\{e_0, e_1, e_2, \ldots, e_m\}$. These nodes have corresponding weights $\{w_0, w_1, w_2, \ldots, w_k\}$. All nodes and edges form a graph. Here, it treats the network flow set F as a set of edges $edges$ and takes each vehicle or RSU in the network as the node Nd_i in the graph. Therefore, the flow corresponding to the edges passes from the origin Nd_o connected to the destination Nd_d, and the hosts number of each traffic flow corresponds to the weight w_d of destination Nd_d.

Through mathematic construction, it is easy to prove that only and only if there is an optimal solution to the MWVC problem, the SDN RSU cloud can serve all flows with the minimum number of hosts service. The construction can be done in polynomial time, and the minimum weighted vertex cover problem is NP-hard to solve [25]. Hence, the minimization of service hosts number problem is also an NP-hard problem.

Furthermore, $d_{m,n,j}$ is a continuous variable that is larger than 0, while $h_{m,k}$ is an integer variable. Therefore, the proposed model is a mixed-integer linear programming problem (MILP). The MILP problem is solved with the Python-programmed simulation framework which is mixed programmed with Gurobi. All the work done above for minimizing the delay and number of hosts simultaneously in the proposed model is called for short in the abbreviation form JO which means joint optimization algorithm. The single-objective algorithms which are host optimization (HO) and delay optimization (DO), respectively, are also analysed. HO hosts the service in the best way to saving the number of hosts in meeting the network demand, regardless of the delay caused by the service; DO seamlessly hosts the best service to minimize delay, without considering the number of hosting services. □

5. Results and Discussion

In this section, the ITS application scenery in the simulation scenario is the coverage area of fixed RSUs in the downtown area of Longhua District, Shenzhen, China, where there are totally ten RSUs ($N = 10$) in the topology, as shown in Figure 3.

The network topology shown above represents the possible paths among the ten RSUs which can be utilized to configure the allocation of the cloud resources when it comes to matching the demands of the vehicles. The left part of Figure 3 is the scenario map of the simulated area, while the right part describes the detailed topology of the designed SDN RSUs network. In Figure 3, each RSU is decomposed into a OpenFlow switch and a server, the RSU_0 consists of s_0 and h_0, so for the rest RSUs. It is assumed that the paths represented in a list from origin RSU m to destination RSU n are acyclic elementary chains, where the labelled notations are from 0 to 9. For the topology of the RSU cloud, it is easy to work out that there are 500 possible paths from origin RSU m to destination RSU n, $\forall 0 \leq m, n < N = 10$ in summary. From what is shown in Figure 3, considering a confined geographical region, vehicles are entering and leaving the region continuously and the RSUs are deployed along the roads. Taking an example of the flow

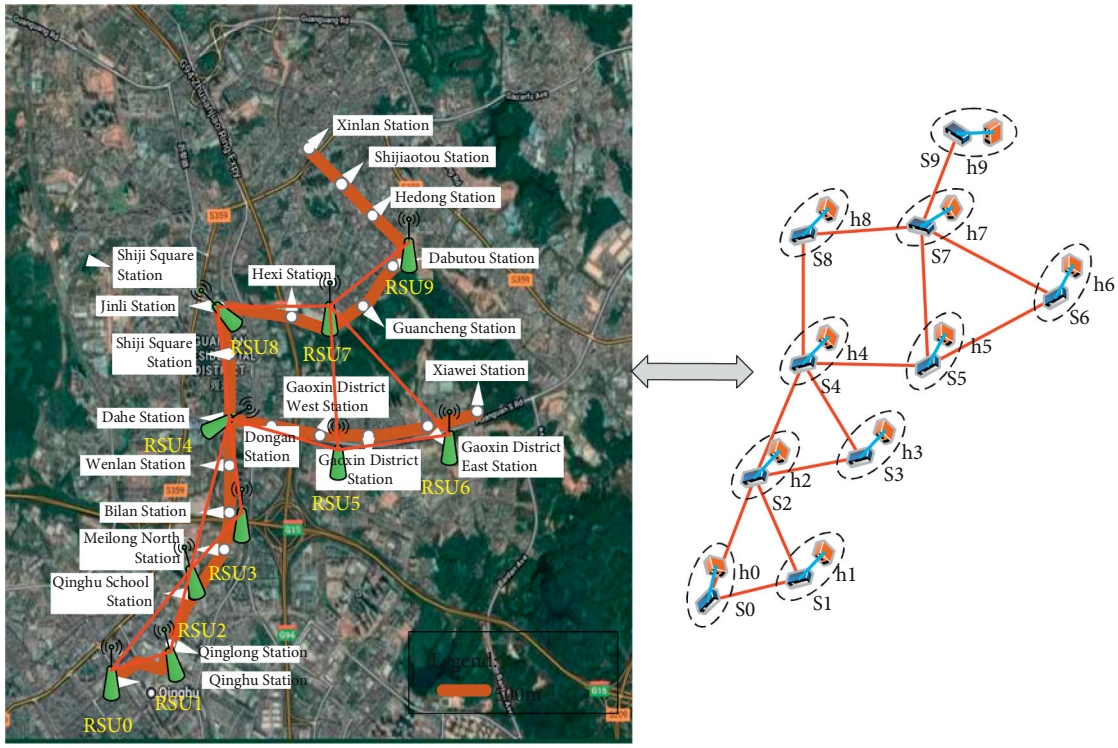


Figure 3: Detailed SDN RSU cloud topology.

$(S_0 - h_0) \longrightarrow (S_9 - h_9)$, which means the traffic flow from RSU_0 to RSU_9, there are totally 12 possible different paths for the flow: $(0, 9, Path_0)$: [0, 1, 2, 3, 4, 5, 6, 7, 9], $(0, 9, Path_1)$: [0, 1, 2, 3, 4, 5, 7, 9], $(0, 9, Path_2)$: [0, 1, 2, 3, 4, 8, 7, 9], $(0, 9, Path_3)$: [0, 1, 2, 4, 5, 6, 7, 9], $(0, 9, Path_4)$: [0, 1, 2, 4, 5, 7, 9], $(0, 9, Path_5)$: [0, 1, 2, 4, 8, 7, 9], $(0, 9, Path_6)$: [0, 2, 3, 4, 5, 6, 7, 9], $(0, 9, Path_7)$: [0, 2, 3, 4, 5, 7, 9], $(0, 9, Path_8)$: [0, 2, 3, 4, 8, 7, 9], $(0, 9, Path_9)$: [0, 2, 4, 5, 6, 7, 9], $(0, 9, Path_{10})$: [0, 2, 4, 5, 7, 9], and $(0, 9, Path_{11})$: [0, 2, 4, 8, 7, 9].

5.1. Parameter Settings. To analyse the experimental results in a representative way, the number of services is set to $S = 1, 2$, which typically represent the single-threading and the multithreading instances, that is, the simple and complex scenario for services, respectively. During period t_1, the average demand D_{t_1} is composed of a multiple of demands $b_{n,k}.\zeta_{n,k}$ at node n for service k. Since that there are many cases in nature that conform to the normal distribution, it can describe a common distribution of samples affected by a large amount of small random disturbances. In this case, the normal distribution is adopted here, similar to the experiments in [19]. For each group of the experiments, the normal distribution is applied with the configuration that the mean value equals to the D (demand), and the standard deviation of the distribution equals to 0.05 * demand. Under this circumstance, one moment of $b_{n,k}.\zeta_{n,k}$ is taken out with $S = 1, 2$, respectively, at 50 and 60 Mbps on behalf of all the average demands ranging from 0 Mbps to 99 Mbps.

In the preliminary experiment, the size of one single packet for the test experiment is 1240 bytes and multiple times are tried with a bunch of same service demands ranging from 0 Mbps to 99 Mbps for each service until there is no feasible solution and $\zeta_{n,k}$ is set to 1 or 0 randomly with the random function of NumPy [26]. In this way, the demands for the services are around D or 0 otherwise, which means the demand differences among the 10 RSUs are great and the SDN RSU cloud resources can therefore be managed.

For $S = 1$, 86 Mbps is the largest flow demand with feasible solution and 97 Mbps is the largest flow demand with feasible solution for $S = 2$, where the entire SDN RSUs cloud is working in saturating state. As a matter of fact, when the flow demands are lower than 30 Mbps, the results of delay and number of hosts change to a small extent, and this is due to the reason that the flow demands are too small. Furthermore, the model is in a multiple dimension when it comes to the circumstance of $S \geq 2$. While $S > 2$ brings the model with nothing but making the network configuration more complex and hard to show the result analysis, therefore $S = 2$ is set when it comes to the circumstance of complex scenario for services. Seven different experimental flow demands $D = \{40$ Mbps, 50 Mbps, 60 Mbps, 70 Mbps, 80 Mbps, 90 Mbps, 97 Mbps$\}$ are set for $S = 2$, and $D = \{30$ Mbps,40 Mbps, 50 Mbps, 60 Mbps, 70 Mbps,80 Mbps, 86 Mbps$\}$ for $S = 1$. All the initial network configuration parameters are shown in Table 3, which depicts all the numerical parameters that have impacts on the experimental results in detail.

As what can be shown in Table 3, the bandwidth capacity of each edge is set to 100 Mbps, and RSU's threshold on number of service hosts is 5. Besides, for one service, its

Table 3: Initial network configuration parameters.

Parameters	Values
C_e	100 Mbps
ϖ_m	5, $\forall 0 \le m < N$.
ϖ_k	5, $\forall 0 \le k < S$.
φ_k	60 ms, $\forall 0 \le k < S$.
ξ	1
$\pi^k_{m,n,j}$	0, $\forall 0 \le k < S$, $0 \le j < \Omega_{m,n}$, $0 \le m, n < N$.
R	1.5

threshold on number of hosts is also set to 5. For the convenience of splitting the network load, the ξ is set to 1, which can split the load into smallest pieces. The initial network configuration assumes that there is no control plane in the network. For the redundancy of the network, R is set to 1.5.

5.2. Experimental Results. In this section, the experiment data are worked out and the results of the cloud delays and hosts numbers for all algorithms are analysed, respectively. To begin with, for the topology of the RSU cloud, the sum of cloud delay refers that all the delays summed for 500 paths from origin RSU m to destination RSU n, $\forall 0 \le m, n < N = 10$ during period t_1.

For $S = 1$, the hosts numbers of joint optimization and delay optimization are all the same with the demands arising from 30 to 86 Mbps. What is more, for $S = 1$, the sum of cloud delay of JO and DO is almost the same as shown in the right part of Figure 4. The above results show that when the service is set to 1, the joint optimization and delay optimization algorithm are the same in essence. On the other side, when compared with the host optimization algorithm, though the number of hosts is less than the other two algorithms, the sum of cloud delay increases explosively.

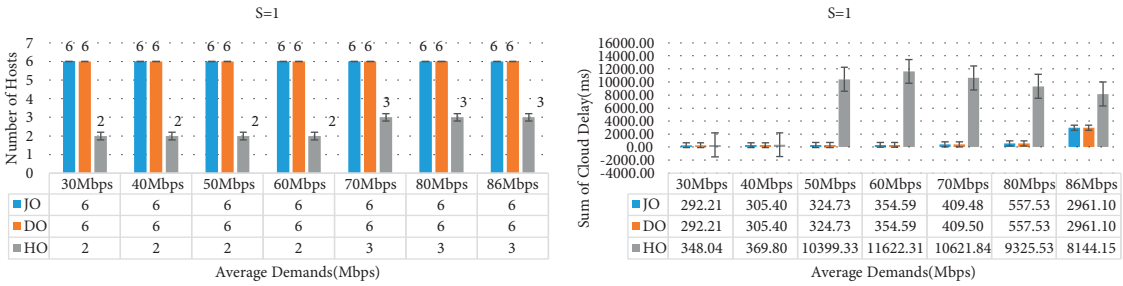
As shown in the left part of Figure 4, when the average demands are below 60 Mbps, the number of hosts can be set as only 2 in HO, greatly less than 6 in JO and HO, which reduces 66.67% number of hosts. What is more, when the average demands are below 80 Mbps, the sums of cloud delay in JO and DO are all much less than HO. When the average demand comes to 86 Mbps which are the maximum bearable demands for each service in the SDN RSU cloud, the sums of cloud delay for all 3 algorithms increase substantially. Though the HO could handle the 86 Mbps situation with only 3 hosts, the sum of delay comes to 8144 ms which is greatly larger than 2961.10 ms in JO and DO. To step further, the HO algorithm represents the minimum number of hosts needed for the SDN RSU cloud.

$S = 2$ represents the complex scenario for services, and the joint optimization algorithm would distinguish from the delay optimization algorithm and host optimization algorithm which is different from the circumstance of $S = 1$. From what can be seen in Figure 5, when the average demands are below 90 Mbps, the JO could achieve near the same sum of cloud delay as DO with the same or even less hosts. Taking 40 Mbps as an example, the sum of cloud delay in JO is 319.40 ms which is only 2.10 ms greater than DO's while the number of hosts in JO is 8 which is 3 less than DO whose number of hosts is 11. For the 40 Mbps circumstance, the delay of JO only increases 0.66% than DO, but the number of hosts in JO decreases 27.27% than JO, which is of great significance for the improvement of the overall cloud efficiency.

From Figure 5, it can be observed that the sum of cloud delays rises at an explosive speed and reaches to over 8000 ms when the network demand arrives at 97 Mbps. The reason why delay grows up drastically when communication demand grows up slightly from 90 Mbps to 97 Mbps accounts for the transfer of the network status. When the network demand is 90 Mbps, the network status is far away from being saturated. However, network status becomes saturated when the network demand comes to 97 Mbps. If the clients and servers are connected with the limited 100 Mbps bandwidth, but there are multiple applications that need to transmit data approximating the bandwidth at the same time, a large amount of data will be waiting in a queue at this time, which makes the delay increase greatly. Furthermore, it can be easily seen that the sum of cloud delays is rising with the demands for the JO and DO algorithms which is in accordance with the normal recognition. As the average demands for each service arise from 40 to 97 Mbps, the number of hosts is rising in HO. However, things get a little bit complex when it comes to the sums of cloud delay in HO. The sum of cloud delay at 50 Mbps in HO is 31.26% larger than 60 Mbps in HO, while the number of hosts of these two is the same. To explain the phenomena, the network configurations Z_{t_1} must be further taken out, which is computed from the HO algorithm in contrast to the initial network configurations Z_{t_0}. The network configurations of the HO algorithm in 50 Mbps and 60 Mbps are shown in Table 4.

When $S = 2$, in Figure 5, the numbers of hosts for 50 Mbps and 60 Mbps equal to 5, where there are 2 hosts running idly for back up and they do not afford any loads. This is due to the reason that $R = 1.5$ is set to ensure a certain amount of service resource redundancy and $[3 \cdot 1.5] = 5$. Therefore, the two idling hosts in Table 4 are removed for the convenience of analysing the network configurations of HO algorithm. For $S = 2$, as shown in Table 4, when the average demands arrive at 50 Mbps and 60 Mbps in HO algorithm, $h_{2,0} = 1$ and $h_{2,1} = 1$ mean there are two services hosted on RSU_2 at 50 Mbps, while the 3 hosts are distributed on different RSUs at 60 Mbps. Due to the reasons depicted above, there are only 2 RSUs working at 50 Mbps, in contrast to 3 RSUs working at 50 Mbps, leading to the fact that the sum of cloud delay at 50 Mbps is greater than at 60 Mbps with the same number of hosts in HO algorithms. What is more, when it comes to designing the RSU cloud, the above phenomenon reminds us that the services should be deployed globally in a distributed manner rather than in the centralized manner for achieving the minimized cloud delay. The services are recommended distributed rather than concentrated on some specific RSUs.

5.2.1. The Analysis of Network Configurations. Since the sum of cloud delay and number of hosts for JO, DO, and HO



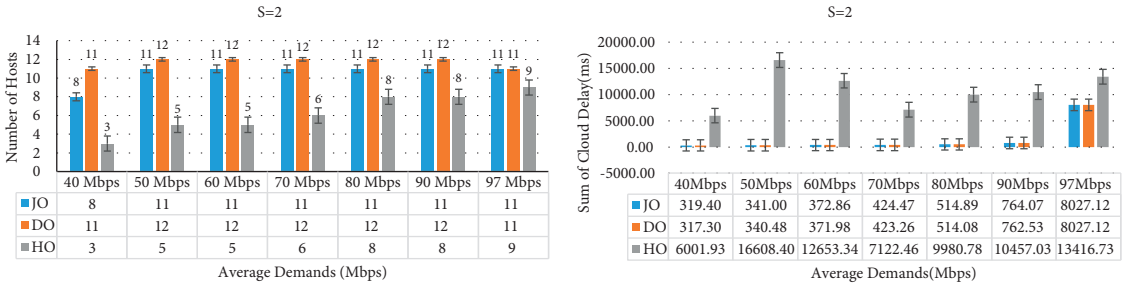
	30Mbps	40Mbps	50Mbps	60Mbps	70Mbps	80Mbps	86Mbps
JO	6	6	6	6	6	6	6
DO	6	6	6	6	6	6	6
HO	2	2	2	2	3	3	3

(a)

	30Mbps	40Mbps	50Mbps	60Mbps	70Mbps	80Mbps	86Mbps
JO	292.21	305.40	324.73	354.59	409.48	557.53	2961.10
DO	292.21	305.40	324.73	354.59	409.50	557.53	2961.10
HO	348.04	369.80	10399.33	11622.31	10621.84	9325.53	8144.15

(b)

Figure 4: Global comparison for each algorithm when $S = 1$.



	40 Mbps	50 Mbps	60 Mbps	70 Mbps	80 Mbps	90 Mbps	97 Mbps
JO	8	11	11	11	11	11	11
DO	11	12	12	12	12	12	11
HO	3	5	5	6	8	8	9

(a)

	40Mbps	50Mbps	60Mbps	70Mbps	80Mbps	90Mbps	97Mbps
JO	319.40	341.00	372.86	424.47	514.89	764.07	8027.12
DO	317.30	340.48	371.98	423.26	514.08	762.53	8027.12
HO	6001.93	16608.40	12653.34	7122.46	9980.78	10457.03	13416.73

(b)

Figure 5: Global comparison for each algorithm when $S = 2$.

Table 4: Network configurations of HO algorithm in 50 Mbps and 60 Mbps when $S = 2$.

50 Mbps	$h_{2,0} = 1$	$h_{2,1} = 1$	$h_{6,0} = 1$	$h_{m,k} = 0 (h_{m,k} \neq h_{2,0}, h_{2,1}, h_{6,0}, \forall 0 \leq m \leq 9, 0 \leq k \leq 1)$
60 Mbps	$h_{1,0} = 1$	$h_{2,1} = 1$	$h_{6,0} = 1$	$h_{m,k} = 0 (h_{m,k} \neq h_{1,0}, h_{2,1}, h_{6,0}, \forall 0 \leq m \leq 9, 0 \leq k \leq 1)$

Table 5: Network configurations of JO, DO, and HO algorithms in 60 Mbps when $S = 1, 2$.

Instances ($S = 1$)	JO		DO		HO		Instances ($S = 2$)	JO		DO		HO	
	50	60	50	60	50	60		50	60	50	60	50	60
$h_{0,0}$	0	0	0	0	0	0	$h_{0,0}$	0	0	0	0	0	0
$h_{1,0}$	0	0	0	0	0	0	$h_{0,1}$	0	0	0	0	0	0
$h_{2,0}$	1	1	1	1	0	0	$h_{1,0}$	1	1	1	1	0	1
$h_{3,0}$	0	0	0	0	0	0	$h_{1,1}$	0	0	0	0	0	0
$h_{4,0}$	1	1	1	1	1	1	$h_{2,0}$	0	0	0	1	1	0
$h_{5,0}$	1	1	1	1	0	0	$h_{2,1}$	1	1	1	1	1	1
$h_{6,0}$	1	1	1	1	0	0	$h_{3,0}$	1	1	1	1	0	0
$h_{7,0}$	0	0	0	0	0	0	$h_{3,1}$	0	0	0	0	0	0
$h_{8,0}$	0	0	0	0	0	0	$h_{4,0}$	0	0	0	0	0	0
$h_{9,0}$	0	0	0	0	0	0	$h_{4,1}$	0	0	0	0	0	0
							$h_{5,0}$	0	0	0	0	0	0
							$h_{5,1}$	0	0	0	0	0	0
							$h_{6,0}$	1	1	1	1	1	1
							$h_{6,0}$	0	0	0	0	0	0
							$h_{7,0}$	0	0	0	0	0	0
							$h_{7,1}$	1	1	1	1	0	0
							$h_{8,0}$	1	1	1	0	0	0
							$h_{8,1}$	0	0	1	1	0	0
							$h_{9,0}$	1	1	1	1	0	0
							$h_{9,1}$	0	0	0	0	0	0

algorithms in 60 Mbps when $S = 1, 2$ are already displayed in Figures 4 and 5, here the idling hosts are removed for the convenience of analysing the network configurations. To further demonstrate the reasons why there are different sums of cloud delay for the 3 different optimization algorithms, the working hosts' configurations are shown in detail at Table 5.

From what can be seen in the left part of Table 5, the network configurations of JO and DO algorithms in 50 and 60 Mbps when $S = 1$ are exactly the same, which further verifies that these two algorithms are the same in essence when $S = 1$. Meanwhile, when $S = 2$ at 50 Mbps, JO algorithm needs 1 less host than DO algorithm where their other host network configurations are the same, except for the $h_{8,1}$. That is, while the $h_{8,1}$ in JO is 0, the $h_{8,1}$ is set to 1 in DO. Under this circumstance, the sum of cloud delay in JO is 341.00 ms while 340.48 ms in DO, showing that JO could achieve almost the same sum of cloud delay as DO even with 1 less host. For the HO results compared with the JO and DO, though the number of hosts is reduced to a great extent, the sum of delay is much larger than JO and DO and unsuitable for use in the application. Still, it gives us the emergency resources management solution in the SDN RSU cloud when the network congestion happened.

6. Conclusions and Future Works

To reduce environmental pollution, achieving a full range of safe, efficient, and green transportation environment, VANETs have been regarded as a key enabling technology for improving the road safety and building intelligent transportation systems. VANETs can provide safety and nonsafety-related applications and services to improve the safety and shared mobility of vehicle users. It is a multidisciplinary technical field involving road traffic, wireless communication, self-organizing systems, etc. The cloud resource management system realizes the interoperability of resource data in the VANETs. It can integrate similar or related resources through resource integration services to reduce the transmission of redundant data. On this basis, through resource data distribution planning services, according to the priority of the resource, under the condition of limited bandwidth, the priority transmission of key data is realized.

In this paper, the deep programmability of SDN is used to dynamically reconfigure the network for hosting, routing the services, and effectively meeting the basic requirements of the VANETs. Multipath is also enabled to forward data streams to balance the loads on the network. To achieve the minimum cloud delay with the least number of hosts, it can be summarized as a mixed-integer linear programming problem (MILP). Based on the research of the architecture technology in cloud resource management system, a cloud resource management system for vehicle collaborative perception scenarios is realized, and the feasibility of the system is verified. When the service is set to 1, which typically represents the single-threading instance scenario, the common optimization and delay optimization algorithms are essentially the same. Besides, the results show that, given the requirements, when the services are larger or equal to 2, the proposed algorithm is superior to single-objective optimization algorithms, namely, the delay optimization (DO) algorithm and the host optimization (HO) algorithms. In addition, the results also indicate that services should be distributed on a global scale, rather than provide services centrally.

For future work, the proposed algorithm can be used to the shared mobility services in connected and automated environment to improve the safety. The inclement road conditions can have impact on the traffic safety, level of service, traffic mobility, and fuel efficiency of the vehicle flows [27], and its communication mode can be further studied and combined with our proposed model in the future research. Besides, the cloud resource management system could be enhanced by comparing it with the centralized resource management system under the scenario of unmanned vehicle collaborative perception. Furthermore, in terms of RSU access, more complex scenarios can be considered in follow-up research, such as overlapping coverage of multiple RSUs, variable speed of vehicles can also be considered. Such in-depth research will be more practical. The follow-up work can further study this special circumstance and combine with other advanced algorithms to propose more effective data distribution algorithms. There are also lots of security and privacy issues and research challenges in RSU clouds.

Data Availability

The data used to support the findings of this research are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest for the publication of the paper.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (52172329) and the 13th Five-Year National Key Research and Development Plan (2018YFB1201403).

References

[1] E. K. Adanu and S. Jones, "Effects of human-centered factors on crash injury severities," *Journal of Advanced Transportation*, vol. 2017, no. 2017, 11 pages, Article ID 1208170, 2017.

[2] Z. Pu, Z. Li, Y. Jiang, and Y. Wang, "Full Bayesian before-after analysis of safety effects of variable speed limit system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 964–976, 2021.

[3] D. Sun, Y. Zheng, and R. Duan, "Energy consumption simulation and economic benefit analysis for urban electric commercial-vehicles," *Transportation Research Part D: Transport and Environment*, vol. 101, Article ID 103083, 2021.

[4] W. Li, Z. Pu, Y. Li, and M. Tu, "How does ridesplitting reduce emissions from ridesourcing? a spatiotemporal analysis in

Chengdu, China," *Transportation Research Part D: Transport and Environment*, vol. 95, Article ID 102885, 2021.

[5] M. Tu, Y. Li, W. Li, M. Tu, O. Orfila, and D. Gruyer, "Improving ridesplitting services using optimization procedures on a shareability network: a case study of Chengdu," *Technological Forecasting and Social Change*, vol. 149, Article ID 119733, 2019.

[6] D. Ou, Y. Yang, L. Xue, and D. Dong, "Optimal connectivity-based deployment of roadside units for vehicular networks in urban areas," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2559, no. 1, pp. 46–56, 2016.

[7] H. Hartenstein and K. P. Laberteaux, "A tutorial survey on vehicular ad hoc networks," *IEEE Communications Magazine*, vol. 46, no. 6, pp. 164–171, 2008.

[8] C. Campolo, A. Molinaro, A. Vinel, N. Lyamin, and M. Jonsson, "Service discovery and access in vehicle-to-roadside multi-channel VANETs," in *Proceedings of the IEEE International Conference on Communication Workshop*, pp. 2477–2482, London, UK, June 2015.

[9] K. Mershad and H. Artail, "Finding a STAR in a vehicular cloud," *IEEE Intelligent Transportation Systems Magazine*, vol. 5, no. 2, pp. 55–68, 2013.

[10] F. S. Behbehani, T. Elgorashi, and J. M. H. Elmirghani, "Power minimization in vehicular cloud architecture," pp. 1–32, 2021, https://arxiv.org/abs/2102.09011.

[11] A. Chinnasamy, B. Sivakumar, P. Selvakumari, and A. Suresh, "Minimum connected dominating set based RSU allocation for smartCloud vehicles in VANET," *Cluster Computing*, vol. 22, no. s5, pp. 12795–12804, 2019.

[12] Q. Wang, P. Fan, and K. B. Letaief, "On the joint V2I and V2V scheduling for cooperative VANETs with network coding," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 1, pp. 62–73, 2012.

[13] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility," *Future Generation Computer Systems*, vol. 25, no. 6, pp. 599–616, 2009.

[14] C. Tang, C. Zhu, X. Wei, H. Wu, Q. Li, and J. J. P. C. Rodrigues, "Intelligent resource allocation for utility optimization in RSU-empowered vehicular network," *IEEE Access*, vol. 8, pp. 94453–94462, 2020.

[15] Q. Zheng, K. Zheng, P. Chatzimisios, and F. Liu, "Joint optimization of link scheduling and resource allocation in cooperative vehicular networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, p. 179, 2015.

[16] E. Lee, E.-K. Lee, M. Gerla, and S. Oh, "Vehicular cloud networking: architecture and design principles," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 148–155, 2014.

[17] S. Olariu, M. Eltoweissy, and M. Younis, "Towards autonomous vehicular clouds," *ICST Transactions on Mobile Communications and Applications*, vol. 11, no. 7–9, p. e2, 2011.

[18] N. McKeown, T. Anderson, H. Balakrishnan et al., "OpenFlow," *ACM SIGCOMM-Computer Communication Review*, vol. 38, no. 2, pp. 69–74, 2008.

[19] M. A. Salahuddin, A. Al-fuqaha, and M. Guizani, "Software-defined networking for RSU clouds in support of the internet of vehicles," *IEEE Internet of Things Journal*, vol. 2, no. 2, pp. 133–144, 2015.

[20] M. Amadeo, C. Campolo, and A. Molinaro, "Enhancing IEEE 802.11p/WAVE to provide infotainment applications in VANETs," *Ad Hoc Networks*, vol. 10, no. 2, pp. 253–269, 2012.

[21] M. A. Salahuddin, A. Al-Fuqaha, M. Guizani, and S. Cherkaoui, "RSU cloud and its resource management in support of enhanced vehicular applications," in *Proceedings of the 2014 IEEE Globecom Work GC Workshops*, pp. 127–132, San Diego, CA, USA, December 2014.

[22] X. Meng, V. Pappas, and L. Zhang, "Improving the scalability of data center networks with traffic-aware virtual machine placement," in *Proceedings of the 2010 IEEE INFOCOM0*, pp. 1–9, San Diego, CA, USA, March 2010.

[23] G. L. Curry and R. M. Feldman, *Manufacturing Systems Modeling and Analysis*, Springer, Berlin Heidelberg, Germany, 2009.

[24] A. P. Wierzbicki, "A methodological guide to multi-objective optimization," in *Proceedings of the Optimization Techniques, Part 1*, Vol. 22 of Lecture Notes in Control and Information Sciences, K. Iracki, K. Malanowski, and S. S. Walukiewicz, eds., Springer, Berlin, Germany, 1980.

[25] D. Li, Y. Yu, W. He, K. Zheng, and B. He, "Willow: saving data center network energy for network-limited flows," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 9, pp. 2610–2620, 2015.

[26] S. Van Der Walt, S. C. Colbert, and G. Varoquaux, "The NumPy array: a structure for efficient numerical computation," *Computing in Science & Engineering*, vol. 13, no. 2, pp. 22–30, 2011.

[27] Z. Pu, C. Liu, X. Shi, Z. Cui, and Y. Wang, "Road surface friction prediction using long short-term memory neural network based on historical data," *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, pp. 1–12, 2020.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 6007485, 12 pages
https://doi.org/10.1155/2022/6007485

Research Article

Efficiency and Reliability Analysis of Self-Adaptive Two-Stage Fuzzy Control System in Complex Traffic Environment

Mingzhi Wang,[1] Xianyu Wu,[2] He Tian,[1] Jie Lin,[2] Meimei He,[2] and Liuqing Ding[2]

[1]*Laboratory for Traffic and Transport Planning Digitalization, Transport Planning and Research Institute Ministry of Transport, Beijing, China*
[2]*Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing, China*

Correspondence should be addressed to Xianyu Wu; wuxy@bjtu.edu.cn

Received 30 November 2021; Revised 24 December 2021; Accepted 11 January 2022; Published 27 January 2022

Academic Editor: Wenxiang Li

In this paper, a self-adaptive, two-stage fuzzy controller is established to realize the real-time online optimization of traffic signal timing plan, which takes multimodels of transportation as the research object to analyze the reliability of the control system at the isolated urban intersection. In this system, the first stage calculates traffic urgency degree for all red phases and selects the red phase with maximum traffic urgency degree as the next green phase. The second stage determines whether to extend or terminate the current signal phase. Aiming at the problems of the parameters of membership functions empirical settings and insufficient response to the real-time fluctuation in traffic flow, the controller introduces an improved hybrid genetic algorithm to solve it and enable the controller to self-learn. Finally, a microsimulation platform is constructed based on the VISSIM and Python language to evaluate the efficiency and reliability of the controller under complex actual traffic conditions. Results showed that the average delay time per vehicle is reduced by 14.59%, while the average number of stops per vehicle is reduced by 0.71% compared with the traditional control method. Results indicate that the traffic signal timing plan generated by the controller can efficiently improve the intersection traffic capacity and has good efficiency and reliability under the condition of medium saturation and unsteady flow.

1. Introduction

At present, with the development of shared mobility and traffic detection techniques theories, many scholars in the field of urban traffic engineering are paying more attention to the collaborative optimization of multi-transport models in traffic safety [1–3], traffic operation efficiency and reliability [4, 5], and environmental protection [6–8]. Driven by the rapid development of the mobile Internet and the sharing economy, dockless bike-sharing systems have grown in popularity in China [9]. For example, in Beijing, the total daily average of cycling trips has reached 5.6 million by 2020, up from 4.62 million in 2017, surpassing the trips by bus and subway and becoming the third-largest way of travel after walking and car [10]. In addition, in the face of congested traffic, people are more likely to choose to use dockless bike-sharing to complete short-distance travel [11]. Shared mobility has led to more walking, and cycling trips significantly changed intersections' traffic flow structure and brought new challenges to the intersection signal timing optimization. Therefore, considering the travel needs of bicycles and pedestrians in the process of signal timing optimization has become an important research content. At the same time, with the emergence of new detection technologies such as microwave radar and high point video, it is possible to analyze the efficiency and reliability of signal timing control systems in the real environment.

In some developing countries, such as China and India, the most signalized intersection' control mode is fixed-time or actuated control due to complex traffic flow and lack of effective observation data [12]. As the fixed-time control cannot respond to the real-time change of traffic demand, vehicle actuated control is better than fixed-time control.

Though it can optimize the timing scheme of signalized intersections according to the real-time traffic demand, the setting of critical parameters relies on urban traffic engineers' personal experience [13]. In the past few decades, to address these deficiencies, a variety of adaptive signal control systems have been proposed by Gartener [14], Lin and Cooke [15], Kronborg and Davidson [16], and Lin and Vijayakumar [17]. Because of its excellent real-time and robustness, fuzzy control has become one of the most common intelligent control methods.

Fuzzy control was introduced into the traffic control field in 1977 by Pappis and Mamdani [18]. Since then, single-level, two-level, and three-level fuzzy controllers have been proposed one after another [19–22]. However, as the parameters are pre-defined by experts and unable to change, this fuzzy controller is non-responsive to the real-time fluctuation of traffic demand and disability to learn [23]. To solve the problems above, Henry et al. [24], Bingham [25], Rida et al. [26], and Xu et al. [27] introduce heuristic algorithm, such as reinforcement learning, neural network, and ant colony to optimize fuzzy controller parameters according to the historical traffic data. The simulation result indicates this adaptive fuzzy controller is more effective than a traditional fuzzy controller, but the performance of these traffic signal control methods needed large and valid sample sets. Genetic algorithm (GA) has been widely used with the ability of global search and does not depend on experiential knowledge and a large amount of historical traffic flow data [22, 28]. In light of this, Liu and Zuo [29] take a four-phase isolated signalized intersection as a research object and uses a genetic algorithm to optimize fuzzy rules and membership functions of the fuzzy control. Shahsavari et al. [30] take the total length of traffic queue and pedestrian queue in all phase of isolated intersection as optimization goals, construct a fuzzy multi-objective optimization model, and use a genetic algorithm to optimize the cycle time of the traffic light systems. Yu et al. [31] established a multi-objective optimization model with traffic capacity, average parking delay, and exhaust emissions as optimization goals and used genetic algorithms to optimize the signal cycle time and effectively green light time. However, in the above study, the traffic flow distribution is generally assumed to obey a particular traffic arrival distribution, which does not reflect the fluctuation of actual traffic demand. In addition, pedestrians and non-motor vehicles are also almost not taken into account. Except that, the simulation experiment designed in the existing research has the disability to reflect actual vehicle driving behavior and mixed traffic flow characteristics.

The aim of this paper was to take the travel needs of non-motor vehicles, pedestrians, and motor vehicles into account during the optimization process of signal timing scheme. In addition, this paper introduces a hybrid genetic algorithm to make the signal control strategy have the ability of autonomous learning. At the same time, to verify the efficiency and reliability of this control system, the simulation parameters of the platform are calibrated according to the actual vehicle trajectory data.

2. Research Scope and Model Structure

2.1. Experimental Test Intersection. The plan of the signalized isolated intersection is shown in Figure 1. As the intersection has through-left movement, this paper selects a split-phase to control the intersection. Since there is a dedicated right-turn lane at this intersection, right-turn traffic flow is not considered in the research.

As shown in Figure 1, each approach has one microwave radar and many cameras to detect the vehicles' status information of each lane, such as cross ID, latitude, and longitude, with an average sampling interval of 80 milliseconds. According to the trajectory data, vehicles' temporal and spatial distribution characteristics can be obtained [32], making it possible to get the real intersection traffic flow information and analyze the driving characteristics of drivers. The daily file size of high-precision track data used in the research is about 1G, and about 8 million track point records are generated. The primary data attributes are shown in Table 1.

Due to the complexity of the real environment, possible packet loss during data transmission, buffer data overflow and on-board equipment and roadside unit failure, there will be problems in the collection of high-precision track data, such as missing field values, missing track information, and change of track ID. At the same time, roadside parking, bad weather, or temporary traffic control may also cause the obtained high-precision trajectory data to fail to reflect the normal operation state of vehicles at the intersection, so the data need to be preprocessed before use. Data cleaning work has been done in a previous work [33], and hence the details are not included to avoid redundancy.

2.2. Self-Adaptive Two-Stage Fuzzy Control Algorithm. The primary process of the self-adaptive two-stage fuzzy control algorithm is as follows: (1) set a minimum green time to a phase until the remainder of the green time is 1 s; (2) calculates the red urgency and green urgency of different approaches according to the real-time motor vehicle, non-motor vehicle, and pedestrians traffic flow data; (3) select the red phase with greatest phase urgency as the next candidate phase; (4) comparing the selected red phase urgency with green phase urgency, decided whether to extend or terminate the current signal phase; (5) if the length of the green phase exceeds maximum green duration, then terminal the current signal phase, otherwise, continue; (6) according to the evaluation of the scheme, decided whether to optimize the parameters of membership functions; (7) if the simulation time does not reached, skip to (2) and continue.

2.3. Structure of Self-Adaptive Two-Stage Fuzzy Controller. Figure 2 shows the structure of the self-adaptive two-stage fuzzy controller. As we can see from the figure, the controller has two main components: (1) two-stage fuzzy controller, which includes four parts: red urgency judgment module, pedestrian urgency judgment module, green urgency judgment module, and decision module; (2) parameter optimization module based on improved hybrid genetic algorithm.

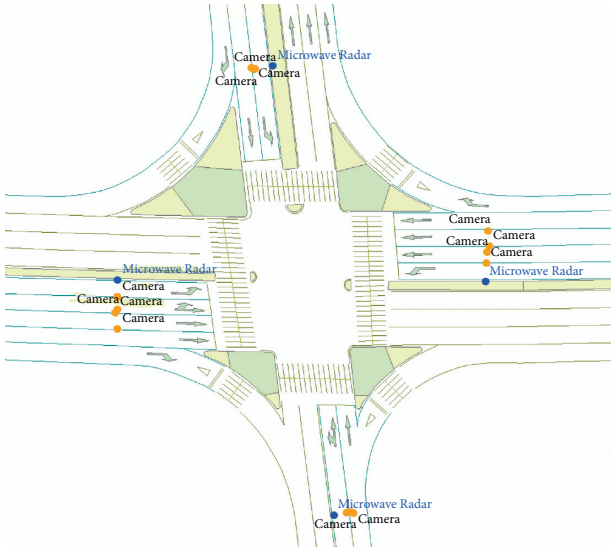


Figure 1: Intersection plan.

Table 1: Introduction of the source data.

Field name	Attribute description
Time	Unix timestamp, in millisecond
crossID	Intersection ID
Sequence	Track serial number
trackID	Track ID, unique for a short time
entryRoad	Entry road direction
exitRoad	Exit road direction
objectType	Object type: 0 unknown, 1 motor
vehicleType	Vehicle type: 1 car, 2 truck, etc
plateNumber	Plate number
N	EPSG:32650—WGS 84/UTM
E	EPSG:32650—WGS 84/UTM
Z	EPSG:32650—WGS 84/UTM

2.4. Computation of the Queue Lengths. When the phase is red, the queue lengths in the approach direction i are expressed as

$$Q_i = Q_{ri} + \sum_{j=1}^{t} q_{ij}, \tag{1}$$

where Q_i is the queue lengths of an approach direction i; Q_{ri} is the number of remaining vehicles when the phase is switched to red phase; t is the duration of the red phase of approach direction i; and q_{ij} is the amount of arriving vehicles of approach direction i duration interval.

When the phase is green, the queue lengths in the approach direction i are expressed as

$$Q_i = Q_{gi} + \sum_{j=1}^{t} q_{ij} - \sum_{j=1}^{t} q'_{ij}, \tag{2}$$

where q'_{ij} is the number of vehicles through stop-line of approach direction i duration interval and Q_{gi} is the number of remaining vehicles when switched to the green phase.

3. Design of Self-Adaptive Two-Stage Fuzzy Control

3.1. First Stage. Since the first stage has three fuzzy inputs, theoretically, the number of rules necessary for gathering all the possible input combinations for the three-term fuzzy controller is $5 \times 5 \times 5 = 125$. In practice, the design of a fuzzy controller with so many fuzzy control rules will occupy a large amount of memory and long processing time [34, 35]. Therefore, this study divided the first stage into three sub-modules modules, namely: red urgency judgment module, pedestrian urgency judgment module, and green urgency judgment module, each with two fuzzy inputs and one output variable.

In these sub-modules, phase urgency on each approach direction is determined by the maximum red/green urgency and pedestrian urgency. For example, if approach direction i is in red and its red urgency is 2.3, and pedestrian urgency is 2.2, then i would have a phase urgency of 2.3 and is named red urgency. The calculation process of the urgency of each submodule is shown below.

3.1.1. Red Urgency Judgment Module. In this module, the input variable is the queue length of motor vehicles r, the red phase duration time tr. The output variable is urgency Ur. The domain of qr is (0, 16), the linguistics are {Very Few Vehicles, Rarely Few Vehicles, Medium Vehicles, Rarely More Vehicles, Very More Vehicles}, short as {VFV, RFV, MV, RMV, VMV}. The domain of tr is (0, 120), the linguistics are {Very Short Time, Rarely Short Time, Medium Time, Rarely Long Time, Very Long Time}, short as {VST, RST, MT, RLT, VLT}. The domain of Ur is (0, 4), the linguistics are {Very Low, Low, Medium, Height, Very Height}, short as {VL, L, M, H, VH}.

With linguistic descriptions of qr, tr as input, the urgency in red direction, Ur, is determined by the fuzzy rules given in Table 2. For example, if tr is VSF, and qr is VFV, then Ur is VL.

3.1.2. Pedestrian Urgency Judgment Module. In this module, the input variable is the traffic flow of non-motor and pedestrian on each approach direction qp, the phase duration time is tp, and the output variable is urgency Up. The domain of qp is (0, 60), the linguistics are {Very Few Pedestrians, Rarely Few Pedestrians, Medium Pedestrians, Rarely More Pedestrians, Very More Pedestrians}, short as {VFP, RFP, MP, RMP, VMP}. The domain of tp, Up is (0, 120) and (0, 4), respectively. The linguistic subsets of tp, Up are same as red urgency judgment module. The rules of these modules are shown in Table 3.

3.1.3. Green Urgency Judgment Module. In this module, the input variable is the number of remaining vehicles qg, and green extension time tg. The domain and terms of each variable are shown in Table 2. The domain of qg, tg, Ug is (0, 16), (0, 120), (0, 4), respectively. The linguistic subsets of

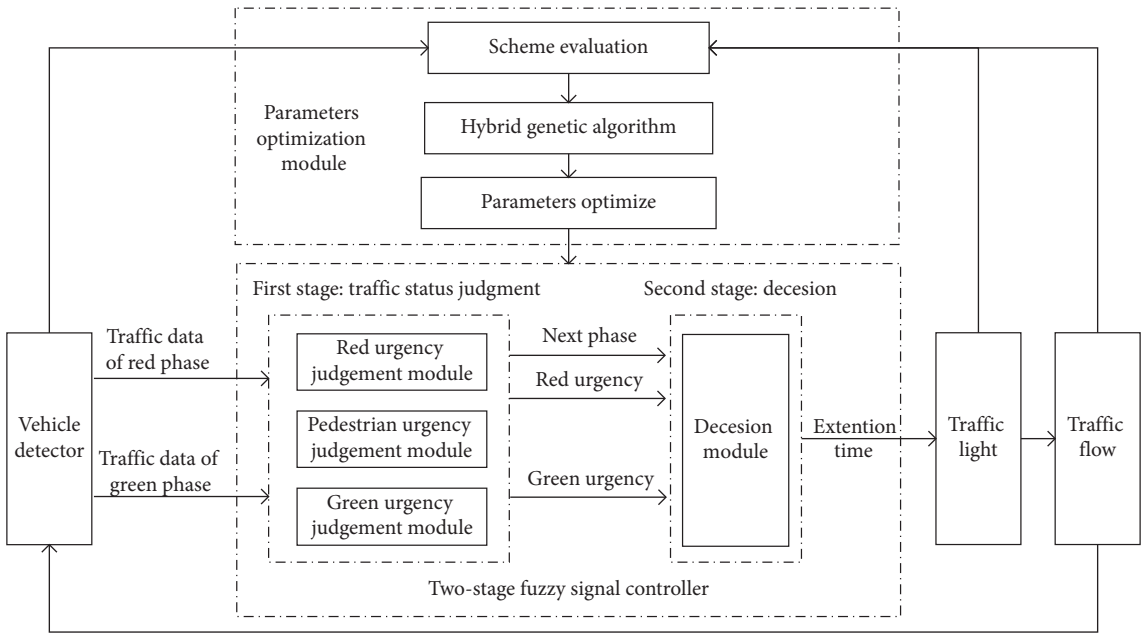


Figure 2: The diagram of adaptive two-stage fuzzy controller.

Table 2: Fuzzy rules for the red urgency judgment module.

Ur		qr				
		VFV	RFV	MV	RMV	VMV
tr	VST	VL	VL	VL	L	M
	RST	VL	VL	L	M	H
	MT	L	M	M	H	VH
	RLT	M	H	H	VH	VH
	VLT	VH	VH	VH	VH	VH

Table 3: Fuzzy rules for the pedestrian urgency judgment module.

Up		qp				
		VFP	RFP	MP	RMP	VMP
tp	VST	VL	VL	VL	L	M
	RST	VL	VL	L	M	H
	MT	L	M	M	H	VH
	RLT	M	H	H	VH	VH
	VLT	VH	VH	VH	VH	VH

each variable are the same as the red urgency judgment module. The rules of these modules are shown in Table 4.

The research uses the Gaussian fuzzy membership function to describe all parameters and uses the center of gravity method to de-fuzz output results.

$$f(x,\mu,\sigma) = e^{-\left((x-\mu)^2/2\sigma^2\right)},$$
$$Z = \frac{\int x f(x)\mathrm{d}(x)}{\int f(x)\mathrm{d}(x)}, \tag{3}$$

where μ represents the membership function center; σ denotes the membership function width; and x indicates the element; and $f(x)$ is the membership function.

Table 4: Fuzzy rules for the green urgency judgment module.

Ug		qg				
		VFV	RFV	MV	RMV	VMV
tg	VST	L	M	H	VH	VH
	RST	L	L	M	H	VH
	MT	VL	L	L	M	H
	RLT	VL	VL	L	L	M
	VLT	VL	VL	VL	L	L

3.2. Second Stage. The decision module takes Ur and Ug as an input variable, and the output variable Dc represents whether to extend the green time or deliver the right of way to the selected phase (Pr). The domain of Dc is (0, 1). It is divided into two linguistic subsets {No, Yes}, abbreviated as {N, Y}. The rules of these modules are shown in Table 5.

3.3. Parameter Optimization Module. To enable the two-stage fuzzy controller have autonomous learning ability, this paper uses a hybrid genetic algorithm to optimize the membership function of the fuzzy variable and uses a rolling horizon to set the parameters of these membership functions [28, 36]. The framework of the module is shown in Figure 3. Furthermore, the process of parameter optimization is as follows: (1) Detect real-time traffic flow during the control time horizon, evaluate the scheme, and decide whether to optimize membership functions' parameters. If so, go to step 2; otherwise, end. (2) Optimize the parameters of μ in membership function by hybrid genetic algorithm (HGA) based on historical traffic flow data. (3) Update the μ parameters of the fuzzy controller in time, according to the optimized membership function. (4) Repeat the above process until the simulation time is reached.

Table 5: Fuzzy rules for the decision module.

Dc		Ur				
		VL	L	M	H	VH
Ug	VL	Y	Y	Y	Y	Y
	L	N	Y	Y	Y	Y
	M	N	N	Y	Y	Y
	H	N	N	N	Y	Y
	VH	N	N	N	N	Y

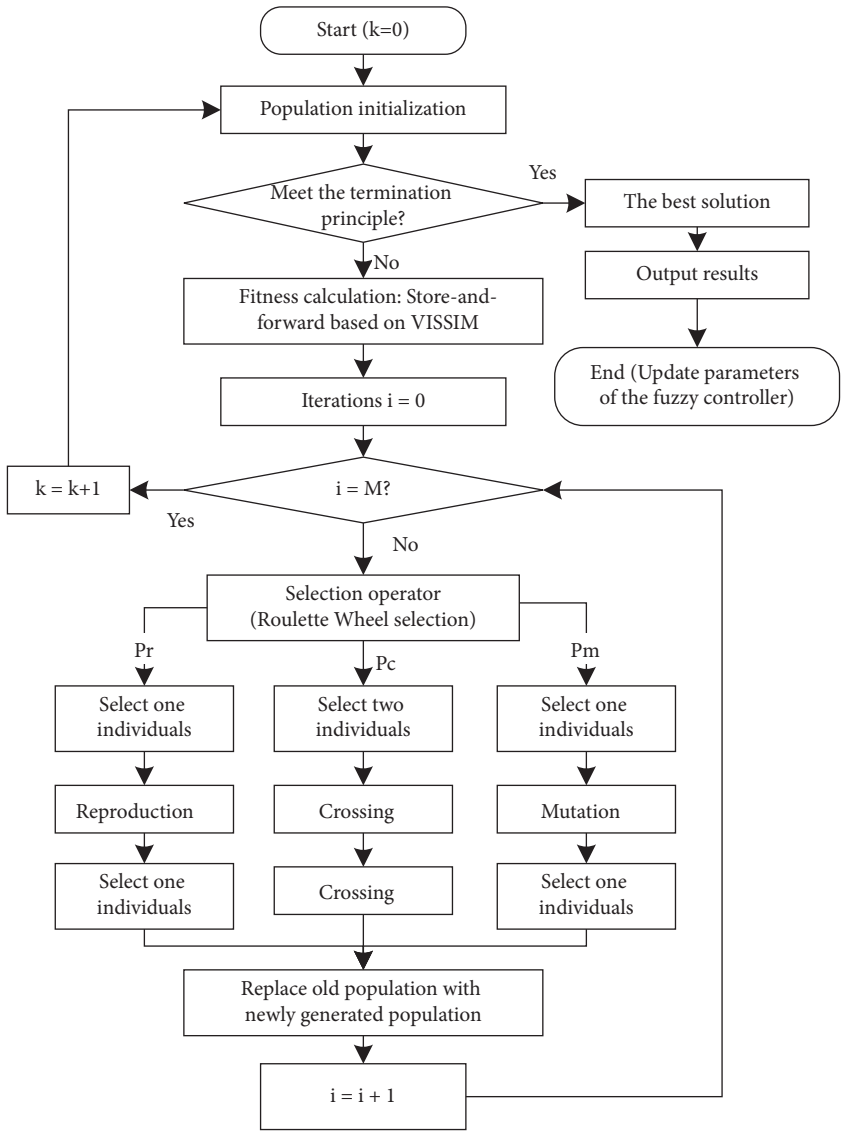


Figure 3: Framework of the parameter optimization module.

As shown in Figure 3, this study uses floating-point coding to encode the chromosome and sets a limited value range for each control parameter to avoid the generation gap phenomenon so that no fuzzy rules are activated after the parameters were input. For example, if the parameters of μ' are the k^{th} element in the linguistic sets of qr, as the domain of qr is (0, 16), so the range of μ' is $((k-1)*(16-0/5-1)-2, (k-1)*(16-0/5-1)+2)$. The fitness value of C, denoted by $f(C)$, is computed by

$$f(C) = \frac{1}{(1+d)}, \tag{4}$$

where d is the average delay time per vehicle obtained by the simulation platform.

The parameters of Pr, Pc, Pm are calculated by

$$Pr = \frac{f(i)}{\sum_{n=1}^{k} f(C)},$$

$$Pc = \begin{cases} Pc_1 \times \dfrac{1}{(Pc_1 - Pc_1) + N}, & N_1 \geq N_2, \\ 0.9, & N_1 < N_2, \end{cases}$$

$$Pm = \begin{cases} Pm_1 \times \dfrac{1}{(Pm_1 - Pm_1) + N}, & N_1 \geq N_2, \\ 0.2, & N_1 < N_2, \end{cases} \quad (5)$$

$$N = \exp\left(\frac{(N_1 - N_2)}{(N_3 - N_2)}\right),$$

where k is the number of individuals; $f(i)$ is fitness value of individual of i; Pr is select probability; Pc is crossing probability; Pc_1 is the minimum crossing probability, 0.2; Pc_2 is the maximum crossing probability, 0.9; Pm is mutation probability; Pm_1 is the minimum crossing probability, 0.01; Pm_2 is the maximum crossing probability, 0.2; and N_1, N_2, N_3 are the rank index of minimum, average, and maximum fitness, respectively.

4. Efficiency and Reliability Analysis

4.1. Simulation Environment

4.1.1. Structure of the Simulation Platform. The proposed control strategy was verified in a prevailing microscopic traffic simulation environment, VISSIM. The control strategy and optimization algorithm are compiled on the visual studio code platform and transmitted to the VISSIM via the COM interface. At each step, according to the current simulation information, the adaptive two-stage fuzzy controller decided whether to terminate or extend the current green signal phase and judgment whether to optimize the membership function or maintain the status quo. The structure of the VISSIM simulation platform is shown in Figure 4.

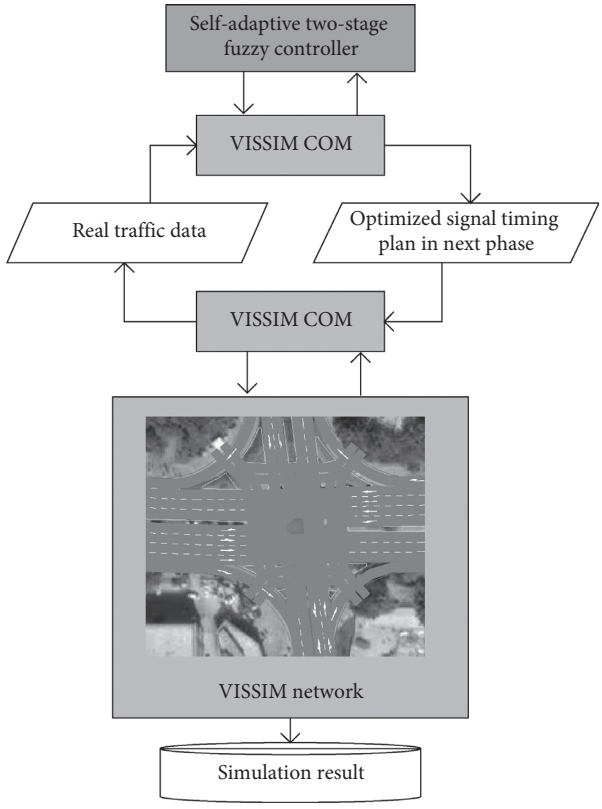


Figure 4: The structure of the VISSIM simulation platform.

4.1.2. VISSIM Parameter Calibration. VISSIM simulation software uses a large number of independent simulation parameters to reflect real vehicle operation, traffic flow characteristics, and driving behavior. However, these parameters are calibrated according to Germany and other western countries' measured data and may not be suitable for China's actual situation. As these parameters have an important impact on the simulation results, it is necessary to calibrate the default simulation parameters according to the actual situation of the intersection before the simulation. In this study, the seven parameters that significantly impact the simulation results are calibrated, including desired speed distribution, maximum acceleration/deceleration, expected acceleration/deceleration, average standstill distance, additive part, and multiple parts of safety distance [37–39]. The specific calibration process can refer to the literature [33].

4.2. Experimental Design

4.2.1. Traffic Flow. The traffic flow data of motor vehicles at the intersection extracted based on high-precision trajectory data are shown in Table 6. Data extraction work has been done in a previous work [33], and hence the details are not included to avoid redundancy.

Since trajectory data cannot obtain the traffic flow data of non-motor vehicles and pedestrians, this study assumes that the arrival rates of non-motor vehicles and pedestrians in each approach direction obey Poisson distribution [40]:

$$P(x) = \frac{\lambda t^x e^{-\lambda t}}{x!}, \quad x = 0, 1, 2, \ldots, \quad (6)$$

where $P(x)$ is the probability of x non-motor or pedestrian arrival in t interval; λ is the average arrival rate per unit time, 0–0.4; and t is phase duration of current approach direction.

4.2.2. Traffic Flow Structure. The traffic flow structure data of motor vehicles at the intersection extracted based on trajectory data is shown in Table 7.

4.2.3. Simulation Parameters Calibration. The simulation parameters are calibrated based on the intersection geometry (Figure 1) and traffic flow characteristics. The parameter calibration results are shown in Table 8.

Table 6: Traffic flow (pcu/15 minutes).

Time	WBL	WBT	SBL	SBT	NBT	NBL	EBT	EBL
0:00	6	18	6	2	5	4	35	9
0:15	3	22	4	4	3	3	26	4
0:30	9	14	1	2	2	3	16	6
...	...	...	...	...	...	...	...	...
7:30	8	134	22	0	3	16	75	39
7:45	5	199	28	5	13	14	87	60
8:00	14	236	33	14	28	16	102	85
8:15	26	245	45	13	32	26	136	100
8:30	12	206	49	16	31	16	130	132
8:45	18	230	36	12	24	20	128	110
9:00	15	181	38	17	10	11	130	54
9:15	14	161	37	13	10	17	117	41
9:30	14	147	19	10	6	24	119	40
9:45	9	156	20	5	2	9	78	29
...	...	...	...	...	...	...	...	...
19:00	1	34	15	1	2	5	22	4
19:15	7	36	19	3	0	0	45	8
19:30	4	20	16	4	3	0	2	0

In this table, WBL means westbound left, WBT means westbound through, and so on.

Table 7: Traffic flow structure.

Structure	Vehicle			Nonmotor (%)	
	Car (%)	HGV (%)	Minibus (%)	Nonmotor (%)	Pedestrian (%)
WBL	93.1	3.1	3.8	60	40
WBT	98.1	1.0	0.9	60	40
SBL	83.9	4.5	11.7	60	40
SBT	85.4	5.5	9.1	60	40
NBL	88.1	4.2	7.7	60	40
NBT	88.1	4.6	7.3	60	40
EBL	90.8	3.2	6.0	60	40
EBT	83.1	8.5	8.4	60	40

Table 8: Simulation parameters.

Parameters	Desired value
Duration of simulation	19.5 h
Green extension	2 s
Maximum cycle length	120 s
Amber time	3 s
All red time	2 s
Minimum duration of the WB/EB approach	21 s
Minimum duration of the SB/NB approach	13 s
Through capacity per lane	1650 pcu/h
Left-turn/right-turn capacity per lane	1550 pcu/h
Simulation runs	20

5. Results and Discussion

Figures 5–7 show the simulation results of fixed-time control (FTC), traditional fuzzy control (TFC), and hybrid genetic algorithm fuzzy control (HGAFC). Compared with FTC, the average delay time per vehicle of HGAFC is reduced by 14.59%, and the average number of stops per vehicle is reduced by 0.71%; the average delay time per vehicle of TFC is reduced by 11.08%, and the average number of stops per vehicle is reduced by 0.33%.

The following inferences can be drawn based on Figures 5–7:

(1) In the early morning, the effect of the control algorithm is the same. The reason is that when the traffic flow of each approach of the intersection is low, the vehicle queue length is short. The fuzzy control's main influence factor is each phase's red/green light duration, and the fuzzy control is in a fixed-time control state.

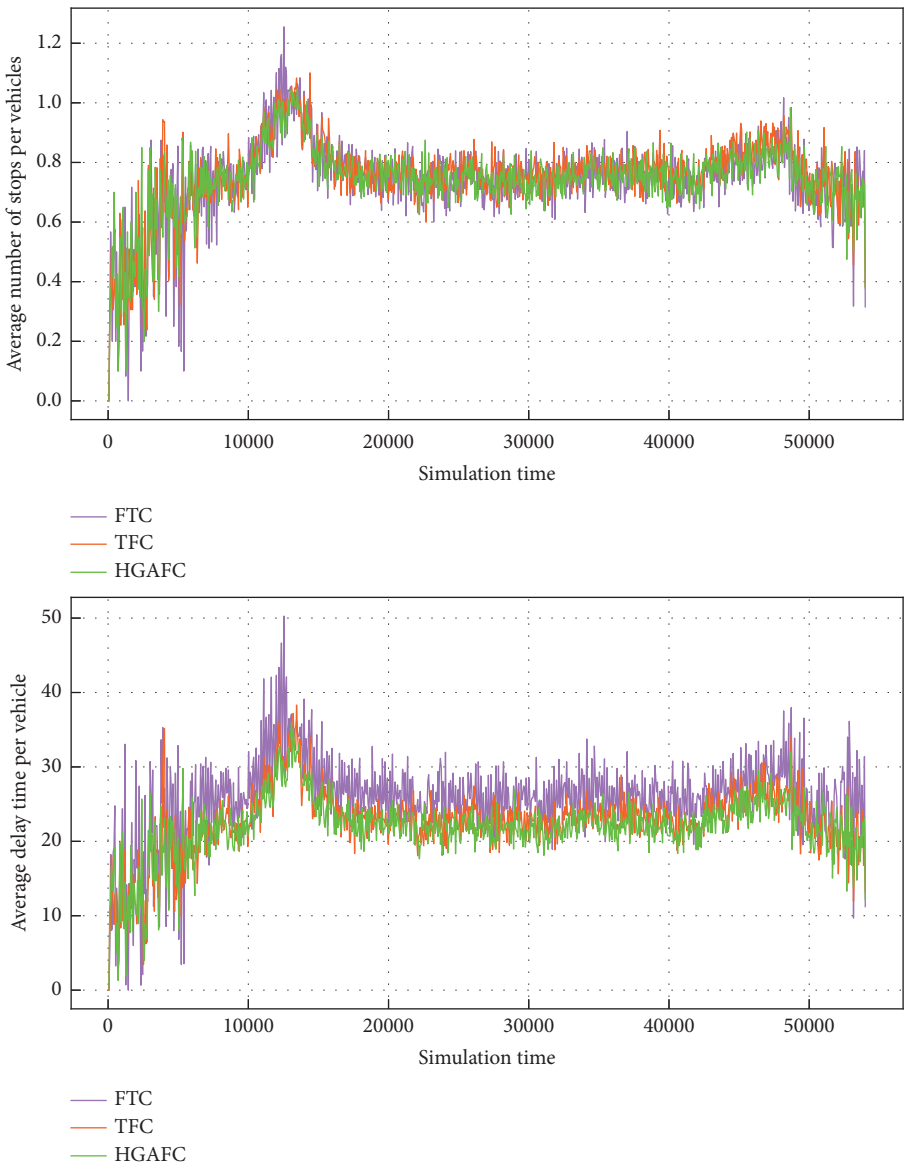


Figure 5: Simulation result.

(2) The control algorithm can effectively reduce the average vehicle delay and queue length during peak hours. The reason is that when the traffic flow at the intersection is large, the queuing vehicles cannot get through the parking line in the shortest green time, and the queuing phenomenon begins to appear. The queuing length begins to affect the decision process of the fuzzy controller, and the entrance lane with long queuing begins to obtain the green light extension time, which fully reflects the flexibility and superiority of fuzzy control over timing control.

(3) In the morning rush hour, the optimization effect of fuzzy control on the intersection is weaker than in another period, but it is still better than Fixed-time control. The reason is that when there is more traffic flow at the intersection, the queue length of each approach is longer. In this case, the queue length and phase duration are the main influencing factors of fuzzy control, but the phase duration occupies the dominant position.

(4) Figures 5 and 6 show the changes of critical parameters in the fuzzy control system under different real traffic conditions. In the simulation model, there are 2911 phases in 54000 time steps, i.e., about 727 cycles. The average cycle length of each cycle is 74 seconds, of which 152 cycles change, accounting for

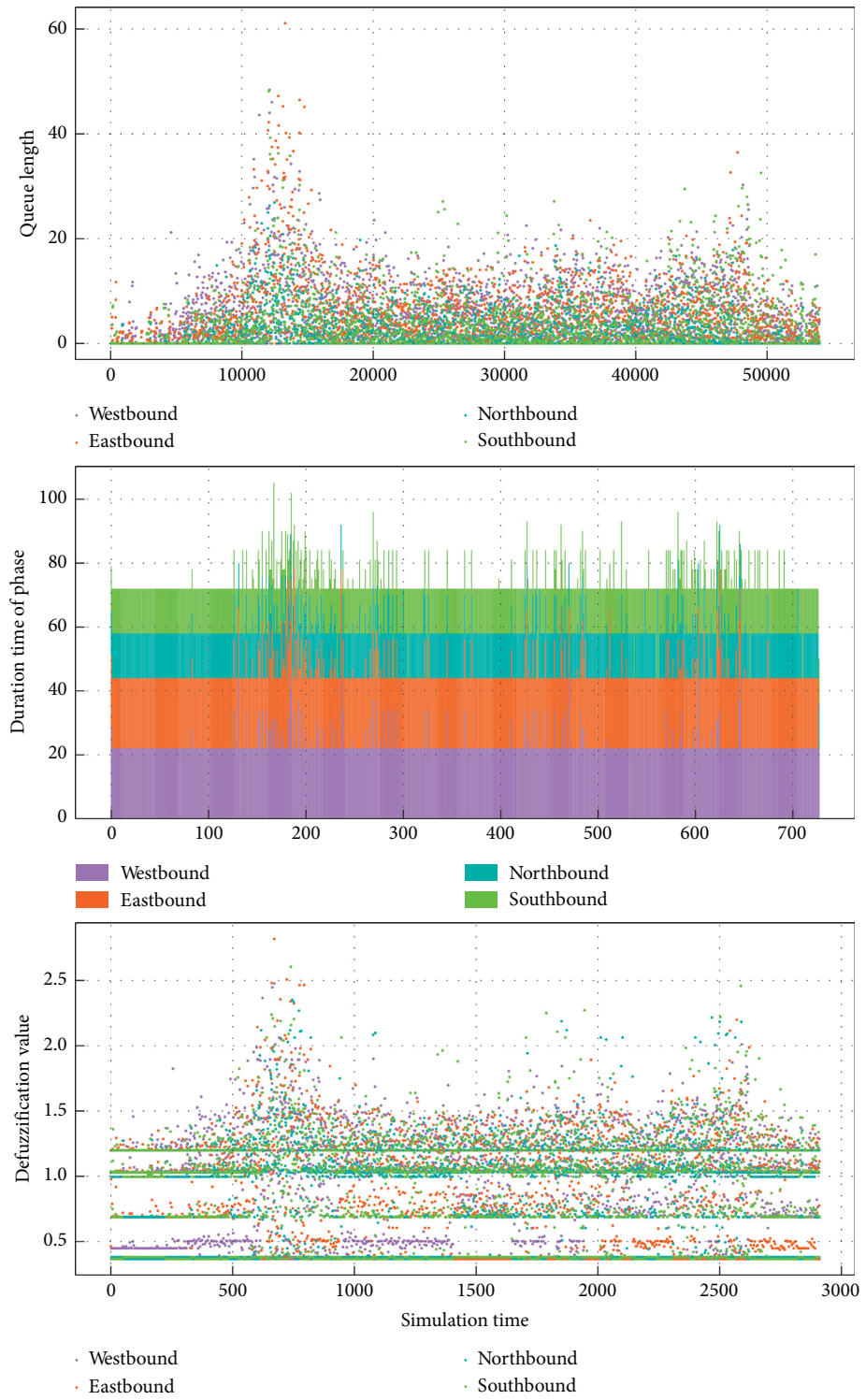


Figure 6: Parameter variation under HGA fuzzy control algorithm.

20.91% of the total cycles. It can be seen from the simulation results that the larger the traffic volume at the intersection, the more significant the signal cycle change, which shows that the signal control strategy generated by the algorithm can effectively adapt to the real-time evolution of intersection traffic flow. Among the 727 cycles, the maximum cycle length is about 100 seconds, and the minimum is 72 seconds. Each phase's maximum green light time is also within a reasonable range, and there are no extreme cases, indicating that the genetic fuzzy control algorithm has good reliability.

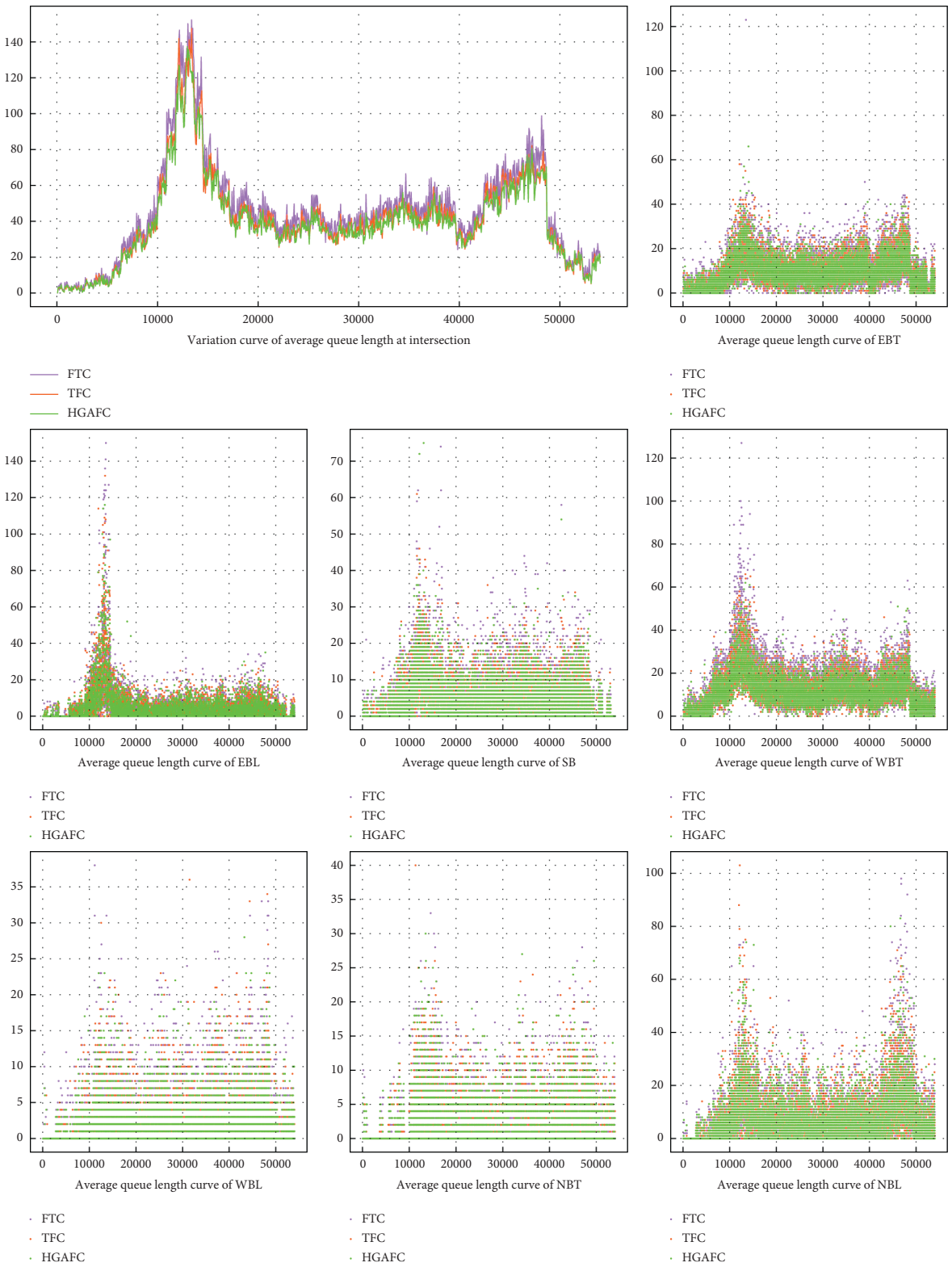


Figure 7: Average queue length curve of different approaches.

6. Conclusion

In this study, a self-adaptive two-stage fuzzy controller is proposed, which considers multiple modes of transportation: motor, non-motor, and pedestrian. In order to simulate the driving behavior characteristics and real-time traffic flow fluctuation characteristics of vehicles in the real environment, the simulation parameters of the VISSIM simulation platform are calibrated through the intersection vehicle high-precision trajectory data, and the vehicle traffic

information extracted based on the trajectory data is used as the input of the simulation platform. Then, its control efficiency and reliability are evaluated based on the microsimulation platform. Results showed that the average delay time per vehicle is reduced by 14.59%, while the average number of stops per vehicle is reduced by 0.71% compared with the traditional control method. In addition, by analyzing the effectiveness of signal control strategy under different traffic conditions, it is found that the fuzzy controller can only achieve a good control effect under medium saturation or unbalanced traffic flow but has a poor control effect under low or high saturation. This paper researched the optimization question of traffic signal timing in natural traffic conditions, but the optimization module's timeliness was not, so it is best to introduce the traffic flow prediction module in future research.

Data Availability

Data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Acknowledgments

This study was supported by project National Key R&D Program of China (No. 2019YFF0301403).

References

[1] Y. Yang, Z. Z. Yuan, D. Y. Sun, and X. L. Wen, "Analysis of the factors influencing highway crash risk in different regional types based on improved apriori algorithm," *Advances in Transportation Studies*, vol. 49, pp. 165–178, 2019.

[2] S. Feng, Z. Li, and Y. Ci, "Risk factors affecting fatal bus accident severity: their impact on different types of bus drivers," *Accident Analysis & Prevention*, vol. 86, pp. 29–39, 2016.

[3] R. Aldred, "Cycling near misses: their frequency, impact, and prevention," *Transportation Research Part A: Policy and Practice*, vol. 90, pp. 69–83, 2016.

[4] Y. Yang, Z. Z. Yuan, J. Y. Li, and Y. H. Wang, "Multi-mode public transit OD prediction and scheduling model," *Advances in Transportation Studies*, vol. 3, 2018.

[5] M. Drapalyuk, S. Dorokhin, and A. Artemov, "Estimation of efficiency of different traffic management methods in isolated area," *Transportation Research Procedia*, vol. 50, pp. 106–112, 2020.

[6] Y. Yang, Z. Yuan, J. Chen, and M. Guo, "Assessment of osculating value method based on entropy weight to transportation energy conservation and emission reduction," *Environmental Engineering & Management Journal (EEMJ)*, vol. 16, no. 10, 2017.

[7] T. Litman, *Smart Transportation Emission Reduction Strategies*, Victoria Transport Policy Institute, Victoria, Canada, 2017.

[8] W. Li, Z. Pu, and Y. Li, "How does ridesplitting reduce emissions from ridesourcing? A spatiotemporal analysis in Chengdu, China," *Transportation Research Part D: Transport and Environment*, vol. 95, Article ID 102885, 2021.

[9] W. Li, S. Chen, J. Dong, and J. Wu, "Exploring the spatial variations of transfer distances between dockless bike-sharing systems and metros," *Journal of Transport Geography*, vol. 92, Article ID 103032, 2021.

[10] Beijing Transportation Research Center, *Beijing Transport Development Annual Report*, Beijing Transportation Research Center, Beijing, China, 2021, in Chinese.

[11] W. Li, Z. Pu, and Y. Li, "Characterization of ridesplitting based on observed data: a case study of Chengdu, China," *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 330–353, 2019.

[12] K. Tang, M. Boltze, H. Nakamura, and T. Zong, *Global Practices on Road Traffic Signal Control: Fixed-Time Control at Isolated Intersections*, Elsevier, Amsterdam, Netherland, 2019.

[13] Y. Wenchen, L. Zhang, Z. He, and Y. Yan, "Adaptive two-stage fuzzy logic controllers for urban traffic signals at isolated intersections," in *Proceedings of the 2012 International Conference on Artificial Intelligence*, La Palma, Canary Islands, April 2012.

[14] N. H. Gartner, "A demand-responsive strategy for traffic signal control," in *Proceedings of the 62nd Annual Meeting of the Transportation Research Board*, pp. 75–81, Transportation Research Board, Washington, DC, USA, January 1983.

[15] F. Lin and D. J. Cooke, "Potential performance characteristics of adaptive control at individual intersections," in *Proceedings of the 65th annual meeting of the Transportation*, pp. 30–33, Transportation Research Board, Washington, DC, USA, January 1966.

[16] P. Kronborg and F. Davidson, "MOVA and LHOVRA: traffic signal control for isolated intersections," *Traffic Engineering and Control*, vol. 34, pp. 195–200, 1993.

[17] F. Lin and S. Vijayakumar, "Adaptive signal control at isolated intersections," *ASCE Transportation Journal*, vol. 114, no. 5, 1988.

[18] C. P. Pappis and E. H. Mamdani, "A fuzzy logic controller for a traffic junction," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 7, no. 10, pp. 707–717, 2002.

[19] R. Hoyer and U. Jumar, "An advanced fuzzy controller for traffic lights," *Annual Review in Automatic Programming*, vol. 19, pp. 67–72, 1994.

[20] M. B. Trabia, M. S. Kaseko, and M. Ande, "A two-stage fuzzy logic controller for traffic signals," *Transportation Research Part C: Emerging Technologies*, vol. 7, no. 6, pp. 353–367, 1999.

[21] Y. S. Murat and E. Gedizlioglu, "A fuzzy logic multi-phased signal control model for isolated junctions," *Transportation Research Part C: Emerging Technologies*, vol. 13, no. 1, pp. 19–36, 2005.

[22] S. M. Rahman and N. T. Ratrout, "Review of the fuzzy logic based approach in traffic signal control: prospects in Saudi arabia," *Journal of Transportation Systems Engineering and Information Technology*, vol. 9, no. 5, pp. 58–70, 2009.

[23] B. M. Nair and J. Cai, "A fuzzy logic controller for isolated signalized intersection with traffic abnormality considered," *IEEE Intelligent Vehicles Symposium*, pp. 1229–1233, 2007.

[24] J. J. Henry, J. L. Farges, and J. L. Gallego, "Neuro-fuzzy techniques for traffic control," *Control Engineering Practice*, vol. 6, no. 6, pp. 755–761, 1998.

[25] E. Bingham, "Reinforcement learning in neurofuzzy traffic signal control," *European Journal of Operational Research*, vol. 131, no. 2, pp. 232–241, 2001.

[26] N. Rida, M. Ouadoud, and A. Hasbi, "Ant colony optimization for real time traffic lights control on a single

intersection," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 14, no. 2, pp. 196–214, 2020.

[27] D. Xu, J. Fang, and S. Shao, "A fuzzy controller of traffic system and ITS neural network implementation," *Information and Control*, vol. 21, pp. 74–79, 1992.

[28] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: past, present, and future," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 8091–8126, 2021.

[29] J. Liu and X. Zuo, "Research on fuzzy control and optimization for traffic lights at single intersection," *Journal of System Simulation*, vol. 32, pp. 2401–2408, 2020.

[30] N. P. Shahsavari, H. Asadi, and M. P. Kheradmand, "Fuzzy multiobjective traffic light signal optimization," *Journal of Applied Mathematics*, vol. 2013, Article ID 249726, 7 pages, 2013.

[31] D. Yu, X. Tian, X. Xing, and S. Gao, "Signal timing optimization based on fuzzy compromise programming for isolated signalized intersection," *Mathematical Problems in Engineering*, vol. 2016, Article ID 1682394, 12 pages, 2016.

[32] Y. Yang, Z. Yuan, and X. Fu, "Optimization model of taxi fleet size based on GPS tracking data," *Sustainability*, vol. 11, no. 3, 2019.

[33] M. Z. Wang, *Research on the Adjustment of Signal Timing Based on High-Precision Trajectory Data: Taking the Intersection of Beier and Chongzhi Road in Shenzhen as an Example*, Beijing Jiaotong University, Beijing, China, 2021.

[34] N. Siddique and H. Adeli, *Computational Intelligence: Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing*, John Wiley & Sons, Hoboken, NJ, USA, 2013.

[35] C. Urrea, J. Kern, and J. Alvarado, "Design and evaluation of a new fuzzy control algorithm applied to a manipulator robot," *Applied Sciences*, vol. 10, no. 21, Article ID 7482, 2020.

[36] Y. Nie, X. Wu, J. F. Dillenburg, and P. C. Nelson, "Reliable route guidance: a case study from Chicago," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 2, pp. 403–419, 2012.

[37] B. Park and J. D. Schneeberger, "Microscopic simulation model calibration and validation: case study of vissim simulation model for a coordinated actuated signal system," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1856, no. 1, pp. 185–192, 2003.

[38] J. Hourdakis, P. G. Michalopoulos, and J. Kottommannil, "Practical procedure for calibrating microscopic traffic simulation models," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1852, no. 1, pp. 130–139, 2003.

[39] L. Chu, H. X. Liu, J. S. Oh, and W. Recker, "A calibration procedure for microscopic traffic simulation," in *Proceedings of the 2003 IEEE International Conference on Intelligent Transportation Systems*, vol. 2, Shanghai, China, October 2003.

[40] M. Hosamo, "A study of the source traffic generator using Poisson distribution for ABR service," *Modelling and Simulation in Engineering*, vol. 2012, Article ID 408395, 6 pages, 2012.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 3001392, 11 pages
https://doi.org/10.1155/2022/3001392

Research Article

A Comparative Study of Passenger Multitasking Activities on Commuting and Leisure Electrified Intercity Railways

Xing Yao, Chunhui Jing, Yaoxuan Huang, and Jinyi Zhi

Department of Industrial Design, Southwest Jiaotong University, Chengdu 611756, China

Correspondence should be addressed to Xing Yao; yancy216@my.swjtu.edu.cn and Jinyi Zhi; zhijinyi@swjtu.edu.cn

Received 15 November 2021; Accepted 24 December 2021; Published 21 January 2022

Academic Editor: Wenxiang Li

The electrification of intercity railways plays a significant role in energy conservation and emission reduction. Research on passenger travel activities to optimize vehicle services can help to understand how best to improve passenger attraction and the use of intercity railways. In this study, we conducted observational research on two electrified intercity railways that were segregated by the attributes, leisure, and commute. The purpose was to determine the influence of line attribute, passenger gender, age, and seat availability on the types of activities performed onboard, with specific attention placed on the use of information and communication technology (ICT). Using structured observations, the travel multitasking activity data of 467 passengers were collected on two intercity railways in real-life situations. Using the chi-square test and binary logistic regression analysis, it was found that line attribute, gender, age, and seat availability have an impact on passenger activities. Differences in factors affecting passenger activities were also found according to the nature of their travel, whether for commute or leisure. Our results suggest that passengers on the leisure line prefer to engage in some social activities. For example, the probability of conversation among passengers on the leisure line was 3.47 times that of the commuting line, and the middle-aged and elderly travelers on this line were more likely to be in a daze and look around. The probability of taking a break for passengers on the commuting line was 3.625 times that of the leisure line, and passengers who were not seated on this line were found to be more likely to be idle. In addition, male travelers and young travelers preferred to engage in ICT immersive activities, such as using mobile phones, while women, middle-aged, and elderly travelers were more likely to engage in non-ICT immersive activities. Seated passengers were more likely to engage in simultaneous multitasking activities, rest, and conversations than passengers without seats.

1. Introduction

Local policies [1, 2], actual cases [3], or related technology research [1] have shown that electrification has become an important issue to which academics, local administrative agencies, and the transportation industry officials have all begun paying attention. Electrification transformation in the field of mobile transportation is already one of the productive choices for urban sustainable development. The increase in municipal deployment of environmentally friendly vehicles, such as electric buses, trams, and trolleybuses, can help a city's energy model to shift toward a low-pollution model [4–6]. The railway is one of the four major areas of the transportation industry. Taking China as an example, the rapid development of electrified railways and high-speed trains has contributed to a fluctuating downward trend in carbon dioxide emissions from the railway transportation industry [7]. Furthermore, with the growth of national leisure and entertainment demand, tourism has become the fastest-growing sector in the Chinese economy, and people's demand for railways is rapidly increasing.

It can be seen that the electrification of railways provides a greener and more environmentally friendly means of transportation for urban residents. However, in addition to the existence of sustainable transportation tools, people's willingness to use these tools is another necessary factor for the sustainable development of urban transportation [8]. The European Commission, for example, found that it was necessary to not only upgrade technology to reduce the impact of transportation on the environment but also to increase the attractiveness of trains as a mode of

transportation. Moreover, with the increasing popularity of information and communication technology (ICT), passenger experience could be enriched with more diverse forms of activities during travel. Collecting evidence about how railway passengers spend their travel time can provide advice to decision-makers, help operators adjust services to better meet passenger needs, and increase the attractiveness of sustainable transportation.

Past experience has shown that the mode of travel was the main determinant of the type of multitasking activity [9, 10]. Different travel modes lead to differences in the cognitive and physical energy consumption of passengers, which lead to differences in multitasking activities [11]. In the past, comparative studies on multitasking activities between different modes of travel were mostly seen between private and public transportation. Lyons et al. [12] found that passengers had the highest level of productivity on trains, followed by cars and buses. Russell et al. [13] found that there were more people looking ahead/out of the window on the bus than on the train, and more people were reading, using a computer, sleeping, or writing on the train. Singleton et al. [14] found that bus and car passengers engaged in more types of activities, while vehicle (cars and bicycles) drivers engaged in fewer types of activities, most of which were passive activities. Keseru et al. [15] compared the multitasking activities of passengers on buses, trams, and subways and found no significant differences. Utsunomiya [16] surveyed the residents of the two railway areas and found that, compared with the traditional nonelectrified railway areas, the residents of the electrified railway areas were more inclined to shop frequently and participate in local festivals. It can be seen that the existing travel-based multitasking information in the field of public transportation mostly comes from trains, and there are few studies on short- and medium-distance vehicles, such as subways and intercity trains [10], and more detailed comparisons (such as line attribute) are lacking.

In addition, gender and age are considered to be important factors affecting passengers' multitasking activities. Berliner et al. [17] found that male passengers were more likely to use ICT devices (listening to music, watching videos), older travelers were more inclined to activities that did not require ICT devices, such as reading, and younger travelers were more likely to use computers, mobile phones, and so on. However, Lyon et al. [18] found that young travelers were more likely to be bored than middle-aged and elderly travelers. Keseru et al. [19] and Frei et al. [20] found that women were more likely to participate in social activities, such as conversation and reading. Many researchers called for further research on the use of travel time in different countries and different cultural backgrounds [21] because compared with developed countries, there is a lack of research on multitasking in developing countries [10], a deficit in similar research that is necessary to undertake in developing countries.

In the context of sustainable development and electrification revolution in urban transport, this study took Chengdu, a typical city in China (the largest developing country), as an example and selected two electrified intercity railways with an attribute for leisure and commuting. The study obtained the universal evidence and influencing factors of passenger multitasking activities based on the real-world observed data and provided service suggestions related to the passengers' behaviors, railway operation, and policies for relevant government decision-makers to help transition toward urban electrification. This research attempts to answer the following questions:

(1) Does the attribute of the electrified intercity railways (leisure line, commuting line), passenger gender, age, and availability of seats affect passengers' multitasking activities?

(2) Are there any differences in the factors affecting passenger multitasking activities in electrified intercity express rail lines between the attributes leisure and commuting?

The next section of the article introduces the sources and collection methods of passenger multitasking activity data. Section 3 describes the results of the data analysis. Section 4 discusses the results in combination with the previous research and provides some service suggestions. Section 5 summarizes the research results and presents the limitations of this article.

2. Methods

2.1. Research Method. In previous studies on passenger multitasking activities, qualitative methods (such as qualitative interviews and focus groups) and quantitative methods (such as questionnaires and observations) have been used. More recently, in this field, structured observation and questionnaire surveys have become the main research methods [10].

This study mainly uses structured observation and instantaneous sampling to collect and record data. Compared with questionnaire surveys, observation has been considered a nonobtrusive method [13] and can be used to collect a large amount of data in a short time to provide researchers with a wealth of information. Under these research conditions, subject behavior tends to be more natural, and researchers can more readily avoid the memory effect inherent to the use of questionnaires [22]. One key issue arising when using the observation method is setting the time of the passenger observation window [23] or how long the passenger's activity frequency and duration are recorded. A longer observation window means a smaller sample size and more observation costs. Gamberini et al. [24] found that 84.2% of travelers only performed one activity during the observation period, which is consistent with the findings of Van Der Waerden et al. [25]. Therefore, the use of instantaneous sampling—that is, making instantaneous recordings about the activities that passengers are engaged in—is a way to balance cost and efficiency.

2.2. Data Source and Collection. Chengdu is a central city in southwestern China and an important junction and transportation corridor between Southeast Asia and South

Asia [26]. The geographic locations of the two intercity lines selected are shown in Figure 1, as detailed in the following:

(1) Intercity line from Chengdu to Qingcheng Mountain: the whole journey takes 42 min and passes through four stops, including famous local attractions. It is presumed to be a leisure line, and the passenger flow is mainly leisure travel, hereinafter referred to as the "leisure line."

(2) Intercity line from Chengdu to Jiangyou: the whole journey takes 1 h and 13 min and passes through five stops, which are urban cities around Chengdu (excluding famous tourist cities). It is presumed to be a commuting line. The passenger flow is mainly commuting and business, hereinafter referred to as the "commuting line."

Because of the strong volatility of railway passenger flow, there is a substantial difference in passenger flow between the peak and nonpeak periods. Previous studies have shown that it is almost impossible to conduct an organized observation during peak hours [13, 27]. Hence, most peak hours were not included. However, the method of instantaneous sampling greatly reduces the difficulty and cost of data collection, and since one of our purposes was to study the difference between the two lines, collecting data from peak hours offered the opportunity to make the results more representative. According to previous studies [28], the passenger flow of leisure travel showed its peak during holidays, and the traffic was the largest. The peak of commuter passenger flow occurred between 7 : 00 and 9 : 00 and between 17 : 00 and 19 : 00 every day. However, commuters taking intercity trains are mostly cross-city, weekly, or monthly commuters. There are few daily commuters [28]. Thus, observations for the commuting line were collected from 16 : 00 to 19 : 00 on Friday, December 20, 2019. For the leisure line, the observation was conducted from 9 : 00 to 12 : 00 on Saturday, December 15, 2019. Both investigations were conducted in fair weather without any abnormal conditions.

We began with a preinvestigation of the travel purpose of passengers on the two lines. We randomly distributed questionnaires to passengers on the two lines (103 on the leisure line, 155 on the commuting line). The results of the questionnaire showed the tourist flow of the leisure line to be 67.92%, of which a vast majority of passengers were foreign tourists. The commuter flow of the commuting line was 65.82%, of which most travelers were migrant workers, and the travelers on business accounted for only 7.27%. Both conformed to the original hypothesis of line attribute.

Data collection process: investigator A held a sports camera (GoPro Hero8 Black) in front of his chest and walked from the first carriage to the rear carriage at a normal speed to avoid deliberate data collection for passengers, activities which not only ensured a strong adherence to passenger privacy but also reduced any "observation effect" [29] to obtain the most real data possible.

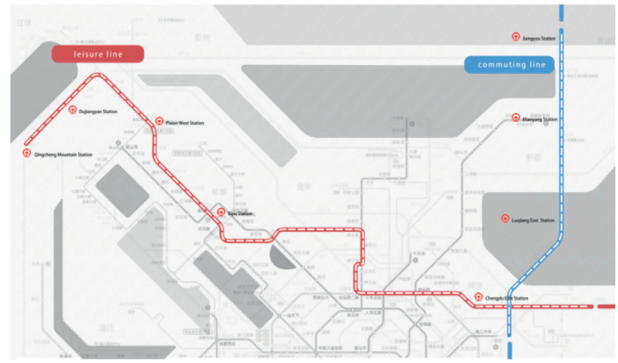


Figure 1: Two electrified intercity lines for data collection.

In recording the data, the researchers conducted preobservations on the collected data and discussed and formulated the observation structure (Table 1), including the specific definitions of variables, such as passenger gender, age, availability of seats, passenger multitasking activities, and standards of invalid filtering information. Two observers were selected from the preobservation to conduct formal observations at the same time to discuss some ambiguous information (such as age) and discard controversial data in a timely manner. On the two lines, a total of 605 passengers were observed (leisure line: 203 times, recording time: 5″56′; commuting line: 405 times, and recording time: 4″18′). Among them, activities that did not meet the specified definitions were not included. As the purpose of passengers walking in the carriage cannot be understood by observation, these observations were discarded. To avoid the impact of seat level differences, the data of the first-class carriage were not included. In addition, a carriage was in a parked state during the observation, and as the observers were strongly interfered with by the passengers getting on and off, these observations were not included in the final dataset. Finally, 507 valid observations were obtained (leisure line: 175 times, commuting line: 332 times). According to previous studies [10], the number of observations is in an intermediate position, which is considered appropriate. The data were strictly kept confidential during the whole process to protect the privacy of passengers.

3. Results

Descriptive analysis, single-factor analysis, and multifactor analysis were performed on the observed data. In single-factor analysis, the chi-square test was performed on multitasking activities and independent variables to initially test the correlation between independent variables and dependent variables, and variables with small correlations would be excluded to make the results more reliable. Then, the binary logistic regression analysis method was used, and odds ratio (OR), which is often used in similar studies [30, 31], was selected to test the relationship between the independent variable and activities.

Table 1: A grid for structured observation of passengers on intercity lines.

Item		Explanatory variable
Line attribute		(Leisure line/commuting line)
Gender	Male	Male passengers identified directly through observation
	Female	Female passengers identified directly through observation
Age	Children	Child passengers aged 13 or younger perceived through observation
	Youth	Young passengers aged 14–39 perceived through observation
	The middle-aged and elderly	Middle-aged and elderly passengers aged 40 and above perceived through observation
Availability of seats	Yes	Passengers seated in the seat area of the train
	No	Nonseated passengers standing in the train
Multitasking activity	Effective activity	Synchronous multitasking: doing two or more activities at the same time
		Being in a daze/looking around
		Using the cell phone
		Wearing headphones
		Talking with other passengers
		Resting: resting and sleeping
		Talking on the phone
		Using computer: using computer or tablet
		Reading: reading paper books
		Writing: writing using pen and paper
		Eating: eating or drinking
		Companion care: caring for a child, elderly people, or a companion
	Invalid activity	Incomplete records or unrecognizable activities

3.1. Descriptive Statistics. To determine the factors influencing passengers' multitasking activities, the independent variables created were as follows:

(1) Line attribute: use 1 and 2 to represent leisure line and commuting line, respectively

(2) Gender: use 1 and 2 to represent male and female passengers, respectively

(3) Age: use 1 and 2 to denote younger and older passengers, respectively

(4) Availability of seats: use 1 and 2 to indicate seated and nonseated, respectively

Table 2 shows the descriptive statistics. There were slightly more women than men in the sample, which was consistent with the slight imbalance in gender distribution commonly observed among public transport passengers [32]. Through preobservation, children accounted for the smallest proportion, and most of them did not engage in multitasking activities. Hence, the data on children's multitasking were not included. There were fewer middle-aged and elderly travelers. Hence, the two groups were analyzed together. Most of the observed passengers (66.6%) belonged to the youth group, and the middle-aged and elderly group accounted for a relatively smaller proportion (34.4%). There were more seated passengers (77.8%) than nonseated passengers (22.2%). The number of activities (581) was more than the number of passengers observed (507), indicating that some passengers performed two or more activities at the same time (24.6%).

Figure 2 shows the distribution of passenger activities. Using mobile phones, wearing headphones, being in a daze or looking around, and taking a break were the most popular activities. Talking with other passengers and eating were also distributed but in a small proportion. In contrast, activities, such as companion care, talking on the phone, writing, reading, and using a computer, were rarely distributed.

As passengers had more types of multitasking activities and some activities were less frequent, to better study the links between activities, the activities were further aggregated based on the use of ICT equipment and previous studies [15, 33] (Table 3). The results were also entered into the regression analysis. Figure 3 shows the overall distribution: ICT-based activities still occupied the main body, and active ICT activities were the most distributed. Figure 4 shows the comparison of the two lines.

3.2. Single Factor Analysis of Passenger Multitasking Activities. We used a four-grid table chi-square test to exclude some variables with a lower degree of fit to ensure a more accurate subsequent multifactor analysis. The standard of significance level in this step was appropriately relaxed: the variable with a p-value of 0.15 or less was included in the multifactor analysis. The results of the single factor analysis are shown in Table 4. Line attribute, passenger gender, passenger age, and seat availability had a significant impact on passengers' multitasking activities. Hence, they were all retained, while variables, such as reading, writing, talking, and companion care, showed a small frequency, or the result was not significant. Hence, they were not included in the multifactor analysis.

3.3. Multifactor Analysis of Passenger Multitasking Activities. We used binary logistic regression to conduct a subsequent multifactor analysis, the results of which are shown in Table 5.

On the leisure line, it was easier for passengers to talk with other passengers (OR = 0.288), while on the commuting

Table 2: Descriptive statistics of independent variables.

Variable	Value	Number	Proportion (%)
Line attribute	Leisure line: 1	132	28.4
	Commuting line: 2	332	71.6
Gender	Male: 1	230	49.6
	Female: 2	234	50.4
Age	Youth: 1	309	66.6
	The middle-aged and elderly: 2	155	33.4
Availability of seats	Yes: 1	361	77.8
	No: 2	103	22.2

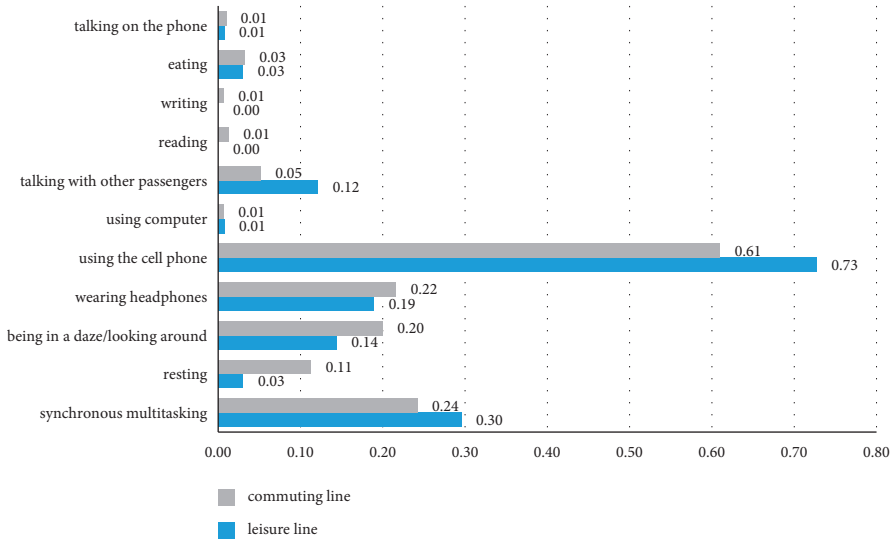


Figure 2: The observed distribution of passenger multitasking activities.

Table 3: Classification of aggregated multitasking activities.

Aggregated multitasking activities		Activities
Non-ICT activities	Paperwork	Reading
		Writing
	Social activities and people care	Talking with other passengers
		Talking on the phone
		Eating
		Companion care
	Passive non-ICT activities	Resting
		Being in a daze/looking around
ICT activities	Active ICT-based activities	Using the cell phone
		Using the computer
	Passive ICT-based activities	Wearing headphones

line, it was easier for passengers to sleep and rest (OR = 3.625).

Male passengers were more likely to engage in simultaneous multitasking (OR = 0.630) and wear headphones (OR = 0.422), while female passengers were more likely to engage in resting (OR = 1.820) and eating (OR = 5.611).

Young travelers were more likely to perform simultaneous multitasking (OR = 0.370), wear headsets (OR = 0.208), and use mobile phones (OR = 0.437) than middle-aged and elderly travelers. The middle-aged and elderly travelers were more likely to be in a daze and look around (OR = 2.705).

Seated passengers were more likely to engage in simultaneous multitasking (OR = 0.462), rest (OR = 0.101), and conversations with other passengers (OR = 0.087) than nonseated passengers. Nonseated passengers were more likely to be in a daze and look around (OR = 1.736).

Analyzing the multitasking activities of the two lines separately, the results showed that on the leisure line, men were more likely to engage in simultaneous multitasking

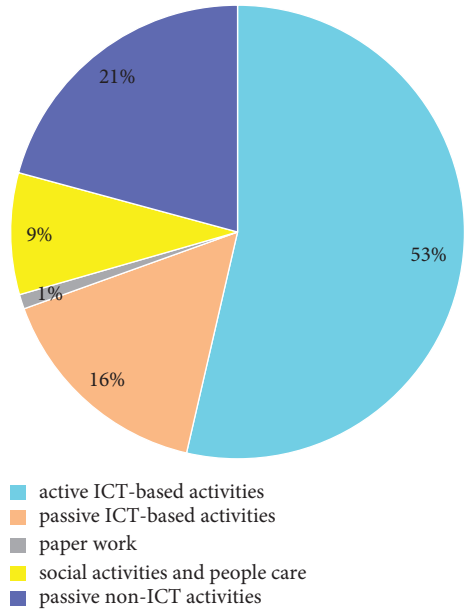


Figure 3: Distribution of aggregated multitasking activities.

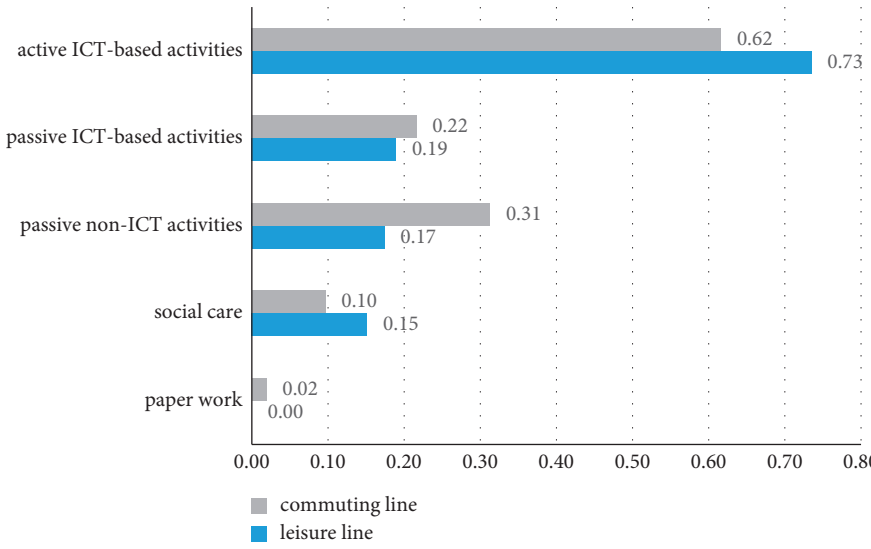


Figure 4: Distribution of aggregated multitasking activities on two lines.

(OR = 0.38) while wearing headphones (OR = 0.26), and women were more likely to take breaks (OR = 1.3). Young travelers were more likely to perform multiple activities at the same time (OR = 0.14), wear earphones (OR = 0.13), and use their mobile phones (OR = 0.18), while middle-aged and elderly passengers were more likely to be in a daze/looking around (OR = 12.12). Nonseated passengers were more likely to wear earphones (OR = 2.27).

On the commuting line, men were more likely to wear headphones (OR = 0.51), while women were more likely to eat (OR = 10.58). Young passengers were more likely to carry out multiasking activities at the same time (OR = 0.45), wear headphones (OR = 0.21), and use mobile phones (OR = 0.55), and middle-aged and elderly passengers were more likely to be in a daze, looking around (OR = 1.74) or be and in conversations with others (OR = 5.1). Seated passengers were more likely to carry out multitasking activities at the same time (OR = 0.37) and rest (OR = 0.1), while nonseated passengers were more likely to be in a daze or were more likely to look around (OR = 1.9).

3.4. Multifactor Analysis of Aggregation Multitasking Activities. Next, we performed a binary logistic regression analysis on the aggregate multitasking activities of the passengers, the results from which are shown in Table 6.

Passengers on the leisure line were more likely to engage in social care activities (OR = 0.5), and passengers on the commuting line were more likely to engage in passive non-ICT activities (OR = 1.77). Men were more likely to engage in

Table 4: Chi-square test results of multitasking activities.

Type	χ^2			
	Line attribute	Gender	Age	Availability of seats
Synchronous multitasking	2.465**	—	13.520***	3.594*
Resting	6.923***	3.184*	—	9.504***
Being in a daze/looking around	—	—	15.010***	—
Wearing headphones	—	9.733***	21.386***	—
Using the cell phone	4.007**	—	18.762***	2.948*
Using the computer	—	3.305*	—	—
Talking with other passengers	7.843***	—	—	7.240***
Reading	—	—	—	—
Writing	—	—	—	—
Eating	—	7.189***	—	2.900*
Talking on the phone	—	—	—	—
Companion care	—	—	—	—

Note. Only significant models at 10% level tested using chi-square tests are shown. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$

Table 5: Logistic regression results of multitasking activities.

Type	Explanatory variable			
Total	Line attribute (leisure line: 1, commuting line: 2)	Gender (male: 1, female: 2)	Age (youth: 1, the middle-aged and elderly: 2)	Availability of seats (yes: 1, no: 2)
Synchronous multitasking	—	0.630**	0.370***	0.462**
Resting	3.625**	1.820*	—	0.101**
Being in a daze/looking around	—	—	2.705***	1.736*
Wearing headphones	—	0.422***	0.208***	—
Using the mobile phone	—	—	0.437***	—
Computer use	—	—	—	—
Talking with other passengers	0.288***	—	—	0.087**
Eating	—	5.611**	—	—
Leisure line				
Synchronous multitasking	—	0.38**	0.14***	—
Resting	—	1.3*	—	—
Being in a daze/looking around	—	—	12.12***	—
Wearing headphones	—	0.26***	0.13*	2.27*
Using the mobile phone	—	—	0.18**	—
Computer use	—	—	—	—
Talking with other passengers	—	—	—	—
Eating	—	—	—	—
Commuting line				
Synchronous multitasking	—		0.45***	0.37**
Resting	—	—	—	0.1**
Being in a daze/looking around	—	—	1.74*	1.9**
Wearing headphones	—	0.51**	0.21***	—
Using the mobile phone	—	—	0.55***	—
Computer use	—	—	—	—
Talking with other passengers	—	—	5.10***	—

Note. Only significant models at 10% level tested using logistic regression are shown. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$.

active ICT activities (OR = 0.66) and passive ICT activities (OR = 0.43), while women were more likely to engage in passive non-ICT activities (OR = 1.55). Young passengers were more likely to engage in active ICT activities (OR = 0.4) and passive ICT activities (OR = 0.22), and middle-aged and elderly travelers were more likely to engage in passive non-ICT activities (OR = 2.25). Passengers with seats were more likely to engage in social care activities (OR = 0.25).

Analyzing the two lines separately, the results showed that on the leisure line, male passengers (OR = 0.26), young passengers (OR = 0.13), and nonseated passengers (OR = 2.27) were more likely to engage in passive ICT activities, and young passengers were more likely to engage in active ICT activities (OR = 0.17), while middle-aged and elderly passengers were more likely to engage in passive non-ICT activities (OR = 9.6), and seated passengers were more likely to engage in social care activities (OR = 0.11).

On the commuting line, male passengers were more likely to engage in active ICT activities (OR = 0.65) and passive ICT activities (OR = 0.51), while female passengers

Table 6: Summary of logistic regression results for aggregated multitasking activity types.

Aggregated multitasking activity types	Explanatory variable			
Total	Line attribute (leisure line: 1, commuting line: 2)	Gender (male: 1, female: 2)	Age (youth: 1, the middle-aged and elderly: 2)	Availability of seats (yes: 1, no: 2)
Non-ICT activities	—	1.52**	2.33***	—
Paperwork	—	—	—	—
Social activities and people care	0.50**	—	—	0.25***
Passive non-ICT activities	1.77**	1.55**	2.25***	—
ICT activities	—	0.62***	0.35***	—
Active ICT-based activities	—	0.66**	0.40***	—
Passive ICT-based activities	—	0.43***	0.22***	—
Leisure line				
Non-ICT activities	—	—	3.20***	0.31**
Paper work	—	—	—	
Social activities and people care	—	—	—	0.11**
Passive non-ICT activities	—	—	9.6***	—
ICT activities	—	0.40**	0.14***	—
Active ICT-based activities	—		0.17***	—
Passive ICT-based activities	—	0.26***	0.13**	2.27*
Commuting line				
Non-ICT activities	—	1.55**	2.02***	—
Paperwork	—	—	—	—
Social activities and people care	—	—	3.02***	—
Passive non-ICT activities	—	1.54*	1.58*	—
ICT activities	—	0.66*	0.47***	—
Active ICT-based activities	—	0.65*	0.55***	—
Passive ICT-based activities	—	0.51**	0.21***	0.31***

Note. Only significant models at 10% level tested using logistic regression are shown. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$.

were more likely to engage in passive non-ICT activities (OR = 1.54). Young passengers were more likely to engage in active ICT activities (OR = 0.55) and passive ICT activities (OR = 0.21), while middle-aged and elderly passengers were more likely to engage in social care activities (OR = 3.02) and passive non-ICT activities (OR = 1.58). Seated passengers appeared to find it easier to engage in social care activities (OR = 0.31).

4. Discussion

4.1. Factors Affecting Passenger Multitasking Activities on Electrified Intercity Railways

4.1.1. The Impact of Line Attribute on Passenger Multitasking Activities. The results of binary logistic regression analysis show that passengers on the leisure line were more likely to engage in social care activities (talking and eating), and passengers on the commuting line were more likely to engage in passive non-ICT activities (rest). The reason may be that passengers on the leisure line take leisure as the main purpose, and they tend to eat and chat with their companions. Passengers on the commuting line mainly commute to work, and passive activities, such as rest and daze, are beneficial to their preservation and recovery.

In recent years, only a few studies have compared the two different lines. However, similar studies focused on the analysis of the travel purpose of passengers. For example, Lyons et al. [18] found that commuters were more likely to engage in relaxation, leisure, and social activities, such as reading, listening to music, or making a phone call. Business travelers preferred to work, call, send text messages, and check e-mails, while leisure travelers preferred to relax and talk with others. We have found similar results in this article.

4.1.2. The Impact of Gender, Age, and Availability of Seats on Passenger Multitasking Activities. As our results show, male travelers were more likely to perform simultaneous multitasking, active ICT activities, and passive ICT activities, the results of which are consistent with previous studies [17, 24, 34]. Female passengers were more likely to engage in passive non-ICT activities, such as resting and eating. Hence, power bank rental and sales services may be provided on the train, along with low-fat and low-sugar snacks and beverages, which are more likely to be preferred by women.

Age also had a strong influence on the types of multitasking activities. Compared with middle-aged and elderly travelers, young travelers were more likely to engage in active ICT activities (using mobile phones) and passive ICT activities (wearing headphones), and middle-aged and elderly travelers were more likely to be in a daze, look around, or engage in other passive non-ICT activities during their journeys. Interestingly, there was no difference in the resting activities between young and old travelers, indicating that compared with younger travelers, middle-aged and elderly travelers were not more likely to feel tired and need to rest during the trip, which differs from our initial assumptions. It indicates that middle-aged and elderly travelers have the energy to carry out activities during the journey, however, compared to ICT activities that require more energy, they may be more comfortable engaging in passive activities. Therefore, more passive non-ICT services could be provided for elderly passengers, such as magazines or interesting carriage interior design. For example, the interior design of trains on the leisure line can be derived from the local history and culture. Although young travelers tended to engage in ICT activities during their travels, Lyon et al. [18] found that they were easily bored, indicating that they may not really want to engage in ICT activities during the trip. ICT activities may just represent a limited choice under restricted conditions and can be further studied through questionnaires in the future.

We also compared the difference between seated and nonseated passengers. Many passengers buy station tickets for short journeys when seats are sold out. The results of our study show that the seated passengers were more likely to engage in social care activities, such as simultaneous multitasking, resting, and talking, than passengers without seats. There is also a certain logic to this phenomenon: generally speaking, when sitting on a train, it is easier to carry out many of these activities than while standing, which is consistent with the content reported by Zhang et al. [35].

4.2. Differences in the Factors Affecting Multitasking Activities in the Two Lines

4.2.1. Multitasking Activities on the Leisure Line. On the leisure line, male passengers were more likely to engage in passive ICT activities (wear headphones), while women were more likely to take breaks and rest, indicating that male passengers may prefer immersive ICT experiences or pursue a personal immersive activity, such as playing a game or watching videos while wearing headphones. One reason may be that travel, as a process of moving from the departure place to the destination, provides an illusion of transition between different realities. In this transition time, male travelers may be more inclined to immerse themselves in their personal world and enjoy the leisurely time for themselves. This finding is similar to the concept of "me time" proposed by Jain et al. [36]. Women may also need this immersive experience, however, compared to men, female travelers may be less likely to use ICT devices. Hence, the results suggested that they were more inclined to rest.

As our results suggest, young travelers preferred ICT activities (using mobile phones and wearing headphones) on the leisure line, while middle-aged and elderly travelers preferred passive non-ICT activities (dazing or looking around), and the difference was extremely significant ($P < 0.01$). It shows that middle-aged and elderly passengers have a significant preference for dazing and looking around on the leisure line, and no preference for resting could be determined.

Compared with nonseated passengers, seated passengers were found to be more likely to engage in social care activities on the leisure line, such as talking or eating, while nonseated passengers preferred to wear headphones. Passengers with seats indicated that they were more willing to engage in social activities, while passengers without seats preferred passive personal immersive activities, such as listening to music.

4.2.2. Multitasking Activities on the Commuting Line. On the commuting line, male passengers were more likely to engage in ICT activities, and female passengers were more likely to engage in passive non-ICT activities (eating). Female passengers may be more likely to prepare or buy food on commuter trains. Hence, providing food sales services that are preferred by women is a service addition worth considering. Middle-aged and elderly travelers also preferred to talk to other travelers, daze, or look around, showing that they may prefer non-ICT activities that consume less energy. Compared to nonseated travelers, seated travelers found it easier to carry out social care activities, which is also consistent with the results of the leisure line.

4.3. Service Suggestions on Different Lines. Based on the differences between commuting and leisure lines and their respective characteristics, service suggestions for relevant government decision-makers are provided below to help improve the attraction and use of intercity railways and to further transition toward urban electrification.

4.3.1. On the Leisure Line. Considering that passengers on the leisure line prefer social and leisure activities, it may be possible to provide passengers with automatic vending services and optimize the seat layout to facilitate passengers' social activities. Men and young travelers on this line were more likely to perform multiple activities at the same time and ICT activities (wear earphones, use mobile phones), indicating that they need a power supply or mobile power, which are currently not provided on this line according to our observations. Therefore, power bank rental services targeted to male and young travelers could be considered.

In addition, middle-aged and elderly passengers on this line were found to prefer to be in a daze, and therefore, more passive non-ICT services for middle-aged and elderly groups could be considered, such as providing in-car reading materials or interior design related to local culture.

As the results suggest, seated passengers were more likely to engage in social care activities on the leisure line.

Therefore, the layout of train seats could be optimized for the social needs of passengers, and for passengers without seats, more personal immersive regional services could be provided for their participation, such as the provision of online descriptions of local attractions.

4.3.2. On the Commuting Line. Generally speaking, compared with the leisure line, the commuter line has a more obvious bias toward rest. Hence, it is possible to consider providing rest-related services on this line. For example, Lyons et al. [21] suggested that listening to music could make the journey more enjoyable for passengers. It also helps commuters overcome the pressure of being close to others by creating a psychological distance. Music and lights could be provided on the commuting line, for example, to create a quiet and comfortable environment for passengers. Earplugs, eye masks, and other goods that help people rest could also be provided in the carriage. In addition, to prevent passengers from sleeping past their station, getting off reminders can also be considered.

The same results were found on the leisure line for men and young travelers, who were more likely to perform multiple activities at the same time along with ICT activities (wear earphones, use mobile phones). Therefore, power bank rental services could be considered. Low-fat and low-sugar snacks and beverages that are more likely preferred by women can also be considered because female passengers were more inclined to eat in the commuting line, as observed from the results.

5. Conclusion

This study introduces passenger multitasking activities and influencing factors on electrified intercity railways. It focused on a comparative analysis of the two lines and found that there were differences in the types of passenger activities between the leisure and the commuting line. On the leisure line, passengers preferred to engage in social activities, such as talking with others, and middle-aged and elderly passengers on this route were often found to be doing nothing in the carriage (in a daze or looking around). Passengers on the commuting line, by contrast, were found to prefer rest, and nonseated passengers on this line tended to be idle. We also found that demographic characteristics (age, gender) and availability of seats also affected passenger activities. Male travelers and young travelers were more likely to engage in ICT-related immersive activities (using mobile phones and wearing headphones), while women, middle-aged, and elderly travelers tended to engage in non-ICT immersive activities. Women preferred to rest, while middle-aged and elderly passengers preferred to stay in a daze and look around. Seated passengers were found to be more likely to engage in activities, such as simultaneous multitasking, resting, and talking than the nonseated passengers. The conclusion can be used for reference by relevant departments to improve services and attract more passengers to take electrified intercity lines. However, there are some limitations to this study. Firstly, it is difficult to cover all the time periods of the day. The data during peak periods alone were collected. Secondly, since the main method used was a structural observation, the number of useful variables collected was relatively small. In the future, more information on travel attributes and ride experience of passengers can be collected through questionnaires, and the number of variables can be increased to make the model more accurate.

Data Availability

No additional data were used to support this study.

Additional Points

Highlights. The line attributes of the electrified intercity rail line, passenger gender, age, and availability of seats have a significant impact on passenger multitasking activities. On the leisure line, passengers preferred social activities, while middle-aged and elderly passengers preferred to engage in passive non-ICT activities, such as staring in a daze, and no obvious preference for resting was found with them. On the commuting line, passengers preferred to rest, and passengers without seats tended to do nothing. Male travelers and young travelers preferred ICT immersive activities, while women, middle-aged, and elderly travelers were more likely to engage in non-ICT immersive activities. Passengers with seats were more likely to engage in simultaneous multitasking activities, rest, and conversation than nonseated passengers.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The authors Xing Yao and Chunhui Jing contributed equally to this work.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant no. 52175253); MOE Layout Foundation of Humanities and Social Sciences (grant no. 19YJA760094); General Project of Sichuan Multicultural Research Center (grant no. DYWH2122); Sichuan Key Research Base of Philosophy and Social Sciences-National Park Research Center (grant no. GJGY2021-YB006); General Project of Sichuan Applied Psychology Research Center (grant no. CSXL-212A03); and Practice Base Project in the Graduate School of Southwest Jiaotong University. The authors thank LetPub (http://www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

References

[1] Y. Bousse, M. V. Corazza, G. Sessing, D. S. Arriaga, and D. Salzillo, "Electrification of public transport in Europe: vision and Practice from the ELIPTIC project," in *Proceedings of the 2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, June, 2018.

[2] M. Glotz-Richter and H. Koch, "Electrification of public transport in cities (horizon 2020 ELIPTIC project)," *Transportation Research Procedia*, vol. 14, pp. 2614–2619, 2016.
[3] L. Moretti and G. Loprencipe, "Climate change and transport infrastructures: state of the art," *Sustainability*, vol. 10, no. 11, p. 4098, 2018.
[4] L. Yang, J. Zhou, O. F. Shyr, and D. Huo, "Does bus accessibility affect property prices?" *Cities*, vol. 84, pp. 56–65, 2019.
[5] L. Yang, X. Chu, Z. Gou, H. Yang, Y. Lu, and W. Huang, "Accessibility and proximity effects of bus rapid transit on housing prices: heterogeneity across price quantiles and space," *Journal of Transport Geography*, vol. 88, Article ID 102850, 2020.
[6] L. Yang, K. W. Chau, W. Y. Szeto, X. Cui, and X. Wang, "Accessibility to transit, by transit, and property prices: spatially varying relationships," *Transportation Research Part D: Transport and Environment*, vol. 85, Article ID 102387, 2020.
[7] W. Xu, J. Zhou, L. Yang, and L. Li, "The implications of high-speed rail for Chinese cities: connectivity and accessibility," *Transportation Research Part A: Policy and Practice*, vol. 116, pp. 308–326, 2018.
[8] R. G. Mugion, M. Toni, H. Raharjo, L. Di Pietro, and S. P. Sebathu, "Does the service quality of urban public transport enhance sustainable mobility?" *Journal of Cleaner Production*, vol. 174, pp. 1566–1587, 2018.
[9] J. Zhou, L. Yang, J. Liu, and C. Zhang, "Beating long trips with a smartphone? A case study of Beijing residents," *Cities*, vol. 73, pp. 36–43, 2018.
[10] I. Keseru and C. Macharis, "Travel-based multitasking: review of the empirical evidence," *Transport Reviews*, vol. 38, no. 2, pp. 162–183, 2018.
[11] G. Circella, P. L. Mokhtarian, and L. K. Poff, "A conceptual typology of multitasking behavior and polychronicity preferences," *Electronic International Journal of Time Use Research*, vol. 9, no. 1, pp. 59–107, 2012.
[12] G. Lyons and J. Urry, "Travel time use in the information age," *Transportation Research Part A: Policy and Practice*, vol. 39, no. 2–3, pp. 257–276, 2005.
[13] M. L. Russell, "Watching passengers: using structured observation methods on public transport," in *Proceedings of the 43rd Universities Transport Study Group Conference*, Milton Keynes, UK, January 2011.
[14] P. A. Singleton, "Multimodal travel-based multitasking during the commute: who does what?" *International Journal of Sustainable Transportation*, vol. 14, no. 2, pp. 150–162, 2019.
[15] I. Keseru, E. Heyndels, T. Dat Ton, and C. Macharis, "Multitasking on the go: an observation study on local public transport in Brussels," *Travel Behaviour and Society*, vol. 18, pp. 106–116, 2020.
[16] K. Utsunomiya, "The impact of regional railways on travel behaviour and social capital," *Research in Transportation Economics*, vol. 83, Article ID 100945, 2020.
[17] R. M. Berliner, A. Malokin, G. Circella, and P. L. Mokhtarian, "Travel-based multitasking: modeling the propensity to conduct activities while commuting," in *Proceedings of the Transportation Research Board 94th Annual Meeting*, Washington, DC, USA, January 2015.
[18] G. Lyons, J. Jain, and I. Weir, "Changing times - a decade of empirical insight into the experience of rail passengers in Great Britain," *Journal of Transport Geography*, vol. 57, pp. 94–104, 2016.
[19] I. Keseru, J. Bulckaen, C. Macharis, J. Minnen, I. Glorieux, and T. P. Van Tienhoven, "Is travel time wasted? Evidence from a time use survey in Flanders, Belgium," in *Proceedings of the 14th International conference on travel behaviour research*, IATBR, Windsor, UK, July 2015.
[20] C. Frei, H. S. Mahmassani, and A. Frei, "Making time count: traveler activity engagement on urban transit," *Transportation Research Part A: Policy and Practice*, vol. 76, pp. 58–70, 2015.
[21] G. Lyons, J. Jain, Y. Susilo, and S. Atkins, "Comparing rail passengers' travel time use in great britain between 2004 and 2010," *Mobilities*, vol. 8, no. 4, pp. 560–579, 2013.
[22] H. Timmermans and P. Van der Waerden, "Synchronicity of activity engagement and travel in time and space," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2054, no. 1, pp. 1–9, 2008.
[23] S. Sudman and N. M. Bradburn, "Effects of time and memory factors on response in surveys," *Journal of the American Statistical Association*, vol. 68, no. 344, pp. 805–815, 1973.
[24] L. Gamberini, A. Spagnolli, A. Miotto, E. Ferrari, N. Corradi, and S. Furlan, "Passengers' activities during short trips on the London Underground," *Transportation*, vol. 40, no. 2, pp. 251–268, 2013.
[25] P. Van Der Waerden, H. Timmermans, and R. Van Neerven, "Extent, nature, and covariates of multitasking of rail passengers in an urban corridor," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2110, no. 1, pp. 106–111, 2009.
[26] A. Schneider, K. C. Seto, and D. R. Webster, "Urban growth in Chengdu, Western China: application of remote sensing to assess planning and policy outcomes," *Environment and Planning B: Planning and Design*, vol. 32, no. 3, pp. 323–345, 2005.
[27] N. Ohmori and N. Harata, "How different are activities while commuting by train? A case in tokyo," *Tijdschrift voor Economische en Sociale Geografie*, vol. 99, no. 5, pp. 547–561, 2008.
[28] K. Wu, C. Fang, M. Zhao, and C. Chen, "The intercity space of flow influenced by high-speed rail: a case study for the rail transit passenger behavior between Beijing and Tianjin," *Acta Geographica Sinica*, vol. 68, no. 2, pp. 159–174, 2013.
[29] A. Willis, N. Gjersoe, C. Havard, J. Kerridge, and R. Kukla, "Human movement behaviour in urban spaces: implications for the design and modelling of effective pedestrian environments," *Environment and Planning B: Planning and Design*, vol. 31, no. 6, pp. 805–828, 2004.
[30] R. C. P. Wong, W. Y. Szeto, L. Yang, Y. C. Li, and S. C. Wong, "Public transport policy measures for improving elderly mobility," *Transport Policy*, vol. 63, pp. 73–79, 2018.
[31] L. Yang, Y. Ao, J. Ke, Y. Lu, and Y. Liang, "To walk or not to walk? Examining non-linear effects of streetscape greenery on walking propensity of older adults," *Journal of Transport Geography*, vol. 94, Article ID 103099, 2021.
[32] E. Cornelis, M. Hubert, and P. Huynen, "Belgian Daily Mobility-BELDAM: Enquete sur la mobilite quotidienne des belges: rapport final," *Politique Scientifique federale*, Brussels, Belgium, 2012.
[33] J. Tang, F. Zhen, J. Cao, and P. L. Mokhtarian, "How do passengers use travel time? A case study of Shanghai-Nanjing high speed rail," *Transportation*, vol. 45, no. 2, pp. 451–477, 2018.
[34] T. Bjørner, "Time use on trains: media use/non-use and complex shifts in activities," *Mobilities*, vol. 11, no. 5, pp. 681–702, 2016.
[35] J. Zhang and H. Timmermans, "Scobit-based panel analysis of multitasking behavior of public transport users," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2157, no. 1, pp. 46–53, 2010.
[36] J. Jain and G. Lyons, "The gift of travel time," *Journal of Transport Geography*, vol. 16, no. 2, pp. 81–89, 2008.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 1570521, 14 pages
https://doi.org/10.1155/2022/1570521

Research Article

A Roadway Safety Sustainable Approach: Modeling for Real-Time Traffic Crash with Limited Data and Its Reliability Verification

Zhenzhou Yuan,[1] **Kun He**,[1] **and Yang Yang**[1,2]

[1]*School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China*
[2]*School of Transportation Science and Engineering, Beihang University, Beijing 100191, China*

Correspondence should be addressed to Yang Yang; 14114222@bjtu.edu.cn

Received 1 December 2021; Revised 19 December 2021; Accepted 27 December 2021; Published 15 January 2022

Academic Editor: Wenxiang Li

With the development of freeway system informatization, it is easier to obtain the traffic flow data of freeway, which are widely used to study the relationship between traffic flow state and traffic safety. However, as the development degree of the freeway system is different in different regions, the sample size of traffic data collected in some regions is insufficient, and the precision of data is relatively low. In order to study the influence of limited data on the real-time freeway traffic crash risk modeling, three data sets including high precision data, small sample data, and low precision data were considered. Firstly, Bayesian Logistic regression was used to identify and predict the risk of three data sets. Secondly, based on the Bayesian updating method, the migration test towards high and low precision data sets was established. Finally, the applicability of machine learning and statistical methods to low precision data set was compared. The results show that the prediction performance of Bayesian Logistic regression improves with the increasing of sample size. Bayesian Logistic regression can identify various significant risk factors when data sets are of different precision. Comparatively, the prediction performance of the support vector machine is better than that of Bayesian Logistic. In addition, Bayesian updating method can improve the prediction performance of the transplanted model.

1. Introduction

In recent years, the potential safety hazard of new energy vehicles has gradually attracted attention, especially the accident of pure electric vehicles [1, 2]. As new energy vehicles and shared vehicles enter the freeway, they are also facing the risk of traffic collisions. As one of the important subsystems of the road system, the freeway greatly facilitates people's travel and improves the transportation efficiency of goods. At the same time, because of the large traffic volume and fast speed of vehicles on freeways, relatively serious traffic crashes are easy to occur, which brings great harm to the safety of people's lives and property. Freeway traffic crash has become one of the problems that cannot be ignored [3].

A large number of scholars have done extensive researches on traffic safety. Some scholars analyzed the internal relationship between the factors causing accidents and the distribution law of accidents based on the historical traffic accident data collected and then put forward the corresponding countermeasures. For example, considering the differences in time, Yuan et al. adopted an improved association rule mining algorithm to analyze the association among the influencing factors of freeway traffic crashes, in which the hidden association rules were found and the accuracy of the algorithm was improved [4]. Tian et al. analyzed the temporal and spatial distribution characteristics of freeway crashes in mountainous areas based on historical crash data, identified the significant influencing factors, and proposed corresponding improvement strategies [5]. Some scholars have analyzed the main influencing factors of accidents according to specific accident scenarios. Mergia et al. analyzed the crash at the junction. It was pointed out that drunk driving and overspeed have increased the severity of crashes in the diverging area, bad weather increased the severity of crashes in the merging area, and adverse linear conditions

would increase the severity of crashes in the diverging area [6]. Xin et al. studied the factors not observed, proving that, under different severities, driving behaviors, environmental characteristics, and other factors have significant differences in the impact of the crash rate [7]. Haghighi et al. studied the impact of road design features on crashes and found that 10-foot wide lanes and narrower shoulder were significantly associated with crash severity and increased vehicle density and guardrail length could reduce crash severity [8]. In order to study the influence of drivers' ages on crash severity, Osman et al. constructed a generalized ordered response probit model to reduce the interference of heterogeneity and found that each variable in different age groups had a different influence on crash [9]. Xu et al. introduced the Bayesian spatial random coefficient model to consider the heterogeneity of spatial structure and unstructured data when studying the spatial variation law of crash rate and cause factors, which improved the fitting effect of the model and verified that the existence of spatial structure heterogeneity would cause bias to parameter estimation [10]. Wang et al. explored risk factors' influence on urban traffic crashes frequency while considering both the spatial and temporal correlation/heterogeneity of traffic crashes. The linear regression model, spatial lag model (SLM), spatial error model (SEM), and time-fixed effects error model (T-FEEM) were established and compared, respectively [11]. To figure out the factors relating to crash risk in different regional types and their inner relation, Yang et al. took three sections of highway (areas of downtown, suburb, and mountain, in Washington State, USA) as the research object and, based on AHP improved Apriori association rule mining algorithm, identified the crash risk influencing factors and their complex association rules were [12]. Li et al. investigated the possibility of using support vector machine (SVM) models for crash injury severity analysis and compared the performance of the SVM model and the order probit model. It was found that the SVM model produced better prediction performance for crash injury severity than did the OP model [13]. In addition, Logit and Tobit models are also widely used in traffic crash analysis [14–22].

With the development of freeway informatization and the improvement of dynamic traffic management, the real-time crash risk model has been widely studied [23–26]. Based on loop detector data and crash data collected by the Shanghai expressway system, Sun et al. established a Bayesian network (BN) model to analyze real-time traffic flow parameters and crash risk of expressway [27]. You et al. established a support vector machine model to analyze highway traffic flow data for rear-end crash. The results showed that the SVM classifier has high practical value and reliability of real-time crash prediction based on traffic flow data of a single volume detector [28]. Xu et al. established a crash risk prediction model based on traffic flow data and meteorological data by using the Logistic model based on American freeway data. The results showed that weather conditions have a significant impact on crash risk [29]. Ma et al. established a crash risk assessment and analysis model using highway crash data and real-time traffic flow data. The significant variables were selected by a random forest algorithm, and the support vector machine model was established. The evaluation ability of models under different kernel functions was compared. The results showed that the model could effectively evaluate road crash risk based on real-time traffic flow [30, 31].

The traditional "postevent" traffic safety analysis can analyze the main influencing factors of crash occurrence, but it is difficult to reflect the influence of dynamic traffic flow characteristics on crash risk. At present, most of the researches on using traffic flow data to establish real-time crash risk models are based on existing data for modeling and analysis. However, different regions have different levels of development, and the data collection of traffic flow and traffic crash will be different. Then does the traffic flow variable also have a significant impact on the occurrence of traffic crashes? If the impact is small, can certain technical means be used to improve the accuracy of the corresponding model?

In order to study the above questions, based on the basic data necessary for the real-time crash risk model, this paper constructs three types of data sets: (1) high precision data set; (2) small sample data set; (3) low precision data set. On the basis of the above three types of data sets, statistical and machine learning methods are used to study the classification and prediction performance of the model under different data sets, and the applicability of the two methods is further analyzed.

From the perspective of data, this paper uses statistical Logistic regression, Bayesian theory, and support vector machine to simulate the impact of different types of data on real-time crash risk modeling. Furthermore, the applicability of different methods under different data types is compared. The conclusions of this paper can be used as a reference for the subsequent practice and research of highway traffic safety.

2. Data Description

2.1. The Data Source. This paper selects the traffic flow and traffic crash data of milepost 100–132 of I-5 in Washington State in 2016. Figure 1 describes the main freeway section in the study area.

In 2016, a total of 332 traffic crashes occurred in this freeway section. In the selected area, 152 groups of loop detectors are arranged bidirectional, and the average

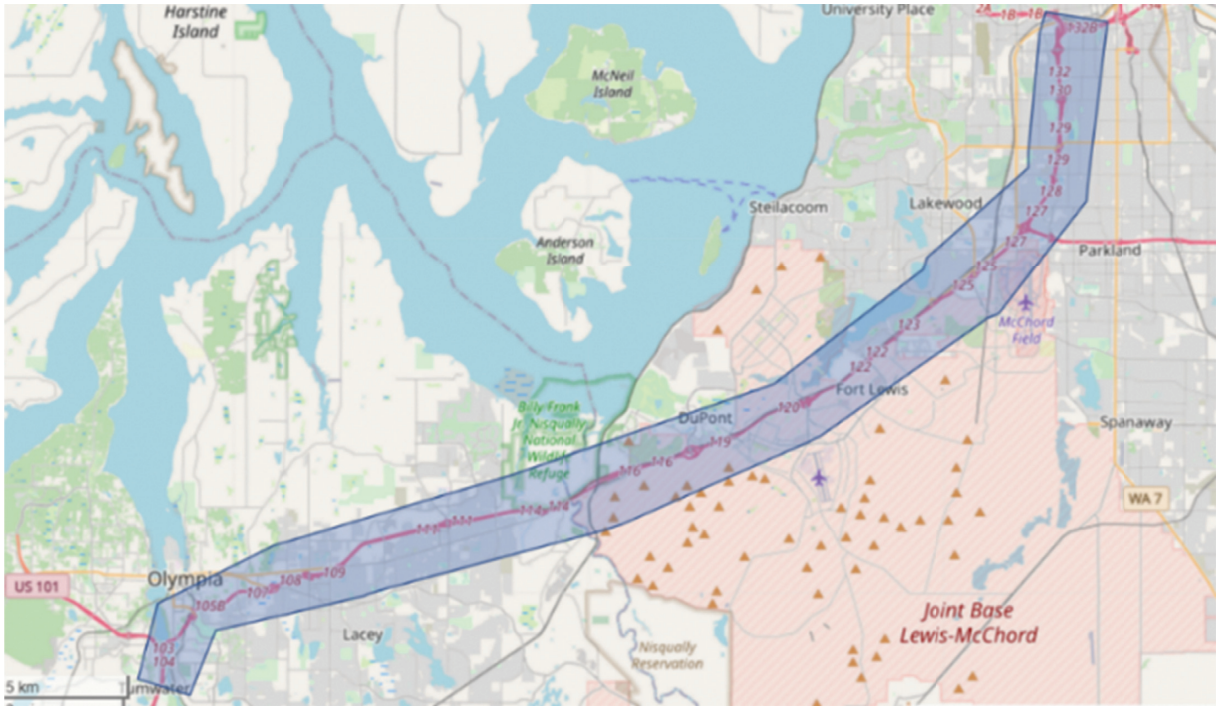


Figure 1: Study section of freeway.

distance between adjacent loops is about 0.7 km. Each loop detector collects average speed, occupancy, and traffic volume in each lane over a 20-second period.

2.2. Variable. In existing studies, traffic flow data of 5-minute lumps are used for analysis, and good research results are obtained [8–13]. Therefore, this paper adopts traffic flow data of 5-minute lumps for analysis, mainly including the volume, speed, and occupancy rate of each lane. The time period of 5–10 minutes before the crash is selected and two groups of upstream and downstream loop detectors were taken into account, as shown in Figure 2.

In addition to the basic data detected by the loop detector above, such as volume, speed, and occupancy of upstream and downstream loops, this paper combines the traffic flow variables as follows [2]. Considering that the difference values in volume, speed, and occupancy between upstream and downstream loops may lead to vehicle crash, the absolute value of the difference values between volume, speed, and occupancy between upstream and downstream loops is constructed. At the same time, the lateral crash between lanes is also one of the main forms of crash. The average difference values of volume, speed, and occupancy between adjacent lanes are constructed to describe the related variables of lateral collision. The specific meanings are shown in Table 1.

2.3. Sample Structure Design

2.3.1. High Precision Data Set Sample. High precision data set refers to the traffic flow data and traffic crash data collected by the American freeway system as the standard. The freeway loop

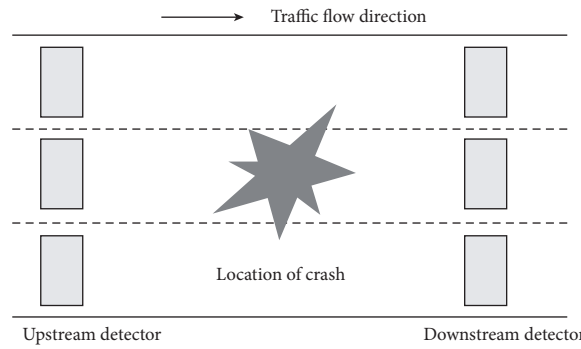


Figure 2: Traffic crash diagram.

detector in the United States has a high laying density, and the traffic crash information is collected completely.

In this paper, the paired sampling method is adopted to match control samples, and noncrash data under the same conditions are extracted for each crash data with a ratio of 1 : 4. The ratio of crash data to noncrash data is 1 : 4 for matching [2]. After data preprocessing, 191 traffic crash data and 764 noncrash data are obtained. The high precision data set sample is shown in Figure 3.

2.3.2. Small Sample Data Set. In order to study the influence of data sample size on the real-time crash risk model constructed, it is necessary to obtain small sample data with different sample sizes. The main ideas of constructing small sample data set in this paper are as follows: obtain and match the high precision data set, and extract data from the high precision data set in a proportion of 5%, 10%, 20%, 30%, and 50%, so as to construct small sample data set of different proportions. The small sample data set is shown in Figure 4.

Table 1: Variable name and physical meaning.

Variable	Meaning	Variable	Meaning
up_v	The upstream volume	*down_o*	Average downstream occupancy
up_s	Average speed of upstream traffic volume	*down_dif_v*	Average absolute value of volume difference between adjacent downstream lanes
up_o	Average upstream occupancy	*down_dif_s*	Average absolute value of speed difference between adjacent downstream lanes
up_dif_v	Average absolute value of volume difference between adjacent upstream lanes	*down_dif_o*	Average absolute value of occupancy difference between downstream adjacent lanes
up_dif_s	Average absolute value of speed difference between adjacent upstream lanes	*abs_dif_v*	Absolute value difference between upstream and downstream volume
up_dif_o	Average absolute value of occupancy difference between upstream adjacent lanes	*abs_dif_s*	Absolute value of speed difference between upstream and downstream
down_v	The downstream volume	*abs_dif_o*	Absolute value difference between upstream and downstream occupancy
down_s	Average speed of downstream traffic volume		

up_v	up_s	up_o	up_dif_v	up_dif_s	up_dif_o	down_v	down_s	down_o	own_dif	own_dif	own_dif	bs_dif_	bs_dif_	bs_dif_	y
−0. 68441	0. 163659	−0. 29517	−0. 32964	0. 329017	−0. 72624	−0. 36068	0. 546614	−0. 16984	−0. 05247	0. 682967	−0. 54752	−0. 68043	−0. 3625	−0. 68998	1
0. 421289	−0. 77021	1. 715074	−1. 41983	−0. 25792	−0. 73558	−0. 09158	−2. 13843	3. 059773	−0. 80973	0. 149726	−0. 53251	−0. 06695	1. 626693	0. 505555	1
−0. 51135	−0. 50202	0. 255297	−0. 69304	−0. 63884	0. 059538	1. 020696	0. 123974	1. 05477	0. 031674	−1. 38998	0. 574897	1. 35454	0. 032906	0. 050371	1
0. 652044	−0. 69696	−0. 16167	0. 985514	1. 47038	0. 245664	0. 626018	0. 006519	−0. 0475	−1. 07618	0. 390673	−0. 24468	−0. 90488	0. 159818	−0. 72195	1
1. 719285	−2. 41331	3. 127089	−0. 918	0. 383031	5. 506444	−1. 5178	−1. 13514	−1. 07261	−0. 09454	1. 33989	−0. 71158	4. 332187	1. 087805	4. 979543	1
0. 450133	−2. 71424	2. 65425	−1. 09104	−0. 92286	1. 530871	−1. 43707	−1. 93982	−0. 11723	0. 578587	−0. 68959	0. 309929	2. 222397	0. 185483	3. 179099	1
1. 738515	−2. 19604	2. 348171	−0. 88339	0. 216537	2. 587597	−0. 87196	0. 230103	−0. 7332	1. 335851	−0. 01299	−0. 21125	3. 284774	3. 120983	3. 511914	1
1. 103939	−1. 20243	1. 242066	−0. 81417	−0. 42792	1. 290966	0. 868207	−0. 5683	−0. 16987	0. 774915	−0. 49772	−0. 11744	−0. 60562	0. 012769	1. 329067	1
0. 969332	−0. 28245	−0. 17012	0. 1895	2. 10045	−0. 12562	1. 110395	0. 63278	−0. 20174	−1. 13227	−0. 62691	−0. 36594	−0. 80014	0. 553284	−0. 54618	1
−0. 29982	−0. 89755	0. 894417	1. 123951	0. 253046	0. 489311	−1. 00651	−2. 61024	4. 226433	−0. 47317	−0. 08751	0. 227479	0. 337053	2. 239837	3. 033843	1
−1. 65551	0. 110034	−0. 97014	0. 224109	0. 215935	−0. 31611	−0. 44141	−0. 05725	−0. 47678	−1. 21641	0. 415572	−1. 00761	0. 696166	−0. 53245	−0. 14774	1
0. 104001	−2. 25757	2. 774571	0. 1895	−0. 1591	3. 804259	1. 594773	−1. 9702	1. 927871	−0. 57133	−1. 01465	0. 292557	1. 35454	−0. 65011	0. 859884	1
−0. 43443	−0. 1565	−0. 07879	−0. 01816	0. 186951	−0. 62702	−0. 25304	−0. 95276	0. 696943	−0. 80973	−0. 40417	−0. 74973	−0. 88992	0. 588801	0. 068833	1
−1. 03054	−1. 50621	0. 592504	−0. 43347	0. 25027	0. 527696	−0. 46832	−1. 04277	0. 795404	−1. 02008	0. 037817	0. 186618	−0. 32132	−0. 30291	−0. 72145	1
−0. 1556	−1. 82467	1. 684171	−0. 84878	0. 567499	−0. 1299	−0. 13643	−1. 82607	1. 919292	−0. 72559	0. 226412	0. 231857	−0. 88992	−0. 72944	−0. 60752	1
0. 200149	0. 756686	−0. 12365	0. 812467	0. 889828	−0. 71253	0. 17752	1. 109791	−0. 09759	0. 031674	0. 169834	−0. 85501	−0. 85999	−0. 38589	−0. 60961	1
−0. 96324	−0. 46531	−0. 31264	−1. 10835	2. 544329	−0. 77833	−0. 57596	0. 275311	−0. 22955	−1. 23043	0. 439101	−0. 96425	−0. 60562	0. 236663	−0. 70559	1
1. 219316	0. 22487	0. 077173	−0. 15659	0. 004811	0. 239609	1. 379494	0. 00871	0. 565516	0. 073744	1. 735347	0. 056289	−0. 74029	−0. 45199	−0. 30206	1
−0. 61711	0. 497055	−0. 75549	−0. 84878	−0. 15521	−0. 50233	0. 168551	0. 430877	−0. 39905	−1. 24446	−0. 74949	−0. 6145	0. 097644	−0. 72963	−0. 34432	1
1. 48853	−0. 86834	0. 931766	−0. 1912	−0. 52749	1. 148165	1. 433313	−0. 81331	0. 34867	−1. 10422	0. 243402	0. 104932	−0. 94977	−0. 87595	0. 279127	1
−1. 51128	−3. 53585	6. 291478	−1. 78323	−1. 6648	0. 774712	−1. 22179	−2. 48337	0. 723412	0. 915149	−0. 84303	2. 572265	−0. 83007	0. 634922	7. 088264	1

Figure 3: High precision data set sample.

up_v	up_s	up_o	up_dif_v	up_dif_s	up_dif_o	down_v	down_s	down_o	own_dif	own_dif	own_dif	bs_dif_	bs_dif_	bs_dif_	y
0. 402059	−1. 81665	1. 619915	2. 231449	0. 799052	3. 621474	1. 191125	−0. 96251	0. 504233	2. 822334	0. 258456	2. 641339	0. 217348	0. 370138	1. 022954	1
−0. 35751	−0. 87518	0. 012122	−1. 57558	0. 680598	−0. 18559	−0. 16334	−0. 80451	−0. 30894	0. 01765	−0. 30431	−0. 64557	−0. 85999	−0. 90317	−0. 16908	0
0. 027082	0. 702722	−0. 46017	−0. 08737	0. 647467	0. 054882	0. 3031	0. 18034	−0. 2712	−0. 79571	0. 300822	−0. 35774	−0. 68043	0. 06398	−0. 58938	0
0. 690503	−0. 0901	0. 096558	0. 656725	−0. 09514	−0. 33982	1. 98048	0. 594676	−0. 55407	−0. 4311	0. 504736	−0. 52743	1. 085205	0. 156749	0. 242889	0
2. 363476	−1. 51613	1. 27373	−0. 57191	−0. 12078	0. 57125	−2. 12776	−3. 29392	7. 387506	−1. 02008	−1. 42513	3. 3015	6. 352199	2. 384989	6. 356577	0
−0. 64595	−0. 44596	−0. 07684	−1. 62749	0. 958155	−0. 29038	0. 26722	−0. 97947	1. 138365	−1. 10422	1. 596822	−0. 47518	0. 307126	0. 140141	0. 601986	1
−1. 57859	0. 54386	−0. 9232	−0. 01816	−1. 38892	−0. 60037	−1. 32943	1. 009968	−0. 87773	2. 219327	−0. 58681	0. 113653	−0. 90488	−0. 19729	−0. 69803	0
0. 267452	0. 172516	−0. 89415	0. 293327	−0. 33743	−0. 30269	−0. 36068	0. 39668	−0. 26415	1. 08343	0. 514419	0. 646959	0. 142533	−0. 64182	0. 007344	0
−1. 02093	0. 941156	−0. 66739	−0. 58921	1. 571614	−0. 87815	−0. 71947	1. 043031	−0. 59838	−0. 51524	−1. 03445	−0. 83104	−0. 75525	−0. 81942	−0. 70566	0
1. 084709	−1. 3259	1. 153409	−0. 36425	−0. 38715	−0. 13133	1. 845931	−0. 76926	0. 673939	0. 129838	−0. 10407	0. 41764	0. 247274	−0. 12984	0. 184718	0
−0. 61711	0. 479366	−0. 50921	0. 085672	−0. 89762	−0. 64201	−0. 37862	0. 619881	−0. 39593	−0. 26282	−0. 45781	−1. 01116	−0. 8151	−0. 77408	−0. 67431	0
−0. 00176	0. 415662	0. 112207	0. 241414	2. 523276	0. 055939	0. 814387	0. 692801	−0. 28082	−0. 34696	3. 318966	−0. 50381	0. 217348	−0. 53652	−0. 06756	0
−0. 45366	0. 961781	−0. 7281	−1. 00452	−0. 42489	−0. 75109	0. 016061	−0. 44247	−0. 38436	0. 354212	−0. 55563	−0. 24947	−0. 4111	1. 604886	−0. 36362	0
0. 238608	0. 48221	−0. 28059	−0. 90069	0. 490578	−0. 67397	0. 096791	−0. 72272	−0. 19109	1. 223664	−0. 01875	0. 348149	−0. 66547	1. 277339	−0. 70883	1
−1. 67473	−0. 27083	−0. 91104	−0. 90069	0. 256389	−0. 45977	−1. 03342	−0. 6536	−1. 18339	0. 213978	0. 654404	−0. 7359	−0. 26147	−0. 13406	−0. 35885	0
1. 123168	−0. 74023	0. 0853	−0. 01816	0. 115887	0. 302784	−0. 74638	−0. 22387	−0. 55407	0. 494447	0. 498116	−0. 46142	2. 117656	−0. 17197	0. 227631	0
−1. 03054	−1. 19417	−0. 17364	−0. 84878	−1. 17074	−0. 79538	−1. 08724	−1. 52766	0. 102975	0. 031674	1. 122836	0. 38967	−0. 66547	−0. 17543	−0. 52356	0
0. 700118	0. 380616	−0. 30148	0. 743249	0. 151765	−0. 01418	0. 25825	0. 816843	0. 310958	−0. 59938	1. 461072	0. 070187	−0. 21658	−0. 25797	−0. 09785	0
0. 82511	0. 077623	0. 247838	2. 577542	0. 543329	0. 502706	1. 218035	−0. 26948	−0. 08738	2. 85038	−0. 69332	0. 730347	−0. 39614	−0. 21406	−0. 11853	0
−1. 28053	0. 966189	−0. 73091	0. 241414	−1. 21338	−0. 53491	−1. 55368	0. 887645	−0. 95545	0. 368236	−1. 19342	−0. 6752	−0. 27643	−0. 73094	−0. 39139	0
0. 834725	0. 390852	−0. 2818	0. 500983	−0. 33818	−0. 48605	0. 599108	0. 089139	−0. 32308	−0. 0104	0. 03159	−0. 57608	−0. 57569	−0. 30944	−0. 55027	0

Figure 4: Small sample data set.

2.3.3. Low Precision Data Set Sample. Low precision data set refers to the data set constructed from the data collected by the detectors with lower density compared to the US freeway system. Considering that the data is difficult to obtain, this paper constructs a low precision data set through certain manual processing methods. Compared with the freeway system in the United States, many freeways in China do not have complete loop detection devices, and the distance between the detectors is relatively long. In this paper, the average distance between detectors of a certain section of freeway in China is taken as a reference, and the freeway data of the US is used to construct a low precision data set. The low precision data set sample is shown in Figure 5.

The main processing ideas are shown in Figure 6. Manually delete part of the loop number in the loop file so that the average distance between the remaining loops is approximately equal to the reference value. Then the processed loop file is used to match the data set to get the low precision data set. After screening, there are a total of 32 bidirectional loops. The low precision data set is screened by paired sampling method, and 161 crash data and 644 noncrash data were obtained.

3. Real-Time Crash Risk Prediction Model

3.1. Bayesian Logistic Regression. Logistic regression is a generalized linear regression model commonly used in statistical methods. Based on the binomial Logistic regression model, this paper establishes a crash risk model between freeway crashes and real-time traffic flow [23, 28]. The crash probability corresponding to a certain data in the research data set is shown as follows:

$$P(x_i) = \frac{1}{1 + e^{-x_i'\beta}}, \quad i = 1, 2, \ldots, n, \tag{1}$$

where x_i represents the i th data; $P(x_i)$ represents the probability value of crash occurrence; $-x_i'\beta$ represents a linear combination of explanatory variables and their coefficients.

The Bayesian method is used to estimate the coefficients of the Logistic regression model. The Bayesian method assumes that all unknown parameters in the model are random variables. Before establishing the Bayesian model M, it is necessary to set the prior probability distribution $\pi(\Theta \mid M)$ of all parameters Θ in the model, which represents the known information of this parameter before obtaining the training data Y. After obtaining the training data Y, the Bayesian statistical model makes statistical inference on Θ through the posterior probability distribution. According to the Bayesian theorem, the posterior probability distribution $f(\Theta \mid Y, M)$ of parameter Θ in model M can be expressed as follows:

$$f(\Theta \mid Y, M) = \frac{f(Y, \Theta \mid M)}{f(Y \mid M)} = \frac{f(Y \mid \Theta, M)\pi(\Theta \mid M)}{\int f(Y, \Theta \mid M) d\Theta} \propto f(Y \mid \Theta, M)\pi(\Theta \mid M). \tag{2}$$

Formula (2) shows that the posterior probability distribution of parameter Θ takes into account both the information contained in the training data Y and the known information of parameter Θ. $f(\Theta \mid Y, M)$ is the posterior distribution of parameter Θ in model M under given training data Y. $f(Y, \Theta \mid M)$ is the joint probability distribution of Y and Θ in model M. $f(Y \mid M)$ represents the marginal probability distribution of model M, that is, the probability distribution of training data Y under given conditions. $\pi(\Theta \mid M)$ represents the prior probability distribution of parameter Θ in model M before obtaining the training data Y. $f(Y \mid \Theta, M)$ is the likelihood function of model M.

3.2. Bayesian Updating Method. Based on the Bayesian updating method, the Bayesian Logistic regression model is established to transmigrate the real-time crash of freeways [32]. The Bayesian method can obtain the posterior probability distribution of each parameter in the model so that the prior probability distribution can be reset during model transplantation.

That is, when low precision data set Y_1 is used to establish a real-time crash risk model, the Bayesian method can be used to obtain the posterior probability distribution of each risk factor. When high precision data set Y_2 needs to establish the Logistic regression model and transplant it to low precision data set, the posterior probability distribution of the risk factors in the previous model can be used as the prior probability distribution of the risk factors in the new model, as shown in the following formula:

$$\begin{aligned} \pi(\beta \mid Y_1, Y_2) &\propto f(Y_1, Y_2 \mid \beta)\pi(\beta) \\ &= f(Y_2 \mid Y_1, \beta) f(Y_1 \mid \beta)\pi(\beta) \propto f(Y_2 \mid \beta)\pi(\beta \mid Y_1). \end{aligned} \tag{3}$$

Schematic diagram of the Bayesian updating method is shown in Figure 7.

$\pi(\beta \mid Y_1, Y_2)$ is the posterior distribution of parameter β under given data sets Y_1 and Y_2; $f(Y_1, Y_2 \mid \beta)$ is the likelihood function; $\pi(\beta)$ is the prior probability distribution of parameter β; $f(Y_2 \mid Y_1, \beta)$ is the likelihood function given data set Y_1 and parameter β; $f(Y_1 \mid \beta)$ and $f(Y_2 \mid \beta)$ are the likelihood functions; $\pi(\beta \mid Y_1)$ is the posterior distribution of parameter β under given data set Y_1.

3.3. Support Vector Machine. Support vector machine (SVM) is a machine learning classification algorithm based on statistical theory [26, 27]. It can obtain the optimal solution through existing information and can deal with small samples or limited samples well. In the sample space, the linear SVM divides the hyperplane by $\omega^T x + b = 0$ to distinguish the labeled data set, where ω is the normal vector and b is the displacement term.

The distance between any point x in the sample space and the hyperplane can be written as

$$r = \frac{\left|\omega^T (x + b)\right|}{\|\omega\|}. \tag{4}$$

up_v	up_s	up_o	up_dif_v	up_dif_s	up_dif_o	down_v	down_s	down_o	own_dif_	own_dif_	own_dif_	abs_dif_	abs_dif_	abs_dif_	y
−0. 35797	0. 626395	−0. 26101	−0. 03055	−1. 25263	−0. 61971	−0. 26454	0. 308747	−0. 08715	−0. 64797	0. 959104	−0. 23376	−0. 96101	−0. 30024	−0. 46374	1
1. 618194	−1. 42883	0. 932333	−0. 56345	1. 293378	−0. 58119	0. 325458	−2. 11207	2. 777052	−0. 68292	−0. 01154	−0. 41908	1. 139127	0. 695528	1. 75901	1
0. 587152	0. 340272	0. 094922	1. 20274	−0. 70182	−1. 02375	−0. 66166	0. 690003	−0. 42753	−0. 26358	−1. 26354	−0. 71554	0. 995936	−0. 79279	−0. 05608	1
1. 306733	0. 896161	−0. 01299	−1. 38564	−0. 17417	−0. 46032	1. 233148	−0. 06932	0. 025452	−0. 90423	0. 210748	−0. 14847	−0. 59508	0. 438862	−0. 54873	1
0. 372352	−2. 06269	1. 594566	−1. 35519	−0. 83324	1. 654453	−1. 47858	−1. 15658	−0. 88232	−0. 08886	1. 086471	−0. 58745	1. 823264	−0. 72722	1. 744638	1
0. 168291	0. 506585	−0. 46551	0. 471904	−0. 81691	0. 041384	−1. 37646	−1. 92292	−0. 0363	0. 470254	−0. 78588	0. 372971	1. 37778	2. 371049	−0. 23084	1
1. 231553	0. 548245	−0. 13686	−0. 92887	−1. 11472	−0. 21109	−0. 66166	0. 143606	−0. 58176	1. 099256	−0. 16166	−0. 11704	1. 950545	−0. 17101	−0. 07998	1
1. 231553	0. 52015	−0. 21716	0. 319646	0. 067754	−0. 10036	−0. 27589	−0. 07157	−0. 42592	0. 225642	−0. 09079	−0. 37185	1. 4096	0. 070218	−0. 3254	1
1. 961874	1. 165155	0. 124981	−1. 17248	−0. 77047	0. 254079	1. 845839	0. 527096	−0. 11114	−0. 95083	−0. 72805	−0. 26248	−0. 48371	−0. 04094	−0. 38884	1
1. 360433	−2. 36298	2. 333208	−0. 83751	0. 330979	4. 174152	−0. 95666	−3. 13368	2. 98878	−0. 50819	−1. 10678	3. 391442	2. 555131	1. 049727	0. 798217	1
−0. 42241	0. 440146	−0. 43721	−0. 85274	−1. 46675	−0. 52668	0. 155266	0. 661721	−0. 45338	−0. 32182	−0. 24746	−0. 33702	−1. 1042	−0. 92546	−0. 48289	1
−0. 75535	0. 285662	−0. 2705	−1. 35519	0. 083988	−0. 55676	−0. 55954	0. 516482	−0. 30366	2. 322317	−0. 01227	0. 911983	−1. 13602	−0. 89652	−0. 50965	1
1. 489313	−0. 15358	0. 365465	−0. 19803	0. 164962	0. 497135	−0. 09435	−1. 78248	−0. 25894	1. 239035	−0. 4738	0. 098564	1. 536881	1. 543322	−0. 01524	1
1. 467833	0. 834087	−0. 18245	−1. 14203	−0. 8546	−0. 27352	0. 166612	0. 297302	−0. 4287	0. 679922	−0. 41383	−0. 46005	1. 139127	−0. 08075	−0. 29249	1
−0. 26131	−3. 87334	5. 34829	−0. 56345	−0. 68677	4. 256876	−1. 45588	−1. 24396	−0. 92407	0. 400365	−0. 87019	−0. 60705	0. 852745	0. 952264	5. 009321	1
0. 887872	1. 315064	−0. 46278	−1. 20293	−0. 90746	−1. 05294	1. 641609	1. 068389	−0. 22264	−0. 74116	−0. 76605	−0. 45744	−0. 96101	−0. 56948	−0. 44433	1
−2. 18377	0. 671476	−1. 1686	0. 334872	0. 239919	−0. 87324	−2. 51108	0. 182772	−1. 01382	−0. 99742	−1. 0001	−0. 64233	−0. 51553	−0. 09823	−0. 47471	1
1. 360433	0. 062559	0. 15608	−1. 50744	0. 446779	−0. 45341	−0. 24185	−0. 31503	−0. 28399	0. 749811	0. 315881	−0. 10724	1. 552791	−0. 0775	−0. 16632	1
1. 489313	0. 746277	−0. 08093	−0. 47209	0. 268719	−0. 13234	−0. 36666	−1. 3948	−0. 48993	0. 889589	−1. 40465	−0. 29345	1. 918725	1. 947669	−0. 1361	1
−0. 63721	−0. 04861	−0. 16601	−1. 44654	2. 508246	−0. 60973	−0. 083	0. 441512	−0. 4357	0. 097512	0. 105883	−0. 18296	−1. 12011	−0. 71924	−0. 27048	1
1. 317473	0. 375682	0. 05136	−0. 95932	0. 768534	−0. 05303	−0. 59358	−1. 33532	−0. 64299	0. 784755	−0. 10643	−0. 57765	1. 982365	1. 50722	0. 150726	1

Figure 5: Low precision data set sample.

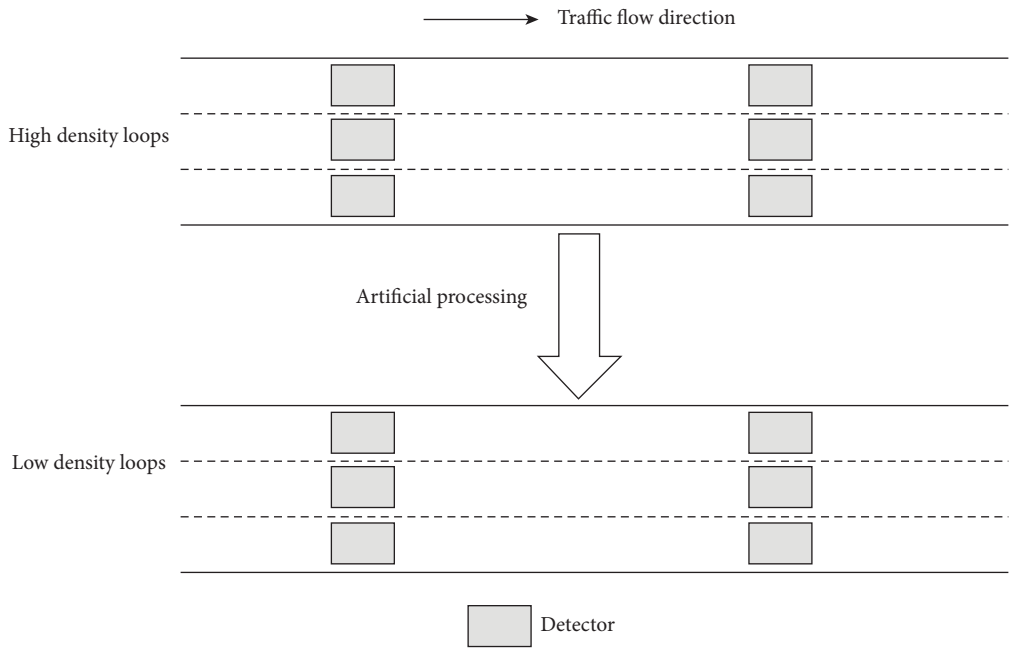


Figure 6: Schematic diagram of freeways with different loop densities.

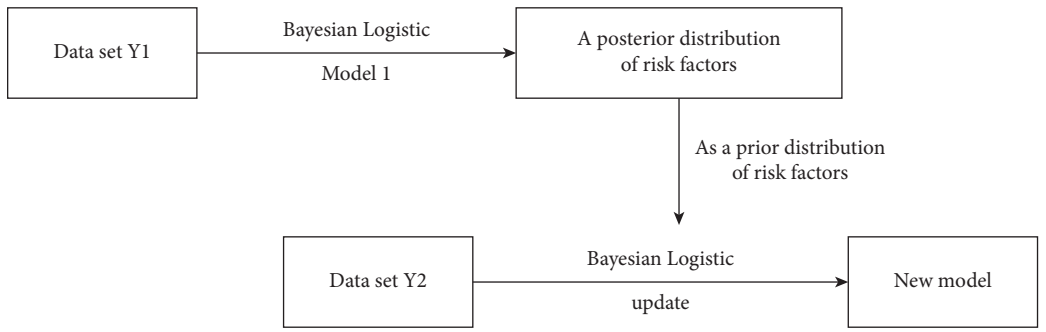


Figure 7: Schematic diagram of Bayesian updating method.

For labeled sample data sets $Y = (y_1, y_2, \ldots, y_n,)$, +1 is accident data, and −1 is nonaccident data. If it can be correctly classified, it can be

$$\begin{cases} \omega^T x_i + b \geq +1, y_i = +1, \\ \omega^T x_i + b \leq -1, y_i = -1. \end{cases} \tag{5}$$

When the sample dimension is high, it may lead to linear inseparability of sample data. The processing method of SVM for this situation is to raise the dimension of the sample data, convert the linear nonfraction data in the low dimensional space into linearly separable data in the high-dimensional space, and then use the linear SVM to find the optimal classification surface in the high-dimensional space.

3.4. The Evaluation Index. In the data classification model, accuracy can intuitively display the overall classification performance of the model, which is expressed as the proportion of the correctly classified sample results in the total sample among all samples as shown in the following formula:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{6}$$

where TP represents the number of samples predicted to be positive; TN represents the number of samples that predicted negative classes as negative classes; FP represents the number of samples that predicted the negative category as the positive category; FN represents the number of samples that predicted positive classes as negative classes.

The confounding matrix shown in Table 2 can directly display the classification results of the model and calculate the corresponding true positive rate (TPR) and false-positive rate (FPR) indexes.

$$\begin{aligned} \text{TPR} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{FPR} &= \frac{\text{FP}}{\text{FP} + \text{TN}}. \end{aligned} \tag{7}$$

The receiver operating characteristic (ROC) curve can be drawn by using TPR and FPR. The ROC represents the curve of the prediction accuracy of the data set under different probability thresholds. The AUC value of the area under the ROC curve can be calculated to measure the quality of the model. The closer the AUC value is to 1, the better the performance of the model.

4. Results and Discussion

4.1. Analysis of the Results from Small Sample Data Sets. In order to study the influence of different sample size data sets on the established crash risk model, Bayesian Logistic regression is used to establish models for the extracted 5%, 10%, 20%, 30%, and 50% high precision data sets, respectively. Table 3 shows the significant risk factors screened by Logistic stepwise regression for each data set.

By comparing the significant risk factors of the models with different sample size data sets, it is found that the sample size does affect the real-time crash risk model. Different models not only share the same (i.e., the same impact factors) but also have their own characteristics. It provides a basis for subsequent analysis.

As can be seen from Table 3, the speed of the upstream loop (*up_s*) is a significant variable of each data set. It shows that, for different data sets, upstream is a significant factor affecting the occurrence of crashes, which plays an important role in explaining the causes of crashes. At the same time, there were differences in other risk factors among each data set. Some risk factors were significant in one small sample, but not in others. This shows that each small sample data set has different characteristics and has certain differences for the establishment of the real-time crash risk model.

Figure 8 is the ROC curve and AUC value diagram of the real-time crash risk model established by the Bayesian Logistic regression method with different small sample data sets.

As can be seen from Figure 8, the change of the AUC value of the real-time crash risk model established by Bayesian Logistic regression does not increase with the increasing of sample size of the data set but is in a state of fluctuation. However, the overall trend of AUC value is decreasing.

For each collision, the traffic flow state is different. When the sample size of the data set is different, the structure of the data set is more complex. The significant risk factors screened by the Bayesian Logistic model established for each set of data mainly explain the predictive classification effect of the data set. And the significant risk factor of each data set is the optimal combination screened by the model. Therefore, there will be different results of collision precursors under different data sets. As mentioned above, as the sample size increases, the data set structure becomes more complex. Under the given combination of significant risk factors in different data sets, the probability of accidents is more complex, so the AUC index of the model will be reduced.

With the increase of sample size, the number of traffic crashes also increases. At this point, the diversity of traffic flow states of traffic crashes increases. From the screening of risk factors, it can be seen that the combination of risk factors changes with different sample sizes. Both the common significance factors (such as *up_s*) and the unique significance factors of each data set are included. Different traffic flow states make the data structure diverse. Therefore, the accuracy of models based on different data sets may be reduced. At the same time, due to the increase of data volume, the amount of traffic crash data increases. Although the AUC value of the overall model decreases, it is still around 0.7 or even a little higher, indicating that the number of traffic crashes correctly classified increases, too. It also indicates that while the sample size increases, though the sample structure is diverse, the law of data can be extracted as the sample size increases. This is the improved classification performance of the model. The increase of sample size can improve the prediction performance of the real-time crash risk model.

Table 2: Confusion matrix.

		Predict		Total
		Positive	Negative	
Actual	Positive	TP	FN	TP + FN
	Negative	FP	TN	FP + TN
Total		TP + FP	FN + TN	TP + FN + FP + TN

Table 3: Significant variables in small sample data sets.

Data sets	Risk factors
5% small sample	*up_s*
10% small sample	*up_s*, *down_v*, *abs_dif_s*, and *up_dif_v*
20% small sample	*up_s*, *down_s*, *up_o*, and *down_v*
30% small sample	*up_s* and *down_dif_v*
50% small sample	*up_s*, *down_dif_v*, *up_dif_s*, and *abs_dif_o*

At the level of significance of 0.05.

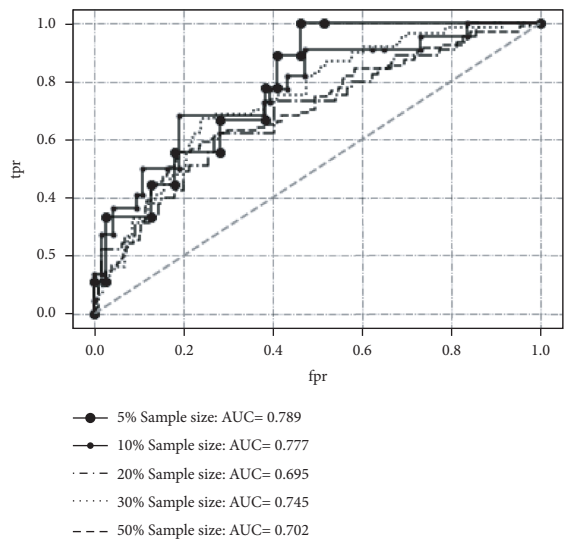


Figure 8: The ROC curve and AUC value of the model were established with a small sample data set.

4.2. Reliability Verification of Model Transferability

4.2.1. Low Precision Data Set and High Precision Data Set Risk Model Comparison. Stepwise Logistic regression is used to screen the factors that had a significant impact on the risk model, and the Bayesian method is used to estimate the coefficients of the model. Table 4 shows the model comparison after the coefficient estimation of significant risk factors.

As can be seen from Table 4, there are partially the same explanatory variables with significant levels between low precision and high precision data sets, such as *abs_dif_o* and *up_s*. These two explanatory variables indicate that, in the two data sets, both the speed of upstream loop and the absolute value of occupancy difference between upstream and downstream loop can effectively explain the causes of crashes. The difference between the two sets of data sets is that *up_dif_o* in the low precision data set has a significant explanatory effect on the model, while *down_dif_v* and *up_dif_s* in the high precision data set have a stronger explanatory effect on the model. In the coefficient estimation, it can be found that the estimated 95% confidence interval of each coefficient does not contain 0, indicating that the coefficient estimation is significant. The average upstream speed (*up_s*) coefficient of the two models is negative, indicating that, within the specified driving speed range of expressways, the decrease of average upstream speed by one unit will lead to an increase in crash risk.

Each piece of data is classified, and the accuracy of the model established by the two sets of data sets is shown in Table 5.

At the same time, the confusion matrix of two data sets for model classification and prediction is shown in Tables 6 and 7.

The ROC curve and AUC values of the model established on the basis of the two data sets are shown in Figure 9.

Through the comparison of the above indicators, it can be found that when the data set established with relatively sparse loop density is used to establish the Bayesian Logistic regression model, the classification accuracy of the model is 70.68%, slightly lower than the classification accuracy of the high precision data set with large loop density which is 73.30%. However, the model AUC value of low precision data set is 0.656, which is much smaller than that of high precision data set. The reasons are as follows: there are few explanatory factors in the low precision data set, and the data information is lost, which affects the accuracy of the model to some extent. In contrast, when the loop density is larger, more traffic flow information can be collected, and traffic variables that have a significant impact on traffic crashes can be screened out, thus making the model more accurate.

4.2.2. The Application of the Model of the High Precision Data Set-Based Model Low Precision Data Set. Applying the model of high precision data set to low precision data set, the classification accuracy is 69.3%, and the confusion matrix is shown in Table 8.

The ROC curve and AUC value obtained are shown in Figure 10.

Directly applying the model established by high precision data sets to low precision data sets, the classification results are worse than those established by previous low precision data sets. When using Logistic regression to screen variables, the optimal variable combinations of the two data sets are different. In the process of parameter estimation, the model obtained is the best fit of the best variables under each set of data. Therefore, when applied to other data sets, there

Table 4: Significant risk factors and parameter estimates for the two data sets models.

Variable	Low precision data set		High precision data set	
	Mean	Confidence interval	Mean	Confidence interval
Constant	−1.477	(−1.657, −1.294)	−1.5734	(−1.755, −1.391)
abs_dif_o	0.183	(0.008, 0.358)	0.276	(0.102, 0.456)
up_s	−0.299	(−0.507, −0.095)	−0.608	(−0.784, −0.427)
down_dif_v	—	—	−0.209	(−0.387, −0.037)
up_dif_s	—	—	0.250	(0.087, 0.405)
up_dif_o	0.257	(0.060, 0.454)	—	—

Table 5: Classification accuracy of Bayesian Logistic regression.

Data sets	Low precision data set (%)	High precision data set (%)
Accuracy	70.68	73.30

Table 6: Low precision data set confusion matrix.

	Positive (predict)	Negative (predict)
Positive (actual)	74	87
Negative (actual)	149	495

Table 7: High precision data set confusion matrix.

	Positive (predict)	Negative (predict)
Positive (actual)	116	75
Negative (actual)	180	584

will be inapplicable situations. It can be seen that direct transplantation of the model cannot achieve a better prediction classification effect.

4.2.3. Bayesian Updating towards High Precision Data Set Model and Low Precision Data Set Model. (a) The Bayesian updating method was used to update and transplant the model established by the original high precision data set. The posterior distribution of parameter estimation of the low precision data set model variables was regarded as the prior distribution of parameter estimation of the high precision data set model and then updated it. The results obtained are shown in Table 9.

The classification accuracy of the updated high precision data model is 68.94%, and the confusion matrix is shown in Table 10. ROC curve and AUC value obtained are shown in Figure 11.

(b) The Bayesian updating method was used to update and transplant the model established by the original low precision data set. The posterior distribution of parameter estimation of the high precision data set model variables was regarded as the prior distribution of parameter estimation of the low precision data set model and then updated it. The results obtained are shown in Table 11.

The classification accuracy of the updated low precision data set model is 70.83%, and the confusion matrix is shown in Table 12. ROC curve and AUC value obtained are shown in Figure 12.

By updating the model established by the high precision data set, it can be found that the prediction accuracy of the model cannot be effectively improved when the model is applied to the low precision data set before and after updating. The prediction accuracy is 69.3% before the update and 68.94% after the update, which decreases by 0.36%, and the AUC value decreases by 0.002.

By updating the model established by the low precision data set, it can be found that the prediction accuracy of the model is 70.68% and 70.83%, respectively, before and after the model is applied to the low precision data set, and the classification accuracy is improved by 0.15%. In addition, the AUC value increases from 0.656 to 0.657.

The classification accuracy of the model is 70.68% in the low precision data set. Based on the evaluation index of the model, the classification accuracy of the model transplant results is improved to a certain extent. Therefore, 69.3% could not meet the requirements, while the result of another model transplantation reached the requirements of 70.83%. In comparison, the improvement of the model is smaller, only 0.15%. However, in the field of traffic safety, it will have practical application significance to improve certain accuracy. In the follow-up research, better models or methods can be further proposed to make the results of model transplantation better. It can be seen from the above that the Bayesian updating method can improve the model transplantation effect to a certain extent, but the overall effect is limited, indicating that this method can indeed carry out model transplantation. The reason for the limited improvement may be that the most significant factor with explanatory effect has been screened out during the stepwise regression, and the difference between the parameters estimated by the Bayesian method and the parameters estimated by maximum likelihood estimation is small. At this time, the parameters of the model have become an excellent combination of parameter values. In the process of Bayesian update, the prior information of the model has little influence on it, so the overall improvement effect of the model is small.

In addition, it is necessary to determine the updated model object. Through the above study and comparison, it can be concluded that updating models with low precision data sets will obtain models with higher prediction performance.

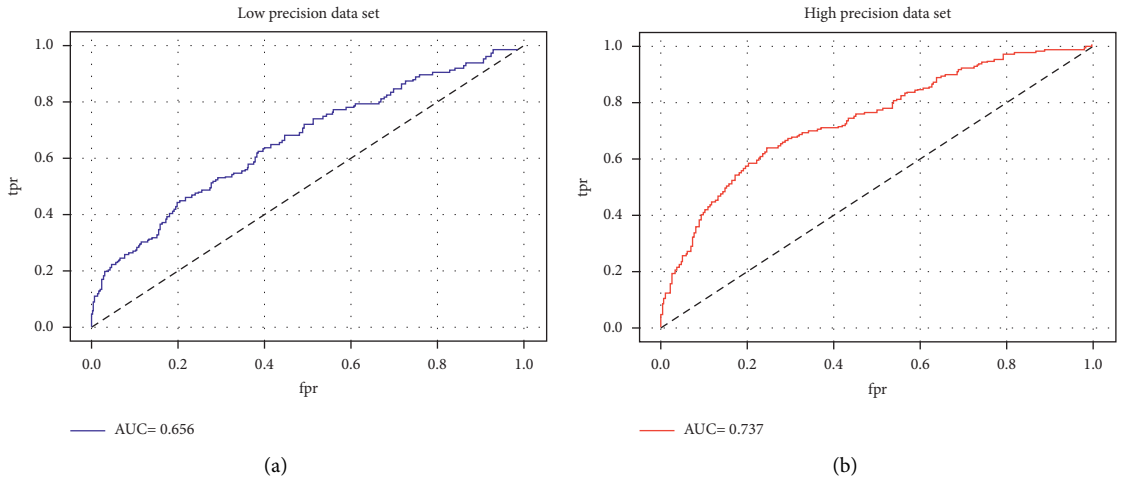


Figure 9: ROC curve and AUC values of the model were established for the two sets of data.

Table 8: The high precision data set model is applied to the low precision data set.

	Positive (predict)	Negative (predict)
Positive (actual)	79	82
Negative (actual)	165	479

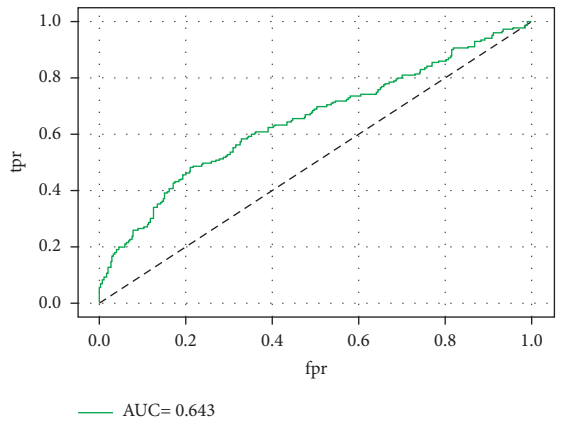


Figure 10: The high precision data set model is applied to the low precision data set.

Table 9: Results of the updated high precision data set model.

Variable	Before		After	
	Mean	Confidence interval	Mean	Confidence interval
Constant	−1.573	(−1.391, 2.474)	−1.540	(−1.715, −1.365)
abs_dif_o	0.276	(0.102, 0.456)	0.267	(0.148, 0.385)
up_s	−0.608	(−0.784, −0.427)	−0.0.494	(−0.624, −0.364)
down_dif_v	−0.209	(−0.387, −0.037)	−0.212	(−0.387, −0.038)
up_dif_s	0.250	(0.087, 0.405)	0.236	(−1.715, −1.365)

4.3. Classification Prediction Model Based on SVM. In order to be consistent with previous research methods, data sets were not divided into training data sets and test data sets. Classification prediction models for high precision and low precision data sets are established based on SVM, and the classification accuracy and confusion matrix are obtained as in Tables 13–15.

ROC curve and AUC values are shown in Figure 13.

Table 10: Confusion matrix of the updated high precision data set model.

	Positive (predict)	Negative (predict)
Positive (actual)	81	80
Negative (actual)	170	474

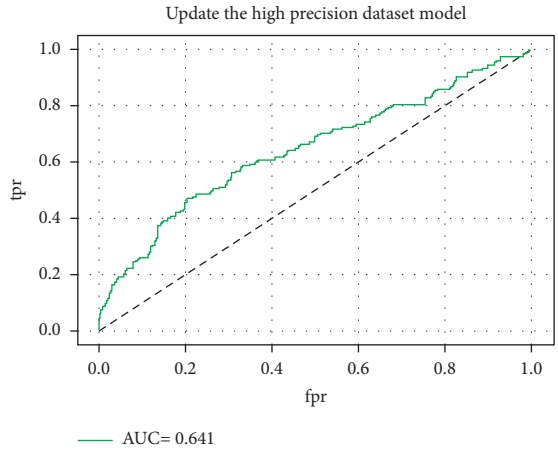


Figure 11: ROC curve and AUC value of the updated high precision data set model.

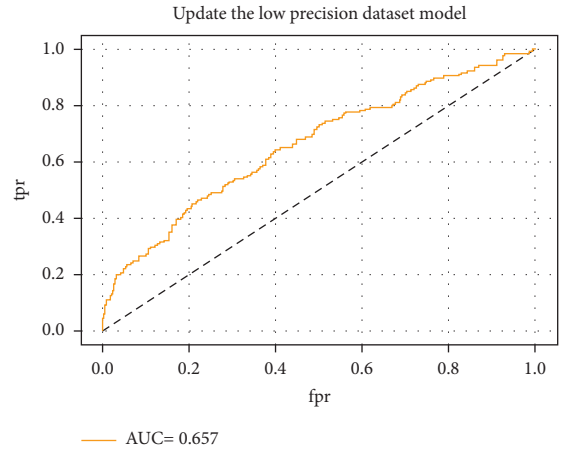


Figure 12: ROC curve and AUC value of the updated low precision data set model.

Table 11: The results of the updated low precision data set model.

Variable	Before		After	
	Mean	Confidence interval	Mean	Confidence interval
Constant	−1.477	(−1.650, −1.283)	−1.470	(−1.663, −1.290)
abs_dif_o	0.183	(0.012, 0.354)	0.189	(0.023, 0.352)
up_s	−0.300	(−0.503, −0.098)	−0.307	(−0.512, −0.110)
up_dif_o	0.252	(0.053, 0.451)	0.253	(0.052, 0.446)

Table 12: Confusion matrix of the updated low precision data set model.

	Positive (predict)	Negative (predict)
Positive (actual)	75	86
Negative (actual)	151	493

Table 13: SVM classification accuracy.

	Low precision data set (%)	High precision data set (%)
Accuracy	76.9	78.7

Table 14: Confusion matrix of the low precision data set.

	Positive (predict)	Negative (predict)
Positive (actual)	81	80
Negative (actual)	46	598

Table 15: Confusion matrix of the high precision data set.

	Positive (predict)	Negative (predict)
Positive (actual)	94	97
Negative (actual)	57	707

By analyzing these two groups of data sets with SVM model, this paper finds that when the loop density is relatively sparse, the precision of the model does have some influence, low accuracy of the data set to establish the SVM model accuracy is 76.9%, the AUC value is 0.8, high precision data set to establish accuracy of SVM is 78.7%, and the AUC value is 0.82. The comparison between the two models shows that the high precision data set is better for modeling. Compared to other machine learning models, Shen considered the weather variables when establishing the random forest real-time accident risk model [33]. In this model, the accuracy of the model reached 82.1%. In this study, the authors screened the characteristics of the data and took the weather into account. Compared with the support vector machine model, the accuracy was improved by 3.4%. In further researches, the data could be processed accordingly, and the parameters of the support vector machine could be adjusted to achieve higher prediction performance.

Compared with the Bayesian Logistic regression model, in the case of the same low precision data set, the overall prediction performance of the established SVM model is better, with the classification accuracy improved by 6.22%, and the AUC value is improved by 0.144. When loop density is small and the loop data information is not rich, a real-time crash risk model based on Bayesian Logistic regression can be established, which can effectively filter out the significant risk factors for crashes and can be explained in detail and quantify the corresponding risk factors. However, strict and

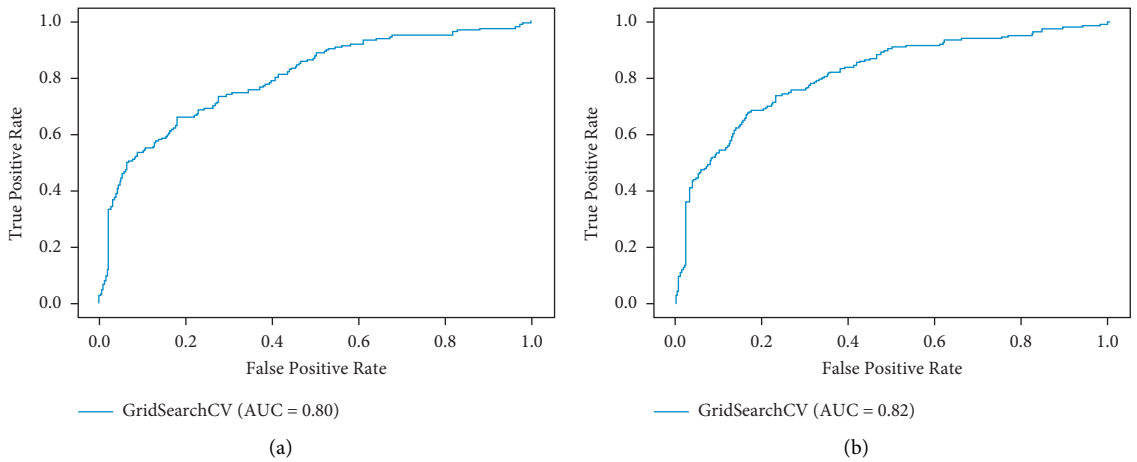


Figure 13: ROC curve and AUC value of low precision and high precision data sets.

mathematical relations limit the overall prediction effect of the model. SVM is a black box machine learning algorithm, which can effectively learn the effects of features on the results and reflect them into the prediction results.

Therefore, when the data set is not accurate enough, it is recommended to use the machine learning algorithm to establish a model to classify and predict the crash risk. When the data accuracy is good, the statistical Logistic regression method can be used to screen out significant risk variables to explain the model and classify the crash risk prediction.

5. Conclusions

Considering the influence of limited data conditions on the real-time freeway traffic crash risk model, this paper constructed high precision data set, low precision data set, and small sample data set. These data sets were modeled and analyzed based on Bayesian Logistic regression, and the reliability of real-time crash risk model transplantation based on Bayesian update was verified. Finally, the advantages and disadvantages of the model established by Bayesian Logistic and SVM were compared. The main conclusions of this paper are as follows:

(1) The significant risk factors of Bayesian Logistic regression established under various sample sizes are different. With the increasing of sample size, the evaluation index of the model decreases. However, the overall performance of the model improves. The increase of sample size can effectively improve the classification and prediction performance of the model.

(2) When the loop detector density of the collected data is small, the prediction performance of the Bayesian Logistic regression model based on low precision data set is weaker than that of the Bayesian Logistic regression model based on high precision data set. In addition, significant risk factors are significantly different in the two models, indicating that Bayesian Logistic regression is not suitable for low precision data set.

(3) Based on the Bayesian updating method, the validity of model migration is verified. Applying the posterior distribution of significant variable parameters of the Bayesian Logistic model based on high precision data set to low precision data set, this approach can improve the prediction performance of the Bayesian Logistic model using low precision data set.

(4) Compared with Bayesian Logistic regression, the crash risk model based on SVM has higher prediction performance. Even under the condition of low precision data set, its prediction performance is significantly improved compared with that of Bayesian Logistic regression, indicating that SVM is a better choice under the condition of insufficient data precision. However, SVM cannot effectively interpret the cause of crash risk. When the data quality is high, Bayesian Logistic regression can be used for modeling and prediction, and the crash risk can be well explained.

In this paper, Bayesian Logistic regression and support vector machine are applied to analyze the impact of various data sets on the traffic crash risk model. Further, other machine learning methods and the enhancement effect of feature engineering on the establishment for the crash risk model can be studied. Some new methods of crash risk model transplantation should also be studied in the future.

Data Availability

The data used to support the findings in this study are available from the corresponding authors upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Acknowledgments

This research was supported by the China Postdoctoral Science Foundation (2021M700333).

References

[1] W. Chu, S. Su, and L. Zhou, "Analysis and suggestions on traffic safety hazards of pure electric vehicles," *Road Traffic Management*, vol. 35, no. 4, pp. 32-33, 2020.

[2] W. Li, Z. Pu, Y. Li, and X. Ban, "Characterization of ride-splitting based on observed data: a case study of Chengdu, China," *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 330–353, 2019.

[3] Y. Yang, *Research on the Method of Freeway Crash Risk Identification and Comprehensive Traffic Safety Evaluation Considering the Regional Type difference*, Beijing Jiaotong University, Beijing, China, 2020.

[4] Z. Yuan, C. Lou, and Y. Yang, "Analysis of highway traffic accidents causes under time differences," *Journal of Beijing Jiaotong University*, vol. 45, no. 3, pp. 1–7, 2021.

[5] B. Tian, C. Liang, and Y. Bao, "Spatial and temporal distribution characteristics of traffic accidents on mountain highways and safety improvement measures," *Journal of Wuhan University of Technology*, vol. 42, no. 6, pp. 1014–1018, 2018.

[6] W. Y. Mergia, D. Eustace, D. Chimba, and M. Qumsiyeh, "Exploring factors contributing to injury severity at freeway merging and diverging locations in Ohio," *Accident Analysis and Prevention*, vol. 55, pp. 202–210, 2013.

[7] X. Ye, R. M. Pendyala, V. Shankar, and K. C. Konduri, "A simultaneous equations model of crash frequency by severity level for freeway sections," *Accident Analysis and Prevention*, vol. 57, pp. 140–149, 2013.

[8] N. Haghighi, X. C. Liu, G. Zhang, and R. J. Porter, "Impact of roadway geometric features on crash severity on rural two-lane highways," *Accident Analysis and Prevention*, vol. 111, pp. 34–42, 2018.

[9] M. Osman, S. Mishra, and R. Paleti, "Injury severity analysis of commercially-licensed drivers in single-vehicle crashes: accounting for unobserved heterogeneity and age group differences," *Accident Analysis and Prevention*, vol. 118, pp. 289–300, 2018.

[10] P. Xu, H. Huang, N. Dong, and S. C. Wong, "Revisiting crash spatial heterogeneity: a Bayesian spatially varying coefficients approach," *Accident Analysis and Prevention*, vol. 98, pp. 330–337, 2017.

[11] W. Wang, Z. Yuan, Y. Yang, X. Yang, and Y. Liu, "Factors influencing traffic accident frequencies on urban roads: a spatial panel time-fixed effects error model," *PLoS One*, vol. 14, no. 4, Article ID 0214539, 2019.

[12] Y. Yang, "Analysis of the factors influencing highway crash risk in different regional types based on improved Apriori algorithm," *Advances in Transportation Studies*, vol. 17, pp. 165–178, 2019.

[13] Z. Li, P. Liu, W. Wang, and C. Xu, "Using support vector machine models for crash injury severity analysis," *Accident Analysis and Prevention*, vol. 45, pp. 478–486, 2012.

[14] Y. Guo, Z. Li, and T. Sayed, "Analysis of crash rates at freeway diverge areas using Bayesian to bit modeling framework," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 4, pp. 652–662, 2019.

[15] X. Lv and Q. Wang, "Analysis of traffic accident severity on circular expressway based on logistic," *Journal of Wuhan University of Technology (Transportation Science and Engineering)*, vol. 63, pp. 1–8, 2021.

[16] C. H. Panagiotis, "Anastasopoulos. random parameters multivariate tobit and zero-inflated count data models: addressing unobserved and zero-state heterogeneity in accident injury-severity rate and frequency analysis," *Analytic Methods in Accident Research*, vol. 11, pp. 17–32, 2016.

[17] Q. Hou, X. Huo, J. Leng, and Y. Cheng, "Examination of driver injury severity in freeway single-vehicle crashes using a mixed logit model with heterogeneity-in-means," *Physica A: Statistical Mechanics and Its Applications*, vol. 531, Article ID 121760, 2019.

[18] Y. E. Fan and D. Lord, "Comparing three commonly used crash severity models on sample size requirements: multinomial logit, ordered probit and mixed logit models," *Analytic Methods in Accident Research*, vol. 1, pp. 72–85, 2014.

[19] Y. Yang, Z. Yuan, J. Chen, and M. Guo, "Assessment of osculating value method based on entropy weight to transportation energy conservation and emission reduction," *Environmental engineering and management journal*, vol. 16, no. 10, pp. 2413–2423, 2017.

[20] X. Chen, J. Ling, S. Wang, and Y. S. Yang, "Ship detection from coastal surveillance videos via an ensemble Canny-Gaussian-morphology framework," *Journal of Navigation*, vol. 74, no. 6, pp. 1252–1266, 2021.

[21] Z. Pu, Z. Li, Y. Jiang, and Y. H. Wang, "Full Bayesian before-after analysis of safety effects of variable speed limit system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 1–13, 2020.

[22] H. Yang, C. Liu, M. Zhu, X. Ban, and Y. Wang, "How fast you will drive? predicting speed of customized paths by deep neural network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–11, 2021.

[23] C. Lee, F. Saccomanno, and B. Hellinga, "Analysis of crash precursors on instrumented freeways," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1784, no. 1, pp. 1–8, 2002.

[24] M. Abdel-Aty, N. Uddin, A. Pande, M. F. Abdalla, and L. Hsia, "Predicting freeway crashes from loop detector data by matched case-control logistic regression," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1897, no. 1, pp. 88–95, 2004.

[25] H. Yang, R. Ke, Z. Cui, Y. Wang, and K. Murthy, "Toward a real-time smart parking data management and prediction (SPDMP) system by attributes representation learning," *International Journal of Intelligent Systems*, vol. 36, pp. 1–34, 2021.

[26] G. Fountas and P. C. Anastasopoulos, "A random thresholds random parameters hierarchical ordered probit analysis of highway accident injury-severities," *Analytic Methods in Accident Research*, vol. 15, pp. 1–16, 2017.

[27] J. Sun and J. Sun, "Proactive assessment of real-time traffic flow accident risk on urban expressway," *Journal of Tongji University*, vol. 42, no. 6, pp. 873–879, 2014.

[28] J. You, J. Wang, T. Tang, and S. Fang, "Support vector machines approach for predicting real-time rear-end crash risk on freeways," *Journal of Tongji University*, vol. 45, no. 3, pp. 355–361, 2017.

[29] C. Xu, P. Liu, W. Wang, and Z. B. LI, "Real time crash risk prediction model on freeways under nasty weather conditions," *Journal of Jilin University (Engineering and Technology Edition)*, vol. 43, no. 1, pp. 68–73, 2013.

[30] X. Chen, H. Chen, Y. Yang et al., "Traffic flow prediction by an ensemble framework with data denoising and deep learning model," *Physica A: Statistical Mechanics and Its Applications*, vol. 565, Article ID 125574, 2021.

[31] X. Ma, B. Fan, and S. Chen, "Evaluation and analysis model for freeways crash risk based on real-time traffic flow," *Journal*

of South China University of Technology, vol. 49, no. 8, pp. 19–25+34, 2021.

[32] C. Xu, W. Wang, P. Liu, R. Guo, and Z. Li, "Using the Bayesian updating approach to improve the spatial and temporal transferability of real-time crash risk prediction models," *Transportation Research Part C: Emerging Technologies*, vol. 38, pp. 167–176, 2014.

[33] J. Shen, *Real-time Risk Prediction and Spatiotemporal Impact Analysis for Freeway Accident*, Southeast University, Dhaka, Bangladesh, 2017.

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 2445693, 18 pages
https://doi.org/10.1155/2022/2445693

Research Article

An Empirical Study on the Segmentation of Potential Users of Shared Parking Spaces considering Individual Heterogeneity

Ange Wang[1], **Hongzhi Guan**[1], **Jun Guo**[1], **Yan Han**[2] and **Hangjin Bian**[1]

[1]*Faculty of Urban Construction, Beijing University of Technology, Beijing 100124, China*
[2]*Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing 100124, China*

Correspondence should be addressed to Yan Han; hanyan422@bjut.edu.cn

Received 21 October 2021; Revised 26 November 2021; Accepted 30 December 2021; Published 11 January 2022

Academic Editor: Ziyuan Pu

Shared parking has become the most effective way to utilize existing parking resources. Little attention has been focused on drivers' intention to use shared parking spaces in residential areas considering individual heterogeneity. To fill this gap, this paper explores the influencing factors and mechanism of shared parking use intention (SPUI) and further studies the preferences for the shared parking of different types of drivers. Firstly, based on the extended unified theory of acceptance and use of technology that includes psychological factors, personal attributes, and travel characteristics, the multiple indicator multiple cause (MIMIC) model was employed for parameter estimation and model assessment. Secondly, using MIMIC's output results as input variables, the segmentation method of the latent class model (LCM) was adopted to explore drivers' preferences regarding SPUI. Finally, a quantitative study was carried out through questionnaire data. The empirical results show that: (a) the extended unified theory of acceptance and use of technology has good explanatory power for SPUI. SPUI is directly affected by perceived risk (PR), behavioral habit (BH), social influence (SI), facilitating conditions (FCs), and effort expectancy (EE), while performance expectancy (PE) have no significant effect on SPUI. In addition, some factors of personal attributes and travel characteristics affect SPUI through psychological factors. (b) According to individual heterogeneity, the surveyed driver groups are divided into four segments: sensitive type (36%), conservative type (29.6%), neutral type (24.5%), and approved type (9.9%), respectively. There are significant differences in psychological observation variables such as EE, PE, FC, and SI among the four segments of drivers. According to the influence mechanism of psychological factors and preferences analysis of different types of drivers, the shared parking promotion strategy can be formulated from the aspects of management, operation, and technology.

1. Introduction

With the rapid growth of car ownership, the contradiction between the increasing demand for parking and the limited parking resources has become more and more prominent. Until July 2020, China's parking space gap has reached 80 million [1, 2], which increases search time for parking and creates exhaust emissions and traffic congestions [3–5]. In addition, in the context of COVID-19, travelers are more worried about the safety of public transportation, and the government has adopted a series of antiepidemic measures, such as restricting the passenger capacity of public transportation, which has led to a further decrease in the proportion of citizens taking public transportation and an increase in the proportion of private car trips. Therefore, the contradiction between the supply and demand of parking spaces is further aggravated.

In recent years, the development of the sharing economy has been in full swing and has penetrated into many areas, such as shared cars(Evcard, Gofun), shared accommodation(Airbnb), and ridesharing (Lyft). Parking space has the characteristics of being nonstorable in time and nonmovable in space; besides, the parking demands also vary depending on land properties, so parking demands and spaces are not well matched. Therefore, the concept of shared parking has been proposed; the basic idea of the concept is that the parking space owners sell parking permits for the idle period of their parking space on the network

platform, and drivers with parking demands can purchase parking permits through the network platform [6].

Residents' travel characteristics (often come out early and return late) make the parking spaces in the residential area idle during working hours, which makes the parking spaces in the residential area sharing during the working hours become possible. The number of parking spaces in residential areas accounts for a large proportion of the total number of parking spaces. According to the statistics of the Hong Kong Transport Department (2016), there are 485,000 parking spaces in Hong Kong (70% of the total number of parking spaces are designated for private use). In addition, Yan et al. [7] pointed out that from 2015 to 2019, residential parking resources in Beijing accounted for an average of 53.16% of all parking resources, and nearly 800,000 private parking spaces were idle during working hours. A simulation experiment shows that if 20% of Beijing's existing parking spaces are added to the sharing projects, carbon dioxide emissions can be reduced by 7.3 million tons per year [8]. Therefore, if the idle time of the huge number of parking spaces in residential areas can be fully utilized, the utilization rate of parking spaces can be effectively increased, and traffic congestion and carbon emissions can be further alleviated [9].

Shared parking projects are mainly concentrated in commercial areas in China, while shared parking projects in residential areas are very rare, and the participation is low; the fundamental reason is that drivers' SPUI is not strong [10]. Therefore, it is necessary to explore the drivers' SPUI. In addition, it is worth noting that because of the different sensitivities of drivers to the influencing factors, their performance is different and thus has different effects on the implementation of policy measures [11]. Generally, a segmentation study of travelers is helpful to estimate the differences in their behavior rules and characteristics, to classify travelers with taste heterogeneity, and then to formulate more targeted and effective policy measures [12]. Regarding taste heterogeneity, it has recently received considerable attention in the transportation field [13–15]. To the best of our knowledge, there is a lack of attention to revealing heterogeneity with respect to SPUI although some scholars have studied the SPUI from the perspective of demanders [10, 16], meaning that it would be certainly worthwhile to unveil how the impacts of the variables vary depending on individual heterogeneity as one of the first scientific efforts, adding knowledge to the literature.

This study aims to investigate the intention of drivers to use shared parking spaces in residential areas considering individual heterogeneity. Specifically, it firstly expands the unified theory of the acceptance and use of technology model and uses the extended unified theory of acceptance and use of technology as the theoretical framework. Then, a multiple indicator multiple cause (MIMIC) model is employed for parameter estimation and model assessment to characterize the intention of drivers toward the use of parking space sharing systems. Secondly, based on the results of the MIMIC model, the segmentation method of the latent class model (LCM) was adopted to explore drivers' preferences regarding SPUI, and the selection preferences of different types of drivers are analyzed according to psychological factors, demographic attributes, and travel characteristics, which will help to clarify the nature of SPUI at a deeper level and to formulate targeted inducement measures based on the driver's heterogeneous characteristics.

In addition, what needs to be elaborated is that exploring the sharing intention of parking space owners is also a critical area of study, since some problems with shared parking schemes, such as overtime, increased noise, and possible traffic congestion, need to be coordinated with parking space owners. However, this is out of the scope of this study, and readers are referred to the study proposed by Wang et al. [17–19], where they explored the influence of demographic attributes, built environment, and other factors of owners on their intention to participate in the shared parking schemes.

The rest of this paper is organized as follows: the next section extensively reviews literature concerning shared parking and research methods; in Section 3, the model framework, the multiple indicator multiple cause model and, the latent class modeling approach are briefly introduced; Section 4 profiles the dataset used in this study, including the survey outline, data collection, selected variables, and its descriptive analysis; in Section 5, the results and analysis are presented, including results of MIMIC and LCM estimation. At last, Section 6 concludes this research by summarizing implications and key findings, discussing limitations of the study, and presenting potential research topics.

2. Literature Review

2.1. Shared Parking. As far as we know, Lalani [20] proposed the concept of shared parking firstly, which also marks the transition from the nature of parking spaces to commodities. Shared parking has become an effective way to improve the utilization of existing parking resources. Litman [21] estimated that shared parking could reduce the need for new spaces by 10–30%. Yan et al. [7] combined prospect theory and logit model to identify the intention of suppliers to participate in shared parking considering uncertain demands. Besides, the matching mechanism of shared parking spaces is an important part of the research related to shared parking. Shao et al. [22] firstly developed a parking space matching model between the supply and demand for residential parking spaces, Iman and Hamid [23] and Kim et al. [24] built matching models based on GIS and other technologies. Xiao et al. [25] proposed a dual auction mechanism for shared parking spaces considering multiple periods. In addition, Zhao et al. [26] developed an intelligent parking management system (IPMS) to simulate the operation of the shared parking system. In addition to the participation behavior and the matching of shared parking, there are also studies based on its pricing. Some scholars have constructed pricing strategies for parking charges based on marginal cost theory and suboptimal pricing theory [27, 28]. Chen and Xie [29] established a dynamic allocation model of shared parking spaces for universities in the central city and conducted an effective evaluation. A competitive auction

mechanism also was presented based on berth allocation rules and transaction payment rules [30].

2.2. Unified Theory of Acceptance and Use of Technology. Plentiful theoretical models have been developed in behavioral research to analyze and explain the public's acceptance of a certain technology, including rational behavior theory, technology acceptance model, motivation model, planned behavior theory, technology acceptance model, and planned behavior theory [31]. In the study of the acceptance of new technologies, the most widely used models are the technology acceptance model and the unified theory of acceptance and use of technology.

In 2003, Venkatesh et al. [32] proposed a unified model called "the unified theory of acceptance and use of technology." He pointed out that the eight models mentioned above only explained 17%–53% of the differences in users' intention to use information technology, while the unified theory of acceptance and use of technology model performed better than them. Similarly, Chen et al.'s [33] research also supports this conclusion. Four core variables, i.e., performance expectancy (PE), effort expectancy (EE), facilitating conditions (FCs), and social influence (SI), and exogenous variables such as age, gender, income, and education are proposed in the unified theory of acceptance and use of technology.

Some scholars have explored the public's acceptance of various road transportation systems based on the unified theory of acceptance and use of technology [34–38]. Although the unified theory of acceptance and use of technology has a forceful universality in practice, it cannot fully explain the actual behavior in some scenarios [33]; that is, except the above factors, behavioral intention may also be affected by some other hidden factors that have not been discovered. To better explain the intention and behavior, some new variables have been incorporated in plentiful existing studies according to the characteristics of the research object.

2.3. Latent Class Model (LCM). In recent years, LCM has been applied to traffic segmentation. Teichert et al. [39] recognized the limitations of traditional segmentation techniques, adopted LCM to classify air passengers, and explored the importance of routes and flight segments to airlines' selection behavior; Crouch et al. [40] collected data on people's past vacation experience choices, travel motives, and basic demographic characteristics, established LCM to divide tourists into 5 classes, and explored how tourists choose vacation experiences/activity types. Xiong et al. [41] used the LCM to reveal the significant heterogeneity of Maryland drivers' potential preferences for carpooling, which supports traffic policies and incentive mechanisms related to congestion management strategies (such as HOV/HOT channel use). In addition, LCM is also widely used in other studies, such as long-distance drivers' route selection behavior [42], bicycle users [43], driving behavior on combined road segments [44], the time spent by tourists at destinations [45], the acceptance of self-driving vehicles [15], and the satisfaction with public transportation [13]. All of these studies have proved the advantages of potential category model segmentation research. Traveler segmentation studies that consider psychological factors have higher predictive power than individual attribute segmentation studies [46, 47].

2.4. Summary. Although the related research studies are fruitful, to the best of our knowledge, no previous study has explored the impact of drivers' heterogeneity on the SPUI. In addition, in the above studies on LCM, the explicit variables that determine the classification results are often entered into the LCM estimation without analysis, which will inevitably affect the accuracy of model estimation [48]. Therefore, this study innovatively uses MIMIC's output results as input variables of LCM to analyze the taste heterogeneity of different classes of drivers more accurately. This approach is expected to contribute to the literature by providing first-hand insights into SPUI of different categories of drivers.

3. Methods

3.1. Model Framework. In order to analyze the influence of psychological factors, personal attributes, and travel characteristics on SPUI, a MIMIC model is constructed to explore the causal relationship between various factors and SPUI. After that, based on the results of the MIMIC model, in order to identify the heterogeneity of preference among different categories of drivers, this study constructs LCM to explore potential driver classes and analyzes the characteristics of different classes. The model framework is shown in Figure 1.

Shared parking provides drivers with more parking options, but it also brings them safety and other concerns; therefore, the SPUI may be influenced by the perceived risks [18]. Besides, Xu [49] pointed out that when travelers have accumulated certain travel experiences in intercity travel choices, travelers tend to form the inertia of adopting a specific behavior under specific selection conditions, and such inertial behavior will enhance the effect of behavior habit (BH) on usage intention (UI). So, we assume that BH will also affect SPUI; therefore, we construct the extended unified theory of the acceptance and use of technology model by incorporating PR and BH variables into the original unified theory of acceptance and use of technology to identify the factors influencing SPUI. The structural equation between the psychological factors can reflect the internal influence relationship between the psychological factors, and the structural equation between the exogenous variables and the psychological factors can reflect the degree of influence of the exogenous variables on the psychological factors [50].

Based on the analysis of the relationship between the influencing factors and SPUI, the difference between the driver's choice intention and preference is further explored. LCM is used to subdivide the driver category, as shown in the LCM framework in Figure 1. In traveler segmentation

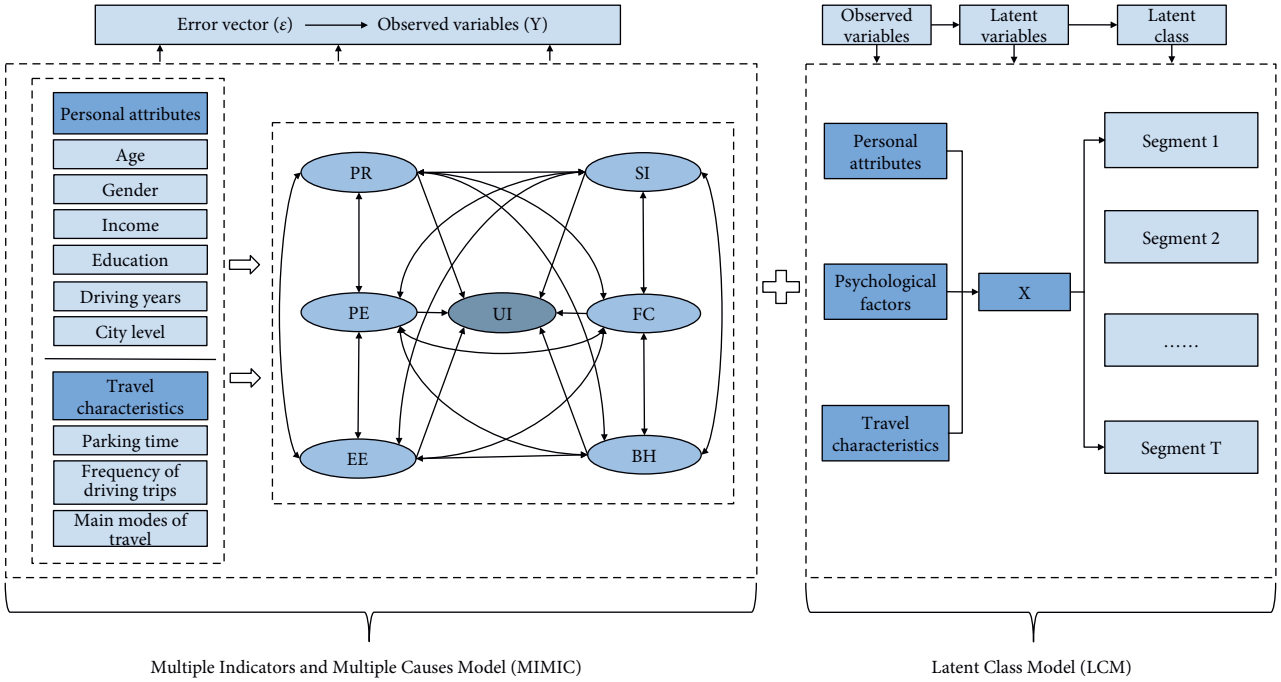


Figure 1: Modeling framework.

studies, demographic factors are usually regarded as basic classification indicators, such as age, gender, occupation, and income [51]. In addition, tourism-related factors [11] and various factor combinations [39] are also used for segmentation research. The driver's choice behavior is closely related to travel characteristics, such as driving frequency and parking time. Therefore, this article comprehensively considers the combination of personal attributes, travel characteristics, and psychological factors for subdivided research. The solution result of the MIMIC model is the influence relationship of various factors on SPUI. Variables that have a significant influence on SPUI are taken as input variables of LCM to obtain the number of types of driver groups and the specific characteristics of each type, such as the distribution and comparison of input variables. Finally, the characteristics of each category are summarized and analyzed.

3.2. Multiple Indicator and Multiple Causes (MIMIC) Model. Compared with the traditional SEM model and other models, the MIMIC model can not only express the exogenous causes and endogenous indicators of latent variables clearly through a strict structural model but also obtain the influence degree of all explanatory variables on latent variables, and this analysis method can make normal statistical relationship test more convenient [52]. Schneider and Enste [53] proposed that the MIMIC model can deal with multiple latent variables and endogenous indicators without strict constraints and assumptions, and allow exogenous causes and endogenous indicators to contain measurement errors. Therefore, its theoretical framework is more flexible than other indirect measurement methods, potentially including all other indirect measurement methods.

Therefore, this study intends to use the MIMIC model to explore the influence of psychological factors and exogenous variables on the SPUI.

The matrix form of MIMIC is shown in the equations as follows:

$$\eta = \Gamma x + \zeta. \tag{1}$$

$$y = \Lambda \eta + \varepsilon, \tag{2}$$

where η is the psychological factor vector; x indicates the exogenous variable vector; y is the observation variable vector; Γ and Λ are the parameter matrixes to be estimated; ζ and ε represent the measurement errors.

We substitute equation (1) into equation (2) to obtain

$$\begin{aligned} y &= \Lambda(\Gamma x + \zeta) + \varepsilon = \Pi x + \nu, \\ \Pi &= \Lambda \Gamma x, \\ \nu &= \Lambda \zeta + \varepsilon. \end{aligned} \tag{3}$$

Assuming that the measurement errors are independent of each other and subject to normal distribution, then $E(\zeta \varepsilon') = 0$, $E(\zeta^2) = \delta^2$, $E(\zeta \varepsilon \varepsilon') = \Theta^2$. Θ^2 is the lower triangular matrix of ν, so the covariance matrix formula is

$$\begin{aligned} \sum(\theta) &= E(\nu \nu') \\ &= \delta^2 \Lambda \Lambda' + \Theta^2. \end{aligned} \tag{4}$$

The total covariance matrix can be calculated by observing the sample values of the variables, so $\sum = \sum(\theta)$; each parameter can be solved.

In the multiindex part of the MIMIC model, it is equivalent to the confirmatory factor analysis of psychological factors. The multicause part can be expressed as

$$\eta_{li} = \gamma_{l1} X_{\text{gender}_i} + \gamma_{l2} X_{\text{age}_i} + \gamma_{l3} X_{\text{inc}_i} + \gamma_{l4} X_{\text{edu}_i} + \gamma_{l5} X_{\text{year}_i} + \gamma_{l6} X_{\text{city}_i} + \gamma_{l7} X_{\text{awareness}_i} + \gamma_{l8} X_{\text{time}_i} + \gamma_{l9} X_{\text{frequency}_i}, \quad (5)$$

where l is the psychological factor, including EE, SI, FC, PE, PR, BH, and UI; γ indicates the parameter to be estimated; i represents the observed individual.

3.3. Latent Class Model (LCM). LCM is a modeling analysis technique to explore latent variables, which is not only a statistical method but also a statistical methodology [54]. The most important feature of the latent category model is that it can process classification data and use the maximum likelihood method to estimate parameters, which can calculate the appropriate number of categories more accurately. In addition, the classification results are relatively stable and can be generalized in other samples of the same population, which is convenient for practical application [55, 56]. It is an important research method to subdivide travelers' travel behavior differences with distinctive characteristics by using the potential category model [57]. It can overcome the shortcoming of traditional cluster analysis that deterministically assign travelers to a single group while ignoring the possibility of misclassification to the wrong group [58], such as K-means++ and DBSCAN. The LCM analysis process consists of three steps: probability parameterization, model fitting and parameter estimation, and latent classification and result interpretation [59].

(1) Probability parameterization. It is the first step that converts the probability of classifying variables into a parameter model [11]. In this paper, personal attributes, psychological factors, and travel characteristics are taken as three explicit variables A, B, and C, whose conditional probabilities are referred to $\pi_{it}^{\overline{A}x}$, $\pi_{jt}^{\overline{B}x}$, and $\pi_{kt}^{\overline{C}x}$. The sum of conditional probability of all observed variables at each level is 1, and the formula is as follows:

$$\sum_i \pi_{it}^{\overline{A}X} = \sum_j \pi_{jt}^{\overline{B}X} = \sum_k \pi_{kt}^{\overline{C}X} = 1. \quad (6)$$

Assuming that there is a latent variable X with t ($t = 1, 2, 3, \ldots, T$) latent categories can explain the relationship between the three latent variables A, B, and C, the LCM can be expressed as follows [60]:

$$\pi_{ijk}^{ABC} = \sum_{t=1}^{T} \left(\pi_t^X \times \pi_{it}^{\overline{A}X} \times \pi_{jt}^{\overline{B}X} \times \pi_{kt}^{\overline{C}X} \right). \quad (7)$$

(2) Model fitting and parameter estimation. Maximum Likelihood (ML) is mainly used for parameter estimation in potential category models, and the commonly used iterative algorithms include Expectation Maximization (EM), and Newton Rapson (NR). In this paper, the expectation maximization algorithm is used for iterative analysis. The ML estimation function is as follows:

$$\pi_{ijkt}^{ABCX} = \pi_t^X \times \pi_{it}^{\overline{A}X} \times \pi_{jt}^{\overline{B}X} \times \pi_{kt}^{\overline{C}X}. \quad (8)$$

(3) Latent classification and result interpretation. By diving π_{ijkt}^{ABCX} and π_{ijk}^{ABC}, the ML probability of observed variables in different potential categories at each level can be obtained, and the formula is as follows:

$$\pi_{ijkt}^{ABC\overline{X}} = \frac{\pi_{ijkt}^{ABCX}}{\pi_{ijk}^{ABC}}. \quad (9)$$

The posterior probability of each sample belonging to different latent classes is calculated, and the one with the highest value is the class to which the observed value belongs. The formula is as follows:

$$\pi_{tijk}^{\overline{X}ABC} = \frac{\pi_{ijkt}^{ABCX}}{\sum_{t=1}^{T} \pi_{ijkt}^{ABCX}}. \quad (10)$$

4. Data

4.1. Sample. To ensure the data quality of the questionnaire, we conducted a presurvey in which we asked people with different levels of knowledge about shared parking to fill in the questionnaire and adjusted the questionnaire according to the presurvey results.

The final questionnaire consists of three parts: the first part is the information about parking on the latest trip; the second part is scene simulation; the third part covers the measurement items of the extended unified theory of acceptance and use of technology model variables, including EE, SI, FC, PR, PE, BH, and UI; the fourth part includes the respondents' socioeconomic attributes, parking characteristics, and travel characteristics. This study chiefly used the third and fourth parts of the questionnaire.

This study uses the Likert five-level scale — from "Strongly Disagree (=1)" to "Strongly Agree (=5)" to measure all psychological factors. It is used to rank respondents' recognition of the relevant descriptions given in each questionnaire measurement item for reflecting drivers' psychological feelings about shared parking. The measurement items constituting the questionnaire were adapted from scales validated in previous literature studies and modified by experts to reflect the specific environment and target population of shared parking. The six psychological factors of the theoretical model contain a total of 31 measurement items. Table 1 presents items related to psychological factors.

We carried out the formal survey using an online questionnaire in February 2021. There is no limit to the location distribution of respondents, and individuals with links to the survey and interest in the study are invited to participate. Individuals were considered eligible for inclusion if they met the following criteria: they were Chinese, had a valid driver's license, and had recently parked in an urban area for longer than 30 minutes. After conditional screening, the number of valid questionnaires in this study is

Table 1: Construction and measurement items of psychological factors.

Psychological factor	Project	Measurement item	Source
Facilitating Conditions (FCs)	FC1	If the operation of shared parking app is easy to understand, I prefer to choose shared parking space in residential areas	[18, 61]
	FC2	If the transaction of shared parking app is safe and convenient, I prefer to choose shared parking space in residential areas	
	FC3	If relevant institutions establish shared parking credit mechanism, I prefer to use shared parking space in residential areas	
	FC4	Shared parking will certainly get policy support	
	FC5	Existing technology is sufficient to support the normal operation of shared parking	
	FC6	I am willing to increase the walking distance to pay less parking fees	
Social Influence (SI)	SI1	If all my friends and relatives are using the shared parking space in the residential area, I will also use it	[61, 62]
	SI2	The encouragement from relatives and friends around me will make me more willing to choose the shared parking space in the residential area	
	SI3	The encouragement and appeal from the government and media will make me more willing to choose shared parking spaces in residential areas	
	SI4	Choosing to use a shared parking space in a residential area makes me feel technologically advanced	
Effort Expectancy (EE)	EE1	Shared parking spaces in residential areas are easy to find	[32]]
	EE2	I can quickly get used to the shared parking system in my residential area	
	EE3	The shared parking system would be simple to use	
	EE4	I feel completely in control of using the residential parking share system	
Perceived Risk (PR)	PR1	It is difficult to find the reserved shared parking space in the residential area	[18, 63]
	PR2	Choosing to use shared parking spaces in residential areas has a high risk of exposing personal privacy (such as vehicle information and location information)	
	PR3	If you park over time (beyond the scheduled time), it will cause more inconvenience for others to use the parking space	
	PR4	At present, the related equipment and policy mechanism of shared parking are not perfect enough	
	PR5	Parking in a shared parking space in a residential area increases vehicle safety risks	
Performance Expectancy (PE)	PE1	Using a shared parking space in a residential area will reduce my parking expenses	[64–66]
	PE2	Using shared parking spaces in residential areas can reduce the time spent looking for a parking space	
	PE3	Using shared parking spaces in residential areas can improve the efficiency and quality of my study, work, and life	
	PE4	Shared parking is good for the sustainable development of the city	
	PE5	Shared parking can alleviate the problem of "hard parking"	
Behavior Habit (BH)	BH1	Next time when I park my car, I will continue to use the public parking lot	[49]
	BH2	Public parking is my usual choice	
	BH3	I usually do not actively consider choosing new parking options (such as shared parking)	
Usage Intention (UI)	UI1	I would like to try to use shared parking space in residential areas in the future	[33, 38]
	UI2	I would like to give priority to using shared parking spaces in residential areas in the future	
	UI3	I will often use the shared parking space in the residential area	
	UI4	I will strongly recommend my relatives and friends to use shared parking spaces in residential areas in the future	

217. The sample size should be at least 5–10 times the variable in multivariate research using structural equation [50], which means the sample size of formal investigation should be more than 155, so the sample size of formal investigation can meet the research needs.

In an effective sample, the drivers surveyed were evenly distributed in cities of different levels, accounting for about 20% in first-tier, second-tier, and third-tier cities, and 35% in fourth-tier cities. The number of male drivers (63.13%) was higher than that of female drivers (36.87%). The respondents were mainly aged from 18 to 30 years old (61.29%), which accords with the age distribution structure of Chinese drivers, and those older than 30 years old were evenly distributed. 42.86% of the respondents had a college or bachelor's degree, and more than 30% had a master's degree or above. More than half of the respondents earned less than 5,000 yuan, which is consistent with the income characteristics of Chinese residents. Table 2 summarizes the main demographic characteristics of the valid sample.

4.2. Data Analysis

4.2.1. Measurement Model Testing. Firstly, exploratory factor analysis was carried out using SPSS, and KMO and Bartlett's sphere tests were performed on the scale. The

Table 2: Demographic characteristics of the participants.

Characteristics	Item	Proportion (%)	Variable naming
Gender	Male	63.13	Gender
	Female	36.87	
Age	18–25	40.09	Old
	26–30	21.2	
	31–40	14.75	
	41–50	11.06	
	51–60	9.68	
	>60	0.46	
Driving experiences	≤1	33.64	Year
	1–3	26.73	
	≥3	39.63	
Education level	High school or below	25.81	Education
	College/undergraduate	42.86	
	Master	26.73	
	Doctor	4.61	
Monthly income	<3000 yuan	31.34	Income
	3000–5000 yuan	23.04	
	5000–7000 yuan	18.43	
	7000–9000 yuan	11.52	
	9000–11000 yuan	3.69	
	>11000 yuan	11.98	
City level	First-tier city	22.58	City
	Second-tier city	22.58	
	Third-tier city	19.82	
	Fourth-tier city	35.02	
Understanding degree of shared parking	Do not understand at all	7.83	Awareness
	Do not understand	24.42	
	General understanding	52.53	
	Better understand	11.98	
	Understand very well	3.23	
Average daytime parking time on weekdays	<2 h	22.12	Time
	2–5 h	23.5	
	5–8 h	30.41	
	>8 h	23.96	
Driving frequency	1–4 times per week	58.53	Frequency
	5–8 times per week	23.96	
	9–12 times per week	9.68	
	>12 times per week	7.83	

results showed that the KMO value was 0.902 (>0.700), and Bartlett's sphericity test value was significant (sig.<0.001), indicating that the questionnaire data met the prerequisite requirements of factor analysis.

Cronbach's α, composite reliability, and average variance extracted were used to evaluate the reliability and validity of the questionnaire. Table 3 shows the reliability and validity of test results. If Cronbach's α is higher than 0.8, the reliability is high; if it is between 0.7 and 0.8, the reliability is acceptable; if it is between 0.6 and 0.7, it is basically acceptable; if it is less than 0.6, the reliability is not good, and the survey scale should be revised [67]. Based on analysis results, only Cronbach's α of BH is 0.784, and Cronbach's α of other psychological factors are all higher than 0.8, indicating that the reliability of the questionnaire is very high.

When the average variance extracted (AVE) value is greater than 0.5, it indicates that the psychological factor has good convergence validity [68]. In addition, when the outer loading coefficient corresponding to each observation variable is greater than 0.6 and $P < 0.05$, it also indicates that the convergence validity is up to standard. Just as the results in Table 4, the minimum AVE value of the psychological factor is 0.567, and the outer loading coefficient of each measurement item is above 0.66, indicating that the data have strong reliability and internal consistency.

4.2.2. Structural Model Testing. The chi-square degree of freedom ratio (χ^2/df), Root Mean Squared Error of Approximation (RMSEA), Goodness-of-Fit Index (GFI), Comparative Fit Index (CFI), and Adjust Goodness-of-Fit Index (AGFI) were selected as evaluation indexes of model fit degree. The MIMIC model was employed for parameter estimation and model assessment to characterize the intentions of drivers toward the use of parking

Table 3: Measurement items for all the relevant constructions.

Latent variable	Item no.	Outer loading	Mean	SD	α	AVE
Facilitating Conditions (FCs)	FC1	0.918	3.83	0.793	0.921	0.683
	FC2	0.943	3.80	0.811		
	FC3	0.916	3.82	0.787		
	FC4	0.735	3.68	0.865		
	FC5	0.741	3.47	0.901		
	FC6	0.661	3.76	0.856		
Social Influence (SI)	SI1	0.923	3.65	0.826	0.928	0.765
	SI2	0.895	3.60	0.861		
	SI3	0.900	3.73	0.808		
	SI4	0.775	3.53	0.892		
Effort Expectancy (EE)	EE1	0.693	3.39	3.39	0.876	0.663
	EE2	0.854	3.56	0.873		
	EE3	0.849	3.50	0.918		
	EE4	0.849	3.39	0.886		
Perceived Risk (PR)	PR1	0.843	3.49	0.849	0.915	0.690
	PR2	0.748	3.40	0.904		
	PR3	0.914	3.73	0.889		
	PR4	0.856	3.71	0.842		
	PR5	0.779	3.49	0.942		
Performance Expectancy (PE)	PE1	0.699	3.44	0.925	0.914	0.684
	PE2	0.824	3.60	0.797		
	PE3	0.706	3.47	0.926		
	PE4	0.954	3.79	0.779		
	PE5	0.920	3.76	0.799		
Behavior Habit (BH)	BH1	0.763	3.62	0.822	0.784	0.567
	BH2	0.776	3.77	0.775		
	BH3	0.718	3.42	0.881		
Usage Intention (UI)	UI1	0.866	3.71	0.774	0.940	0.815
	UI2	0.922	3.45	0.793		
	UI3	0.938	3.43	0.800		
	UI4	0.885	3.39	0.815		

Table 4: Test results of goodness of fit.

Index	χ^2/df	RMSEA	GFI	CFI	AGFI
Actual value	1.537	0.064	0.801	0.950	0.737
Standard value	<3	<0.3	>0.8	>0.9	>0.8

space sharing systems based on AMOS software. Then, we found that the theoretical model and empirical data cannot be entirely fitting. Under the premise of not affecting the completeness of the theoretical model, the model was revised, and the path that had no significant impact was deleted. The calculation results of the test index of the final model fit are shown in Table 4. Except that the AGFI is slightly less than the standard value, other models are proposed. All the compliance indicators meet the requirements.

5. Result Analysis

5.1. Analysis of MIMIC Estimation Results. From the calibration results of the MIMIC model, we can see the relationship between personal attributes, travel characteristics and psychological factors, as well as the influence relationship between the psychological factors.

5.1.1. Analysis of the Influence Relationship between Psychological factors. The path relationship between the psychological factors in the shared parking use intention model is shown in Figure 2. The values on the path in the figure are standardized coefficients. In the figure, * means $P < 0.05$, ** means $P < 0.01$, and ***means $P < 0.001$.

The impact of each psychological factor on UI is ranked from the highest to lowest as follows: FC ($\beta = 0.677$, t-value $= 4.025$, $P < 0.001$), SI ($\beta = 0.452$, t value $= 1.958$, $P < 0.05$), EE ($\beta = 0.443$, t value $= 3.306$, $P < 0.001$), BH ($\beta = 0.229$, t value $= 2.052$, $P < 0.05$), PR ($\beta = -0.214$, t value $= -0.232$, $P < 0.05$), except PE ($\beta = 0.170$, t value $= 1.100$, $P = 0.271 > 0.05$), other psychological factors all have a significant impact on UI, with a confidence interval of 95% (corresponding to $P < 0.05$), the proportion of variance in UI that can be explained by FC, SI, EE, PR, BH is 59.6% ($R^2 = 0.596$). Marcoulides [69] pointed out the following evaluation criteria: R^2 lower

than 0.19 indicates unacceptable explanatory power; 0.19–0.33 indicates weak explanatory power; 0.33–0.67 indicates moderate explanatory power; higher than 0.67 indicates good explanatory power. Besides, the evaluation standard proposed by Cohen et al. [70] pointed out that when R^2 is greater than 0.4, the model is considered to have good explanatory power. According to the evaluation criteria proposed by Marcoulides and Cohen, in addition, the results of Chen's [33] study ($R^2 = 0.48$) also illustrate the validity of our result. So, it can be inferred that the theoretical model we constructed has a good explanatory power for the use intention of shared parking.

5.1.2. Analysis of the Influence Relationship of Exogenous Variables on Psychological factors. A summary of the impact of exogenous factors including population socioeconomic attributes, parking characteristics, and travel characteristics on FC, SI, PE, BH, EE, and PR is shown in Table 5.

Exogenous factors do not directly affect UI but indirectly affect UI by affecting FC, SI, PE, BH, EE, and PR. It can be seen from Table 5 that those exogenous variables do not have a significant impact on every psychological factor. Gender and the city level have no significant impact on all psychological factors, indicating that gender and the city level of the respondent do not affect their SPUI.

Age has a significant positive effect on EE ($P < 0.05$), indicating that within a certain range, older people have to work harder to accept and adapt to the shared parking system. On the contrary, young people have strong confidence in the control of new things, which may be related to the limited network information and physiological factors accepted by the elderly, while modern young people receive more extensive information from the outside world and have more active thinking. Driving experiences have a significant positive impact on PR ($P < 0.05$), indicating that the older the driving experiences, the more sensitive they are to the potential risks of shared parking; income has a significant negative impact on both FC ($P < 0.01$) and SI ($P < 0.05$); this shows that the higher the income, the less sensitive the attitudes of the external conditions and the surrounding people will be and will not be easily influenced by the outside world, this may be because people with higher incomes will look at new things more rationally; parking time has a significant positive effect on PE ($P < 0.001$), FC ($P < 0.001$), BH ($P < 0.05$), SI ($P < 0.01$), and EE ($P < 0.01$); driving frequency has a significant negative impact on PR ($P < 0.05$), PE ($P < 0.05$), FC ($P < 0.05$), and SI ($P < 0.05$), indicating that the higher frequency of driving.

5.2. Analysis of LCM Segmentation Results. In this study, based on the results of the MIMIC model, 11 factors were selected initially to subdivide the driver category. These factors are age, income, driving experiences, education level, parking time, frequency of driving, the degree of confidence in the control of shared parking systems (EE4), the influence of walking distance (PC6), the influence of government and news media (SI3), the attractiveness of preferential fees (PE1), and the influence of parking choice habit (BH3). After calculation, the P value of BH in LCM is greater than 0.05, so the null hypothesis cannot be rejected. This variable cannot be effectively explained by the potential category, so BH is excluded from the explicit variables.

Because the number of classes C is initially unknown, we need to specify the value of C beforehand. According to literature [71], CAIC (Consistent Akaike Information Criterion) and BIC (Bayesian Information Criterion) indicators of LCM are used to determine category C. The smaller the CAIC and BIC values are, the higher the model fit degree is. Therefore, it is assumed that the samples can be divided into 1–10 classes. Then, LatentGOLD 5.0 software is used for LCM fitting and parameter estimation to obtain the CAIC and BIC values under a different number of classes. It can be seen from Figure 3 that when the number of classes is 4, CAIC and BIC have the smallest values, so the sample can be divided into four classes.

Respondents are modeled based on potential category analysis, and model parameters are estimated. The results are shown in Table 6. From the results, the probabilities of the four potential classes are 0.3602, 0.2964, 0.2449, and 0.0988, respectively. Finally, the classification accuracy of the model is $\lambda = 0.95$, which shows that the classification accuracy is high, and the classification results are reasonable [72].

The different observation factors of the latent class are analyzed, as shown in Figures 4–13. It is can be seen that the five observed psychological factors (EE4, SI3, FC6, and PE1) are also different in the latent classes, as shown in Figures 4–7, which shows their different psychological characteristics. There are significant differences among the observed factors age, driving experience, education level, income, parking time, and travel frequency in the latent classes, as shown in Figures 8–13.

According to the results of Figures 4–13 and Table 6, we characterized the segments as follows:

Segment 1 (36% of the respondents) is mainly composed of people with relatively more positive psychological characteristics. This type of driver makes up the largest proportion. They are more likely to accept "long distance in exchange for less parking fees" than segments 2 and 3, and they are more optimistic about the potential economic benefits of shared parking, but they are more susceptible to the influence of the government and the media than segments 2 and 3, so we call them "sensitive type." The age and income distribution of the members of this class are relatively even. 61.1% of the respondents have more than three years of driving experience, more than 98% of the respondents have a bachelor's degree or above, and more than 40% of respondents parked for more than 8h during working hours.

Segment 2 (29.6% of the respondents) is mainly composed of people with relatively more conservative attitudes, which we call the "conservative type." Their degree of influence by the government and media is the least among the four segments, and their degree of confidence in the control of the shared parking system is basically the same as that of segment 3 but lower than that of segments 1 and 4; besides, they have no great expectation of reduced parking fees compared with other segments. The age distribution of this

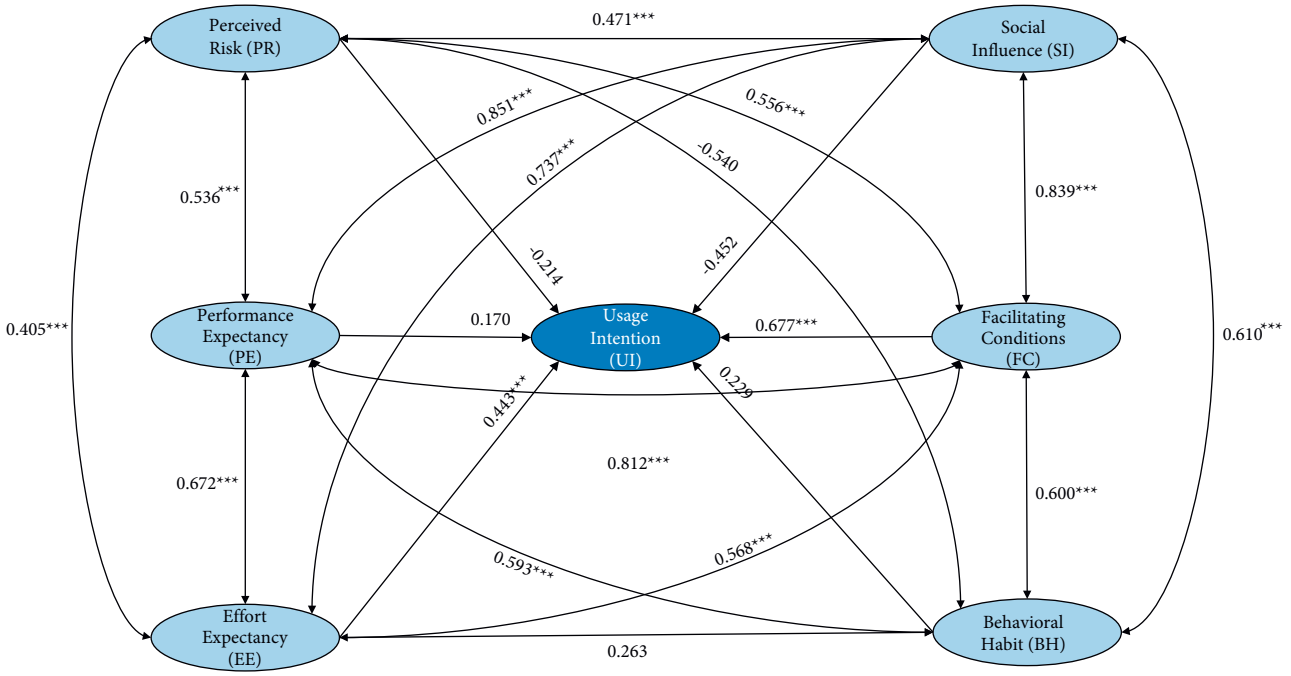


Figure 2: Analysis results of the structural model. Solid line indicates that there is a significant influence relationship; dashed line indicates that there is no significant influence relationship.

Table 5: Results of ANOVA.

Variable	Frequency	Parking time	City level	Income	Education	Driving experiences	Age	Male
PR	**−0.614***	0.395	−0.133	−0.247	0.067	**0.494***	−0.22	0.067
PE	**−0.287***	**0.844*****	−0.069	−0.048	0.221	0.000	0.131	0.148
FC	**−0.279***	**0.734*****	−0.206	**−0.384****	0.222	0.037	0.212	0.155
BH	−0.376	**0.539***	−0.010	−0.384	0.112	0.198	0.000	0.296
SI	**−0.462***	**0.701****	−0.246	**−0.358***	0.045	0.162	0.236	0.017
EE	−0.064	**0.578****	−0.246	0.088	**0.449***	0.156	**0.540***	0.191

Note. If the P value is greater than 0.05, the moderating effect is not significant; if the P value is lower than 0.05, the moderating effect is significant. $^{***}P < 0.001$, $^{**}P < 0.01$, and $^{*}P < 0.05$.

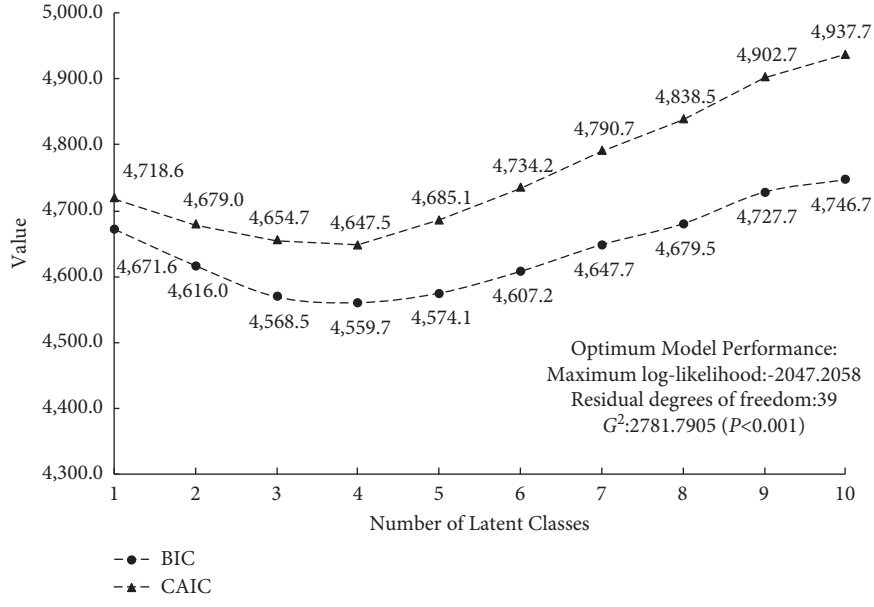


Figure 3: Optimum number of classes in the latent class model.

Table 6: Parameter estimation results of LCM (four latent classes: C1, C2, C3, and C4).

Variable	Observed Variable	Level	Description	Conditional Probability of Latent Class			
				C1	C2	C3	C4
Psychological factors	EE4	1	Strongly disagree	0.0178	0.0431	0.0436	0.0000
		2	Disagree	0.0716	0.1176	0.1182	0.0000
		3	Generally agree	0.4425	0.4913	0.4916	0.0013
		4	Agree	0.4538	0.3407	0.3395	0.1842
		5	Strongly agree	0.0144	0.0073	0.0072	0.8145
		Mean	3.3753	3.1514	3.1487	4.8132	
	SI3	1	Strongly disagree	0.0071	0.0393	0.0334	0.0000
		2	Disagree	0.0080	0.0232	0.0210	0.0000
		3	Generally agree	0.2697	0.4055	0.3929	0.0000
		4	Agree	0.6377	0.5003	0.5176	0.0291
		5	Strongly agree	0.0775	0.0317	0.0350	0.9709
		Mean	3.7704	3.4619	3.4999	4.9708	
	FC6	1	Strongly disagree	0.0074	0.0436	0.0278	0.0000
		2	Disagree	0.0165	0.0495	0.0378	0.0000
		3	Generally agree	0.2713	0.4134	0.3792	0.0043
		4	Agree	0.5673	0.4393	0.4838	0.1746
		5	Strongly agree	0.1375	0.0541	0.0715	0.8211
		Mean	3.8110	3.4108	3.5335	4.8168	
	PE1	1	Strongly disagree	0.0332	0.0758	0.0423	0.0000
		2	Disagree	0.0604	0.0959	0.0694	0.0000
		3	Generally agree	0.4032	0.4469	0.4189	0.0048
		4	Agree	0.4714	0.3643	0.4423	0.2479
		5	Strongly agree	0.0319	0.0172	0.0270	0.7474
		Mean	3.4083	3.1512	3.3423	4.7426	
Personal attributes	Age	1	18–25	0.0038	0.0001	0.0544	0.0020
		2	26–30	0.2752	0.0300	0.7637	0.1862
		3	31–40	0.2916	0.0883	0.1547	0.2528
		4	41–50	0.2408	0.2024	0.0244	0.2675
		5	51–60	0.1309	0.3056	0.0025	0.1864
		6	>60	0.0576	0.3735	0.0002	0.1051
		Mean	3.3929	4.9038	2.1576	3.7654	
	Driving experience	1	≤1	0.1490	0.2572	0.6764	0.1818
		2	1–3	0.2403	0.2760	0.2155	0.2550
		3	≥3	0.6107	0.4668	0.1081	0.5632
		Mean	2.4617	2.2096	1.4317	2.3813	
	Education level	1	High school or below	0.0166	0.7945	0.0178	0.5355
		2	College/Undergraduate	0.4631	0.2038	0.4749	0.4517
		3	Master	0.4319	0.0017	0.4241	0.0127
		4	Doctor	0.0884	0.0000	0.0832	0.0001
		Mean	2.5922	1.2073	2.5727	1.4774	
	Income	1	<3000 yuan	0.0769	0.3299	0.6277	0.2719
		2	3000–5000 yuan	0.1602	0.3281	0.2777	0.3124
		3	5000–7000 yuan	0.2118	0.2069	0.0779	0.2276
		4	7000–9000 yuan	0.1814	0.0846	0.0142	0.1074
		5	9000–11000 yuan	0.0968	0.0215	0.0016	0.0316
		6	>11000 yuan	0.2729	0.0289	0.0010	0.0491
			Mean	3.8797	2.2265	1.4871	2.4617
Travel characteristics	Parking time	1	<2 h	0.0752	0.2609	0.3218	0.1858
		2	2–5 h	0.1258	0.2199	0.2324	0.1945
		3	5–8 h	0.3764	0.3316	0.3003	0.3640
		4	>8h	0.4226	0.1876	0.1456	0.2557
		Mean	3.1463	2.4460	2.2696	2.6896	
	Frequency of driving	1	1–4 times per week	0.6159	0.3487	0.9920	0.5204
		2	5–8 times per week	0.2888	0.3405	0.0080	0.3238
		3	9–12 times per week	0.0665	0.1633	0.0000	0.0989
		4	>12 times per week	0.0288	0.1475	0.0000	0.0569
			Mean	1.5082	2.1096	1.0080	1.6924
Probability of latent class				0.3602	0.2961	0.2449	0.0988

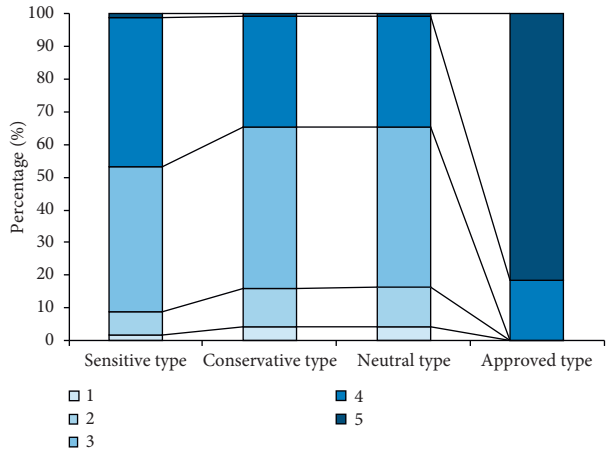


Figure 4: EE of drivers in the four segments.

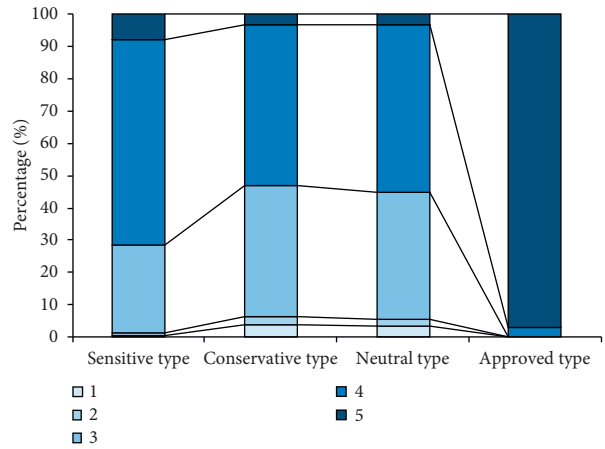


Figure 5: SI of drivers in the four segments.

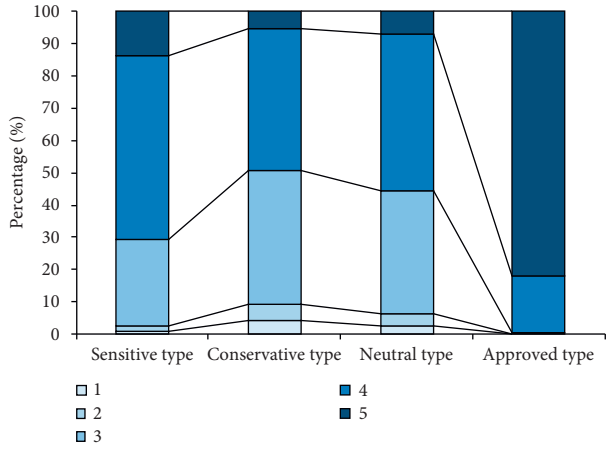


Figure 6: FC of drivers in the four segments.

segment is relatively uniform, and the average education level is the lowest, while the average driving frequency is the highest; among the four segments, their monthly income is concentrated below 7000 yuan (86.5%), but the parking time is lower than segments 1 and 4. Therefore, it can be inferred that the randomness of their trips may be strong.

Segment 3 (24.5% of the respondents) has the least confidence in the control of new things, and their response

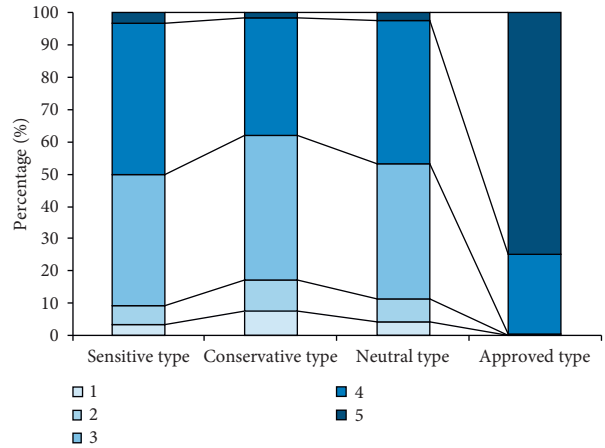


Figure 7: PE of drivers in the four segments.

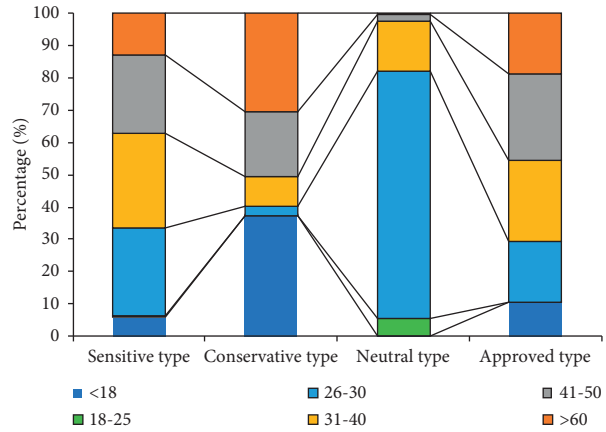


Figure 8: Age of drivers in the four segments.

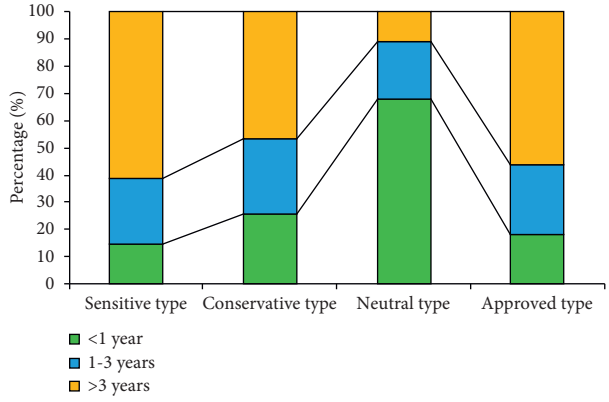


Figure 9: Driving experience of drivers in the four segments.

to the government and the media is basically the same as segment 2. The expectation for reduced parking fees is comparable to segment 1, lower than segment 4 but higher than segment 2, which is why we call it the "neutral type." Members of this segment are more often people younger than 25 years of age (76.4%) and more often have driving experience within 1 year (67.6%), which is also consistent with the distribution of age. Their average educational background is slightly lower than that of segment 1 but much higher than that of segments 2 and 4, and their monthly income is less than 5,000 yuan (90.5%), so it can be inferred that this group may be college students or newly

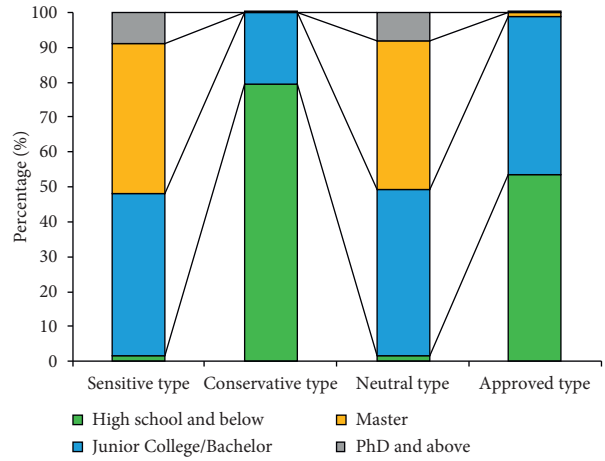


Figure 10: Education level of drivers in the four segments.

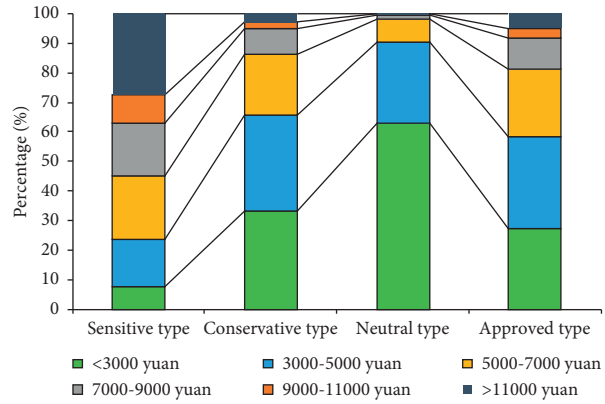


Figure 11: Income of drivers in the four segments.

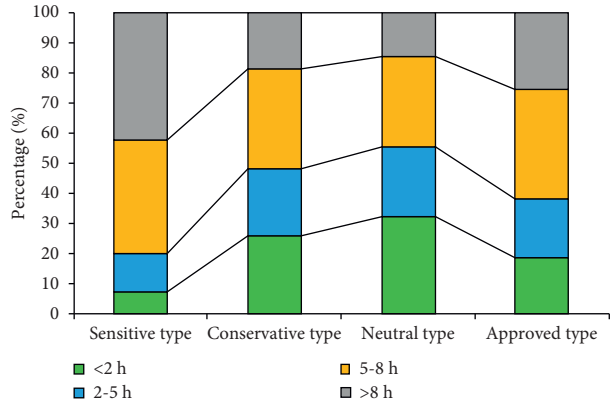


Figure 12: Parking time of drivers in the four segments.

graduated workers. In addition, they also have the lowest amount of parking time and the lowest frequency of driving trips, which reinforced our above hypothesis.

Segment 4 (9.9% of the respondents) has the highest score among all psychological factors. This type of driver has the smallest percentage. People in this segment are generally confident in their ability to control new things and are also willing to respond to the call of the government and the media. 82% of the drivers in this group tend to "replace long distances for lower costs," which also reflects that they are very appreciative of shared parking projects, so we call it the "approved type." In this segment, their ages are evenly distributed, their educational backgrounds are concentrated in undergraduates and below (98%), and their monthly

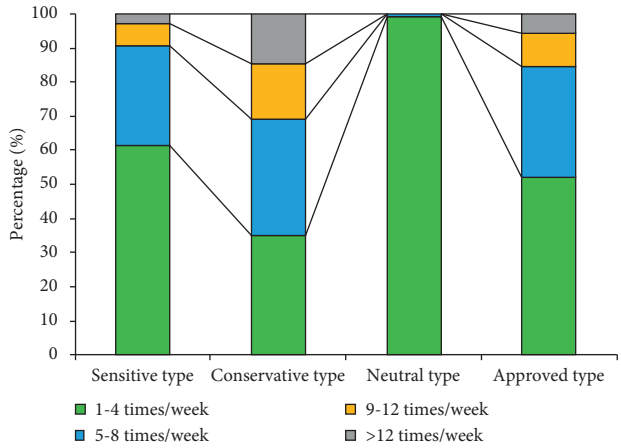


Figure 13: Driving frequency of drivers in the four segments.

income is concentrated below 9000 yuan. In addition, their driving frequency and parking time are both higher than segment 1 and lower than segment 2.

6. Conclusions and Discussion

In this paper, we use the output of the MIMIC model as the input variables of LCM to explore the SPUI of drivers considering individual heterogeneity. The results showed that the surveyed driver groups can be divided into four segments: sensitive type (36%), conservative type (29.6%), neutral type (24.5%), and approved type (9.9%). In exploring those characteristics, we can better understand "the respective natures of the four segments." This study contributes to the literature by providing first-hand insights into SPUI of different categories of drivers.

Firstly, the multiple indicator multiple cause model was used to explore the relationship between psychological factors, personal attributes, travel characteristics, and SPUI. The research shows the following: (1) The extended theoretical model has good explanatory power for SPUI, and SPUI is directly affected by perceived risk (PR), behavioral habit (BH), social influence (SI), facilitating conditions (FC), and effort expectancy (EE), while performance expectancy (PE) have no significant effect on SPUI. Among them, FC has the largest effect and PR has the smallest effect on SPUI. (2) Personal attribute variables including age, education level, monthly income, driving experience and driving frequency, and parking time in travel characteristics all have a significant influence on psychological variables.

Secondly, based on the analysis results of MIMIC, this paper uses LCM to explore the SPUI of different types of drivers from a microperspective. According to individual attributes, travel characteristics, and psychological factors, the potential classes of drivers are analyzed and simulated. Ten observation variables are used including age, education level, monthly income, driving experience, driving frequency, parking time, confidence in the control of shared parking systems (EE), response to the government and media (SI), perceived cost effectiveness (PE), and "long distance in exchange for less parking fees" tendency (FC). Finally, the surveyed driver groups are divided into four segments: sensitive type (36%), conservative type (29.6%), neutral type (24.5%), and approved type (9.9%). The results show that for different types of drivers, their SPUI is closely related to personal attributes, travel characteristics, and psychological factors. Therefore, in order to attract drivers to choose shared parking, it is recommended that managers or operators can start with psychological factors, such as effort expectations, facilitating conditions, performance expectancy, and social influence.

More targeted measures can be implemented for different categories of drivers:

(1) For the first segment (sensitive type), they are more optimistic about the potential economic benefits of shared parking, which may be due to their higher average educational level and stronger acceptance of emerging things, and are exposed to a wider range of online information, so they are more susceptible to government and media influence than segment 2 and segment 3. It is not difficult for shared parking operators to attract users from this largest proportion of the drivers through various promotional methods, such as providing incentives for demanders and suppliers. At the same time, government-level countermeasures are very important to substantially promote dominoreaction. The government can introduce policies and regulations to encourage the development of shared parking, thereby stimulating the development of the shared parking industry.

(2) For the second segment (conservative type), their attitude toward shared parking is more conservative, less susceptible to the influence of the government and the media, and more rational than other segments, but their confidence in the control of the shared parking system is less than that of segments 1 and 4. The perceived potential economic benefit is the smallest among the four segments. Therefore, saving more parking fees may be the most important point to attract them. Reasonable parking fees are

determined, and the negative impact of shared parking fees is reduced on them. For example, those who use shared berths flexibly or who have little parking demand can charge parking fees according to the parking time or the number of use times. For those who use shared parking spaces for a long time and have stable demand, a monthly or yearly fee with preferential rates can be adopted.

(3) For the third segment (neutral type), they have the least confidence in the control of the shared parking system. Therefore, creating a convenient, simple, and fully functional shared parking operating system may be the most important measure to attract them. However, since most of the available commercial maps cannot provide guidance in residential areas, it is important to display the turn-by-turn navigation function and detailed maps of the vacant parking spaces. In addition, some studies advocate the application of ergonomic principles in system design to make it more humane [73]. Finally, well-designed instruction manuals and training courses may help develop users' sense of self efficacy and ability to use technology [74]. These measures may help to improve users' technical compliance.

(4) For the fourth segment (approval type), they are very optimistic about the economic benefits of using shared parking. 82% of drivers have a strong intention to save parking fees even if the walking distance is increased. This group may have been exposed to the actual shared parking projects or in areas with high parking pressure. Therefore, the potential target areas of shared parking projects can be explored according to their spatial distribution characteristics.

However, this study inevitably has limitations that need to be resolved. Firstly, this study only considered the impact of latent variables and individual socioeconomic characteristics on the SPUI in residential areas and did not consider the impact of variables such as parking fees and walking distance, which is crucial for operators to determine reasonable prices. Secondly, this study explores the SPUI rather than the actual user behavior. In the future, the research can be extended to the investigation of the factors affecting actual use behavior. Thirdly, it is more valuable to jointly analyze the factors that affect the participation of suppliers and demanders in shared parking programs. This may promote better cooperation and matching between the two sides.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (No. 71971005) and the project was sponsored by Beijing Municipal Natural Science Foundation (No. 8202003). The supports are gratefully acknowledged.

References

[1] Beijing Transportation Development Research Institute, *2020 Annual Report on Beijing Transportation Development*, Beijing Research Institute of Transportation Development, Beijing, China, 2020.

[2] Beijing Municipal Commission of Communications, *Beijing Parking Resources Survey Report*, Beijing Municipal Commission of Communications, Beijing, China, 2017.

[3] R. Arnott and J. Rowse, "Modeling parking," *Journal of Urban Economics*, vol. 45, no. 1, pp. 97–124, 1999.

[4] D. C. Shoup, "Cruising for parking," *Transport Policy*, vol. 13, no. 6, pp. 479–486, 2006.

[5] D. Ayala, O. Wolfson, B. Xu, B. Dasgupta, and J. Lin, "Parking slot assignment games," in *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 299–308, Chicago, IL, USA, November 2011.

[6] Y. Cai, J. Che, C. Zhang, and B. Wang, "A parking space allocation method to make a shared parking strategy for appertaining parking lots of public buildings," *Sustainability*, vol. 11, no. 1, p. 120, 2019.

[7] Q. Yan, T. Feng, and H. Timmermans, "Investigating private parking space owners' propensity to engage in shared parking schemes under conditions of uncertainty using a hybrid random-parameter logit cumulative prospect theoretic model," *Transportation Research Part C: Emerging Technologies*, vol. 117, Article ID 102776, 2020.

[8] P. Zhao, H. Guan, P. Wang, and H. Yan, "Evaluation of environmental benefits caused by reservation-based shared parking: a case study of Beijing, China," *IEEE Access*, vol. 9, pp. 3744–3751, 2021.

[9] W. Zhang, Y. Su, J. Dai, and L. Wang, "Distributing model for shared parking in the residential zones," *Journal of Transportation Systems Engineering and Information Technology*, vol. 19, no. 1, pp. 89–96, 2019.

[10] X. Zhou, *Empirical Study on Sharing Willingness and Choosing Intention of Private Lots*, Nanjing University, Nanjing, China, 2018.

[11] H. Zhu, H. Guan, Y. Han, and W. Li, "A study of tourists' holiday rush-hour avoidance travel behavior considering psychographic segmentation," *Sustainability*, vol. 11, no. 13, p. 3755, 2019.

[12] S. Haustein, M. Thorhauge, and E. Cherchi, "Commuters' attitudes and norms related to travel time and punctuality: a psychographic segmentation to reduce congestion," *Travel Behaviour and Society*, vol. 12, pp. 41–50, 2018.

[13] S. Choi, J. Ko, and D. Kim, "Investigating commuters' satisfaction with public transit: a latent class modeling approach," *Transportation Research Part D: Transport and Environment*, vol. 99, Article ID 103015, 2021.

[14] C. M. Peterson, T. F. Nelson, and M. A. Pereira, "Driver speeding typologies by roadway behaviours and beliefs: a latent class analysis with a multistate sample of U.S. adults," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 81, pp. 373–383, 2021.

[15] A. Rahimi, G. Azimi, H. Asgari, and X. Jin, "Adoption and willingness to pay for autonomous vehicles: attitudes and latent classes," *Transportation Research Part D: Transport and Environment*, vol. 89, Article ID 102611, 2020.

[16] J. Xie, X. Ye, Z. Yang et al., "Impact of risk and benefit on the suppliers' and managers' intention of shared parking in residential areas," *Sustainability*, vol. 12, no. 1, p. 268, 2019.

[17] A. Wang, H. Guan, Z. Qin, and J. Zhu, "Study on the intention of private parking space owners of different levels of cities to participate in shared parking in China," *Discrete Dynamics in Nature and Society*, vol. 2021, Article ID 9955686, 16 pages, 2021.

[18] A. Wang, H. Guan, Y. Han, and Y. Cao, "Private parking space sharing intention in China: an empirical study based on MIMIC model," *Discrete Dynamics in Nature and Society*, vol. 2021, Article ID 9283686, 13 pages, 2021.

[19] J.-K. Liang, T. Eccarius, and C.-C. Lu, "Investigating factors that affect the intention to use shared parking: a case study of Taipei City," *Transportation Research Part A: Policy and Practice*, vol. 130, pp. 799–812, 2019.

[20] N. Lalani, "Evaluating shared parking for new developments," *Public Works*, vol. 115, no. 2, p. 64, 1984.

[21] T. Litman, "Parking management best practices," *Planning*, vol. 72, no. 9, pp. 40–45, 2006.

[22] C. Shao, H. Yang, and Y. Ke, "A simple reservation and allocation model of shared parking lots," *Transportation Research Part C: Emerging Technologies*, vol. 71, pp. 303–312, 2016.

[23] M. Iman and A. Hamid, "Applying shared-parking turn-time (SPPC) model and geographic information system in the supply and demand analysis of parking space," *Malaysian Journal of Real Estate*, vol. 1, no. 2, pp. 57–76, 2006.

[24] O. T. T. Kim, N. D. Tri, N. H. Tran, and C. S. Hong, "A shared parking model in vehicular network using fog and cloud environment," in *Proceedings of the 2015 17th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pp. 321–326, Busan, Repulic of Korea, August 2015.

[25] H. Xiao, M. Xu, and Z. Gao, "Shared parking problem: a novel truthful double auction mechanism approach," *Transportation Research Part B: Methodological*, vol. 109, pp. 40–69, 2018.

[26] P. Zhao, H. Guan, and P. Wang, "Data-driven robust optimal allocation of shared parking spaces strategy considering uncertainty of public users' and owners' arrival and departure: an agent-based approach," *IEEE Access*, vol. 8, Article ID 24182, 2020.

[27] A. Glazer and E. Niskanen, "Parking fees and congestion," *Regional Science and Urban Economics*, vol. 22, no. 1, pp. 123–132, 1992.

[28] R. G. Lipsey and K. Lancaster, "The general theory of second best," *The Review of Economic Studies*, vol. 24, no. 1, pp. 11–32, 1956.

[29] J. Chen and K. Xie, "Dynamic allocatin model and effect evaluation of campus shared-use parking in central city," *China Journal of Highway and Transport*, vol. 28, no. 11, pp. 104–111, 2015.

[30] K. He, J. Zhang, and Y. Zeng, "Households' willingness to pay for energy utilization of crop straw in rural China：Based on an improved UTAUT model," *Energy Policy*, vol. 140, Article ID 111373, 2020.

[31] G. Sheng, B. Yue, and S. Gong, "A research on consumer's intention to use shared bicycles continuously—based on extended model of TAM theory," *Journal of Northeastern University*, vol. 21, no. 6, pp. 567–574, 2019.

[32] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: toward a unified view," *MIS Quarterly*, vol. 27, no. 3, pp. 425–478, 2003.

[33] J. Chen, R. Li, M. Gan, Z. Fu, and F. Yuan, "Public Acceptance of Driverless Buses in China: An Empirical Analysis Based on an Extended UTAUT model," *Discrete Dynamics in Nature and Society*, vol. 2020, Article ID 4318182, 13 pages, 2020.

[34] R. Madigan, T. Louw, M. Wilbrink, A. Schieben, and N. Merat, "What influences the decision to use automated public transport? Using UTAUT to understand public acceptance of automated road transport systems," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 50, pp. 55–64, 2017.

[35] J. Lan and D. Zhu, "Acceptance and use behavior of sustainable transport: based on a survey of carsharing in Shanghai," *China Population, Resources and Environment*, vol. 26, no. 11, pp. 98–105, 2016.

[36] C. Chen and X. Li, "Research on the effect factors of customer use intention of shared bicycles," *Chinese Journal of Management*, vol. 15, no. 11, pp. 1601–1610, 2018.

[37] H. Aarts, B. Verplanken, and A. V. Knippenberg, "Predicting behavior from actions in the past: repeated decision making or a matter or habit?" *Journal of Applied Social Psychology*, vol. 28, no. 15, pp. 1355–1374, 1998.

[38] J. Chen, Q. Chen, and H.-P. Li, "Psychological influences on bus travel mode choice: a comparative analysis between two Chinese cities," *Journal of Advanced Transportation*, vol. 2020, no. 2, 9 pages, Article ID 8848741, 2020.

[39] T. Teichert, E. Shehu, and I. V. Wartburg, "Customer segmentation revisited: the case of the airline industry," *Transportation Research Part A: Policy and Practice*, vol. 42, no. 1, pp. 227–242, 2008.

[40] G. I. Crouch, T. Huybers, and H. Oppewal, "Inferring future vacation experience preference from past vacation choice," *Journal of Travel Research*, vol. 55, no. 5, pp. 574–587, 2014.

[41] C. F. Xiong, H. Pratt, and L. Zhang, "On Ride-Sharing: A Departure Time Choice Analysis with Latent Carpooling Preference," *Journal of Transportation Engineering*, vol. 140, no. 8, 2014.

[42] W. H. Greene and D. A. Hensher, "A latent class model for discrete choice analysis: contrasts with mixed logit," *Transportation Research Part B: Methodological*, vol. 37, no. 8, pp. 681–698, 2003.

[43] S. Shelat, R. Huisman, and N. van Oort, "Analysing the trip and user characteristics of the combined bicycle and transit mode," *Research in Transportation Economics*, vol. 69, pp. 68–76, 2018.

[44] M. Yanagihara, N. Uno, and T. Nakamura, "Latent class Analysis for driving behavior on merging section," *Transportation Research Procedia*, vol. 6, pp. 259–271, 2015.

[45] J. Alegre, S. Mateo, and L. Pou, "A latent class approach to tourists' length of stay," *Tourism Management*, vol. 32, no. 3, pp. 555–563, 2011.

[46] C. Pronello and C. Camusso, "Travellers' profiles definition using statistical multivariate analysis of attitudinal variables," *Journal of Transport Geography*, vol. 19, no. 6, pp. 1294–1308, 2011.

[47] S. Haustein and M. Hunecke, "Identifying target groups for environmentally sustainable transport: assessment of different segmentation approaches," *Current Opinion in Environmental Sustainability*, vol. 5, no. 2, pp. 197–204, 2013.

[48] J. T. Zhang, C. Jiao, and M. Q. Zhang, "Application of latent class Analysis in psychological research," *Advances in Psychological Science*, vol. 18, no. 12, pp. 1991–1998, 2010.
[49] B. Xu, *Research on the ICLV Model of Intercity Travel Mode Selection*, Beijing Jiaotong University, Beijing, China, 2018.
[50] S. Shu, Y. Bian, J. Rong, and S. Li, "Transition from driving to bicycling in short-distance travel based on MIMIC model," *Journal of Beijing University of Technology*, vol. 45, no. 10, pp. 998–1008, 2019.
[51] M. Hunecke, S. Haustein, S. Böhler, and S. Grischkat, "Attitude-based target groups to reduce the ecological impact of daily mobility behavior," *Environment and Behavior*, vol. 42, no. 1, pp. 3–43, 2010.
[52] P. Jing, J. Zhi-Cai, and Q. F. Zha, "Application of the expanded theory of planned behavior in intercity travel behavior based on MIMIC model," *Journal of Industrial Engineering and Engineering Management*, vol. 30, no. 4, pp. 61–68, 2016.
[53] F. Schneider and D. Enste, *Shadow Economies Around the World-Size, Causes, and Consequences*, IMF Working Papers, Washington, D.C, USA, 2000.
[54] H. Qiu, *Latent Class Modeling Principles and Techniques*, Educational Science Publishing House, Beijing, China, 2008.
[55] L. M. L. Bonilla and J. M. L. Bonilla, "Postmodernism and heterogeneity of leisure tourist behavior patterns," *Leisure Sciences*, vol. 31, pp. 68–83, 2009.
[56] C.-H. Wen and S.-C. Lai, "Latent class models of international air carrier choice," *Transportation Research Part E: Logistics and Transportation Review:Logistics and Transportation Review*, vol. 46, no. 2, pp. 211–221, 2010.
[57] S. Okazaki, S. Campo, L. Andreu, and J. Romero, "A latent class Analysis of Spanish travelers' mobile internet usage in travel planning and execution," *Cornell Hospitality Quarterly*, vol. 56, no. 2, pp. 191–201, 2015.
[58] E. Molin, P. Mokhtarian, and M. Kroesen, "Multimodal travel groups and attitudes: a latent class cluster analysis of Dutch travelers," *Transportation Research Part A: Policy and Practice*, vol. 83, pp. 14–29, 2016.
[59] X. H. Zeng, L. Xiao, and Y. B. Zhang, "Principle of latent class Analysis and case analysis," *Health Statistics Department of Public Health School*, vol. 30, no. 6, pp. 815–817, 2013.
[60] X. F. Ji and D. L. Li, "A model for segmentation of highway passenger groups based on latent class," *Journal of Highway and Transportation Research and Development*, vol. 36, no. 10, pp. 152–158, 2019.
[61] R. L. Tompson, C. A. Higgins, and J. M. Howell, "Personal computing: toward a conceptual model of utilization," *MIS Quarterly*, vol. 15, no. 1, pp. 125–143, 1991.
[62] I. Ajzen, "The theory of planned behavior," *Organizational Behavior and Human Decision Processes*, vol. 50, no. 2, pp. 179–211, 1991.
[63] I. Im, Y. Kim, and H.-J. Han, "The effects of perceived risk and technology type on users' acceptance of technologies," *Information & Management*, vol. 45, no. 1, pp. 1–9, 2008.
[64] M.-C. Lee, "Factors influencing the adoption of internet banking: an integration of TAM and TPB with perceived risk and perceived benefit," *Electronic Commerce Research and Applications*, vol. 8, no. 3, pp. 130–141, 2009.
[65] H. Wang, R. Li, X. Wang, and P. Shang, "Effect of on-street parking pricing policies on parking characteristics: a case study of Nanning," *Transportation Research Part A: Policy and Practice*, vol. 137, pp. 65–78, 2020.
[66] Z. Pu, Z. Li, J. Ash, W. Zhu, and Y. Wang, "Evaluation of spatial heterogeneity in the sensitivity of on-street parking occupancy to price change," *Transportation Research Part C: Emerging Technologies*, vol. 77, pp. 67–79, 2017.
[67] D. George and P. Mallery, *IBM SPSS Statistics 26 Step by Step: A Simple Guide and reference*, Routledge, London, UK, 2019.
[68] S. Ahmad, N. Zulkurnain, and F. Khairushalimi, "Assessing the validity and reliability of a measurement model in structural equation modeling (SEM)," *British Journal of Mathematics & Computer Science*, vol. 15, no. 3, pp. 1–8, 2016.
[69] G. Marcoulides, *Modern Methods for Business Research*, Psychology Press, New York, NY, USA, 1998.
[70] P. Cohen, S. G. West, and L. S. Aiken, *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, Psychology Press, London, UK, 2014.
[71] D. Pacifico and H. I. Yoo, "lclogit: a Stata command for fitting latent-class conditional logit models via the expectation-maximization algorithm," *STATA Journal: Promoting communications on statistics and Stata*, vol. 13, no. 3, pp. 625–639, 2013.
[72] C. C. Clogg, "Some latent structure models for the analysis of likert-type data," *Social Science Research*, vol. 8, no. 4, pp. 287–301, 1979.
[73] Y. Ning, M. Yan, X. S. Xu, Y. Li, and L. Li, "Shared parking acceptance under perceived network externality and risks: theory and evidence," *Transportation Research Part A: Policy and Practice*, vol. 150, pp. 1–15, 2021.
[74] J. A. Jacko, "Human Computer Interaction Handbook:," *Fundamentals, Evolving Technologies, and Emerging Applications*, CRC Press, Florida, FL, USA, Second edition, 2012.

Hindawi
Journal of Advanced Transportation
Volume 2021, Article ID 6243530, 14 pages
https://doi.org/10.1155/2021/6243530

Research Article

Simulation-Based Evaluation of Variation in Left-Turn Paths in the Coordinated Intersection Management

Menglin Yang[1,2] **Hao Yu**[1,2] **and Lu Bai**[1,2]

[1]*School of Transportation, Southeast University, Nanjing 210096, China*
[2]*Jiangsu Key Laboratory of Urban ITS, Nanjing 210096, China*

Correspondence should be addressed to Hao Yu; seudarwin@gmail.com

Received 18 November 2021; Accepted 7 December 2021; Published 23 December 2021

Academic Editor: Wenxiang Li

Coordinated intersection management (CIM) has gained more attention with the advance of connected and autonomous vehicle technology. The optimization of passing schedules and conflict separation between conflicting vehicles are usually conducted based on the predefined travelling paths through the intersection area in the CIM. In real-world implementation, however, the diversity of turn paths exists due to multiple factors such as various vehicle sizes and automation control algorithms. The aim of this paper is to investigate how the variation in left-turn paths affects the feasibility and viability of optimal passing schedules, as well as the safety and efficiency of intersection operation. To do this, we start with identifying six typical left-turn paths to represent the variation. A scenario-based simulation is first conducted by using each of the paths as the nominal path. The optimal schedules and the corresponding alternative schedules are generated to calculate indicators for nominal performance, average performance, and robustness. The best path is selected in terms of schedule optimality and robustness. With schedules obtained by solving CIM models using the selected path, the left-turning CAVs are assumed to travel along one of the six paths randomly to simulate the path divergence. A surrogate safety measure, PET, is utilized to assess the safety of the intersection under CIM. The theoretical PET with the nominal path and the actual PET with the random path are calculated for each conflict event. Comparisons of two PET sets show the increase in conflict risk and vehicle delay. The conclusion can be drawn that the variation in left-turn paths causes the decline in safety level and travelling efficiency and should be considered in the CIM model to ensure safe and efficient implementation in the intersection.

1. Introduction

Connected and autonomous vehicles, which react faster than human drivers and have communication functions, provide a promising approach to improve traffic operations while maintaining the maximum safety level for urban traffic network. With the implementation of techniques including vehicle-to-vehicle (V2V)/vehicle-to-infrastructure (V2I) communication and vehicle automation, coordinated intersection management (CIM) or autonomous intersection management (AIM) is proposed as a new intersection management method for signal-free intersections under 100% connected and autonomous vehicle (CAV) environment [1], which indicates that vehicles communicate with intersections and/or other vehicles to cooperate to pass an intersection without collisions. CIM is commonly implemented in two approaches, including the centralized and decentralized way. In the centralized way, CIM achieves the intersection level objectives, such as ensuring collision avoidance, minimizing total delay, reducing air pollution by assigning the vehicles' entering sequences and reaching times at the intersection [2–10]. More specifically, an intersection manager (IM) is deployed to communicate with CAVs via infrastructure-to-vehicle (V2I) communication and coordinate their passing schedules through the intersection. Here, passing schedules include arrival times at the intersection and the travelling speed through the intersection of all approaching vehicles. In the decentralized way, CAVs usually have their own distributed controller for trajectory optimization and executing vehicle control. They

communicate with ambient traffic or vehicles with conflicting movements to achieve consensus on collision-free trajectories [11–13]. The decentralized approach is also commonly applied in other scenarios such as cooperative metering [14, 15] and lane changing [16, 17]. Obviously, the communication burden is lower than the centralized way, but the coordination might not be optimal in terms of intersection operation due to the deficiency of the global scope of all approaching vehicles. Therefore, in the study, we concentrate on the centralized approach of CIM.

As illustrated in Figure 1, the communication and coordination area of an intersection virtually consists of two functional zones: multiple buffer zones (BZs) where vehicles adjust their speeds to pass the intersection as scheduled by the IM and a conflict zone (CZ) where all conflict points exist. The CIM aims to resolve conflicts of different vehicle movements within the CZ by the effective allocation of the right of way to approaching vehicles and to achieve the operation objectives, for example, increasing throughput and reducing energy consumption simultaneously. Specifically, the IM arranges the passing sequences of conflicting vehicles, such as vehicles travelling straight with those making the left turn from the opposite direction, through their potential conflicting areas at the CZ. Here, passing sequences do not refer to the order by which conflicting vehicles enter the intersection, but the order by which vehicles arriving and travelling through the conflicting area overlapped by two paths. On-Board Units (OBUs) of CAVs receive the assigned arrival times at the CZs and desired travel speeds through the intersections and manipulate the vehicles to realize the IM's schedule.

Lane-based trajectories are usually applied in approaching lane setting, signal timing planning, and traffic conflict analysis at intersections [18–20]. Vehicle movement paths and conflict points within the CZ are also predefined in the formulation of the CIM problem according to the lane-based assumption in previous studies of CIM. For example, Levin and Rey [5] assumed that turning movements are circular lines with determined left-turn and right-turn radius and formulated the conflict-point separation algorithm at intersections based on the predefined movement paths. Li et al. [8] utilized the planes instead of lines to represent vehicle trajectories with the consideration of vehicle sizes, where turning movements were still presented as quarter-circles with determined radiuses. In the research of Yu et al. [7], an intersection area was separated by a grid of tiles where vehicle trajectories were determined and represented by different sets of tiles. The trajectories of turning vehicles were assumed to be elliptical arcs of predefined radiuses to identify the occupied tiles and occupying time durations. With the decentralized approach of CIM, Zhang and Cassandras [13] used the turning radius of the circular path linking the approaching and destination lane to calculate the desirable times for vehicles' making a turn, which was provided for coordinated control of vehicles passing through the intersection. With the lane-based assumption, CAVs only need to design and implement one-dimensional speed/acceleration profiles along the predefined paths according to the received schedules.

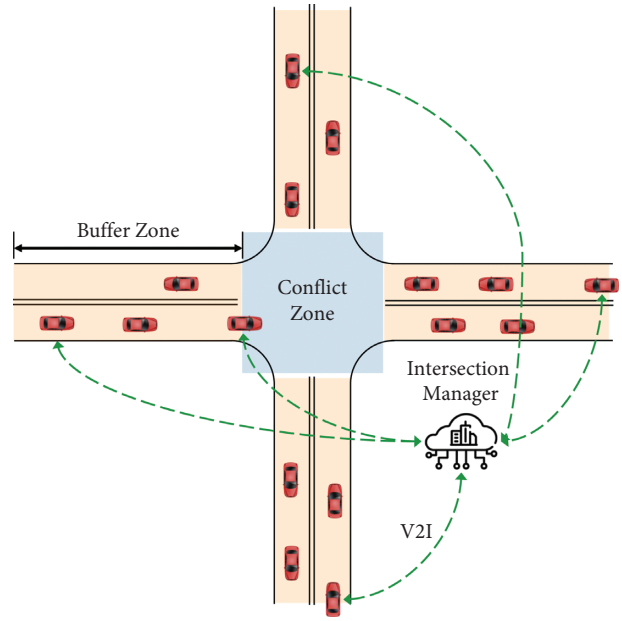


Figure 1: An exemplar scenario of CIM at a four-armed intersection.

However, the variation in vehicle turn paths is inevitable even for CAVs in real-world implementation. In the centralized CIM, in addition to target lane selection which is conducted by the IM, the controlling of CAVs through an intersection mainly consists of two tasks: trajectory planning and calculation and trajectory tracking and maneuvering. For straight travelling vehicles, one-dimensional speed/acceleration profiles are calculated and easy to track due to the few variations in the geometric path. For turning vehicles, however, the tasks are complicated since the geographical trajectory planning and calculation should generate both speed/acceleration and turning angle profiles, which needs to determine the turning pattern and the turning radius first. In this stage, the variety in vehicle sizes [21], planning algorithms [22], intersection geometric property [23–25], and desired turning speeds [21, 26, 27] directly leads to different turning patterns and radiuses, referring to variation in both vehicle position and speed trajectories. In the trajectory tracking and maneuvering, the speed and turning angle are required to be adjusted simultaneously. Diversity in the control performance of different controllers and health status of motor drive systems also cause errors or failures in tracking planned turning trajectories in terms of geometric path and turning speed. Therefore, even with assigned passing schedules and target lanes, the actual turns may not be made along circular arcs linking approaching and target lanes or other predefined paths by the IM.

Due to the significance of predefined paths in conflicting relationship determination and subsequent conflict separation design and the inevitability of path divergence from the predefined ones, it is questioned whether the feasibility and viability of schedules and collision-free operation would remain. Therefore, it is of great importance to investigate two aspects: what effects the variation in turn paths has on the schedules and how the safety and travelling efficiency of

the intersection could vary due to the path divergence. Since the left-turn path is longer and has more conflict points, this study concentrates on the variation in the left-turn path.

The centralized CIM scheduling problem is usually modelled as a mixed integer linear programming problem (MILP), with the passing sequences of conflicting vehicles as binary variables and arrival times at the intersection as continuous variables [5–7]. In the model formulation, the variations of turn paths are reflected in the model parameters. Hence, the primary trial is analysing the impact of the variations on the solutions (passing schedules) by the sensitivity analysis of the mathematical model. However, sensitivity analysis in the integer programming is still not mature due to the lack of optimality criteria, although it is well developed in the linear programming. Dawande and Hooker [28] developed a method of sensitivity analysis for MILP derived from the idea of inference duality to obtain the perturbations of parameters that reduce the optimal value by no more than a prespecified amount. Utilizing the inference-based sensitivity analysis of MILP, Jia and Ierapetritou [29] proposed an integrated framework to address the issue of uncertainty in short-term scheduling problems of batch plants. However, the inference-based sensitivity analysis is not suitable for MILP problems in the CIM. The reason is that parameters in the corresponding MILP are derived and calculated from real-world information (i.e., geometric data of different turning paths). Those parameters have no direct physical meanings and are highly related to each other. Therefore, the achieved perturbations in those parameters from the inference-based sensitivity analysis are not useful for the real-world analysis such as feasibility of passing schedules. Besides, the inference-based sensitivity analysis addresses the MILP solved using the branch-and-bound methodology, making it difficult to be adopted due to computational complexity with the rise in the number of integer variables, which has a positive correlation with traffic demand through the intersection.

Without appropriate mathematical tools, it is impossible to investigate the variation issues in left-turn paths analytically. Therefore, we propose a scenario-based approach to evaluate the effects of variation in left-turn paths on travel efficiency and safety performance of CIM. Six typical left-turn paths are selected from two main turning patterns of different turning radius to be applied in the numerical tests. The scenario-based evaluation is conducted to assess the optimality and robustness of schedules obtained by using different left-turn paths as the nominal path. Based on the evaluation result, schedules are determined using the best path(s) and left-turning CAVs travel along the random left-turn paths according to the scheduled arrival times and desired speeds. Safety and travelling efficiency of the intersection are then estimated based on the schedules and the actual operation with path divergence, respectively, and compared to assess the effects of the variation in left-turn paths.

This paper is organized as follows: the scheduling model to optimize vehicle passing arrangements in the CIM is briefly described and analysed in Section 2. Section 3 provides two simulation-based evaluation approaches for analysing the variation in left-turn paths. Section 4 presents the safety and efficiency analysis based on the CIM simulation results. Conclusive comments are listed, and future research directions are discussed in the last section.

2. CIM Scheduling Model

2.1. Original Model. In this paper, the mathematical model used for vehicle passing schedule optimization in the CIM is formulated based on the vehicle movement paths within the intersection [5]. The paths (denoted by ρ) are discretized by critical points, and the critical points contain two parts: the starting points (denoted by o) and ending points (denoted by f) of the allowed paths on the boundary of the conflict zone and the conflict points (denoted by c) where conflicting paths overlap. The exemplar intersection is illustrated in Figure 2, along with all movements, paths, and critical points within the CZ.

The objective to minimize the integrated passing time of the intersection over all vehicles in one optimization horizon, and other constraints in the MILP model are as follows:

$$z = \min \sum_{i \in \Omega} p_i t_i(f), \tag{1}$$

subject to

$$t_i(c) + \tau_i(c) - t_j(c) \le \left(1 - \delta_{ij}\right) M, \forall i, j \in \Omega,\ c = \rho_i \cap \rho_j, \rho_i \ne \rho_j, \tag{2}$$

$$t_j(c) + \tau_j(c) - t_i(c) \le \delta_{ij} M, \forall i, j \in \Omega,\ c = \rho_i \cap \rho_j, \rho_i \ne \rho_j, \tag{3}$$

$$t_i(o) + \tau_i(o) \le t_j(o), \forall i, j \in \Omega,\ o \in \rho_i \rho_i = \rho_j, \tag{4}$$

$$t_i(f) + \tau_i(f) \le t_j(f), \forall i, j \in \Omega,\ f \in \rho_i, \rho_i = \rho_j, \tag{5}$$

$$\frac{d_i(o,f)}{\overline{U}_i} \le t_i(f) - t_i(o) \le \frac{d_i(o,f)}{\underline{U}_i}, \forall i \in \Omega, o, f \in \rho_i, \tag{6}$$

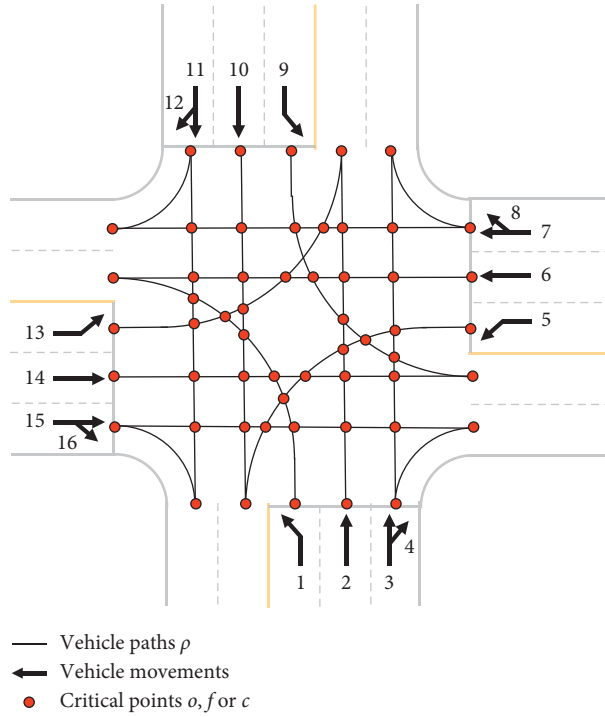


Figure 2: A four-armed exemplar intersection.

$$\frac{t_i(f)-t_i(o)}{d_i(o,f)}=\frac{t_i(c)-t_i(o)}{d_i(o,c)},\forall i\in\Omega,\ o,f\in\rho_i,c=\rho_i\cap\rho_j,\rho_i\neq\rho_j, \tag{7}$$

$$t_i(o)\geq\max\{t_i(B)+e_i,h_i\},\forall i\in\Omega,o\in\rho_i, \tag{8}$$

$$t_i(\cdot)\geq 0,\forall i\in\Omega, \tag{9}$$

$$\delta_{\mathrm{ij}}=\{0,1\},\ \forall i,j\in\Omega,\ \rho_i\neq\rho_j, \tag{10}$$

where Ω is the vehicle set consisting of all vehicles in one optimization horizon and i and j are the indices of vehicles. ρ represents any path illustrated in Figure 2. p_i represents the priority of vehicle passing sequences. $t_i(\bullet)$ refers to the arrival time at one critical point of vehicle i and $\bullet$ can be either a starting point (o), an ending point (f), or a conflict point (c) of one path. $\tau_i(\bullet)$ refers to the reservation time duration of such a point for vehicle i, consisting of the encroachment time of the point by vehicle i and the predefined clearance time, as expressed by equation (11). $d_i(o,\bullet)$ is the distance from the starting point to another critical point along path ρ_i. $t_i(B)+e_i$ denotes the earliest possible arrival time at the CZ of vehicle i, where $t_i(B)$ is its arrival time at the buffer zone and e_i is the estimated shortest time travelling through the BZ. δ_{ij} is a binary variable indicating the passing sequence through the conflict point of vehicles i and j.

$$\tau_i=\frac{L_i}{u_i}+g=\frac{L_i\left(t_i(c)-t_i(o)\right)}{d_i(o,c)}+g,\forall i\in\Omega,o\in\rho_i,c=\rho_i\cap\rho_j,\gamma_i\neq\gamma_j, \tag{11}$$

where L_i denotes the distance that vehicle i entirely passes through a point (i.e., $c=\rho_i\cap\rho_j$) along path ρ_i. g represents the minimum clearance time. The detailed calculation process of L_i can be found in [5].

In general, equations (2) and (3) represent the determination of passing sequences at conflict point c. Specifically, when $\delta_{ij}=1$, equation (2) works, while equation (3) is inactive, indicating that vehicle i passes ahead of vehicle j over point c. When $\delta_{ij}=0$, equation (3) works, while equation (2) is inactive, indicating the reverse passing sequence. Equations (4) and (5) constrain the car-following behaviors of vehicles on the same path. The upper and lower speed boundaries of the intersection are expressed by equation (6), and constant speed limitation within the conflict zone is indicated by equation (7). Equation (8) indicates the estimated earliest arrival time at the conflict

zone of vehicle i. Basic constraints of variables are expressed by equations (9) and (10).

2.2. Reformulation. To clearly display the relationship between MILP parameters and factors at the real-world level, we reformulate the original MILP (equations (1)–(10)). With no priority of any vehicle, let us set $p_i = 1$, $\forall i \in \Omega$. We take $x_i = t_i(f) - t_i(o)$ and t_i (short for $t_i(o)$) as the continuous variables. $t_i(c)$ is transformed to be the linear combination of x_i and t_i based on equation (7). Therefore, the problem is reformulated as

$$z = \min \sum_{i \in \Omega} (t_i + x_i), \tag{12}$$

subject to

$$-t_i + t_j - (\alpha_{ic} + \beta_i) x_i + \alpha_{jc} x_j - M\delta_{ij} \geq g - M, \forall i, j \in \Omega, c = \rho_i \cap \rho_j, \rho_i \neq \rho_j, \tag{13}$$

$$t_i - t_j + \alpha_{ic} x_i - (\alpha_{jc} + \beta_j) x_j + M\delta_{ij} \geq g, \forall i, j \in \Omega, c = \rho_i \cap \rho_j, \rho_i \neq \rho_j, \tag{14}$$

$$\begin{aligned} -\left[1 + \frac{L_i}{L}\right] \cdot t_i + t_j &\geq -\frac{L_i}{L} \cdot t_j(B) + g, \forall i, j \in \Omega, \rho_i = \rho_j, \\ -t_i + t_j - [1 + \beta_i] \cdot x_i + x_j &\geq g, \forall i, j \in \Omega, \rho_i = \rho_j, \end{aligned} \tag{15}$$

$$x_i \geq \underline{\tau_i}, \forall i = 1, \ldots, 7, \tag{16}$$

$$-x_i \geq -\overline{\tau}_i, \forall i = 1, \ldots, 7, \tag{17}$$

$$\begin{aligned} t_i &\geq \gamma_i, \forall i = 1, \ldots, 7, \\ \delta_{ij} &= \{0, 1\}, \forall i, j \in \Omega,\ c = \rho_i \cap \rho_j, \rho_i \neq \rho_j, \end{aligned} \tag{18}$$

where L is the length of the buffer zone, $\alpha_{ic} = d_i(o, c)/d_i(o, f)$, $\beta_i = L_i/d_i(o, f)$, $\tau_i = d_i(o, f)/\overline{U}_i$, $\overline{\tau}_i = d_i(o, f)/\underline{U}_i$, and $\gamma_i = \max\{t_i(B) + e_i, \underline{h_i}\}$.

Equations (16) and (17), which are derived from equation (6), represent the upper and lower bounds of x_i. Vehicle passing arrangements through the CZ can be interpreted from the variable values obtained by solving the MILP problem. To be specific, the passing sequences are indicated directly by values of δ_{ij}, the arranged arrival times at the CZ are achieved directly by values of t_i, and vehicle cruise speeds (v_i) through the CZ can be calculated as follows:

$$v_i = \frac{d_i(o, f)}{x_i}, \forall i \in \Omega. \tag{19}$$

2.3. Variation in Left-Turn Paths. Turning pattern and radius are two main features of a turn path. Left-turn patterns are classified into two types: two-stage pattern (straight and turning) and one-stage pattern (directly turning) to generally model the formation of path diversity [30]. Here, six typical left-turn paths are selected to represent the variation of left-turn paths, including three of the two-stage pattern (Paths $C1$, $C2$, and $C3$ in Figure 3(a)) and three of the one-stage pattern (Paths $E1$, $E2$, and $E3$ in Figure 3(b)), where w and W refer to the width of a vehicle and a lane, respectively. The middle paths (Paths $C2$ and $E2$) are on the centerline of lanes, while the external paths (Paths $C1$ and $E1$) and the internal paths (Paths $C3$ and $E3$) are the paths where vehicles travel along the boundaries of the lane. For each pattern, the most common condition (travelling on the centerline) and two extreme conditions (travelling laterally along the two sides of the lane) are presented and considered in the study. The locations of related conflict points also vary with the deviation of paths. These variations are mainly reflected by the changes of $d_i(o, \bullet)$ and accordingly the changes in values of parameters α_{ic}, β_i, and τ_i in the MILP.

3. Simulation-Based Evaluation

3.1. A Scenario-Based Approach. With the selected left-turn paths, we propose a scenario-based approach to analyse the efficiency and robustness of schedules solved from the CIM model, to investigate whether a schedule obtained with a predefined path can accommodate the variations in actual paths and still holds valid. The proposed scenario-based approach (see Figure 4) comprises the following steps:

Step 1. Six scenarios k ($k = 1, 2, \ldots, 6$) are defined where the typical left-turn paths (Paths $C1$, $C2$, $C3$, $E1$, $E2$, and $E3$) are used as the nominal path, respectively, and the corresponding parameter sets (α_{ic}, β_i, and τ_i) of the MILP problem are calculated with the geometry of different left-turn paths under each scenario.

Step 2. An MILP problem is solved based on the nominal values of the varying parameter set (α_{ic}, β_i, and τ_i) from the scenario k. A set of optimal fixed sequences of vehicle passing through the CZ (δ_{ij}) is obtained, and the solution is denoted as Schedule k. The objective

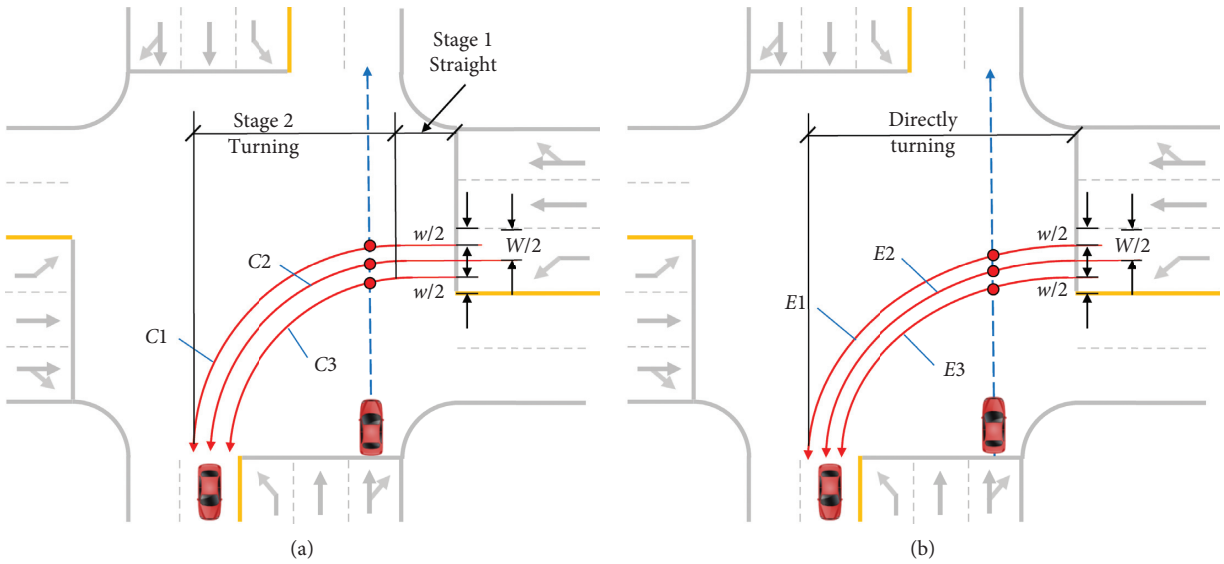


Figure 3: Typical left-turn paths. (a) Two-stage left-turn paths. (b) One-stage left-turn paths.

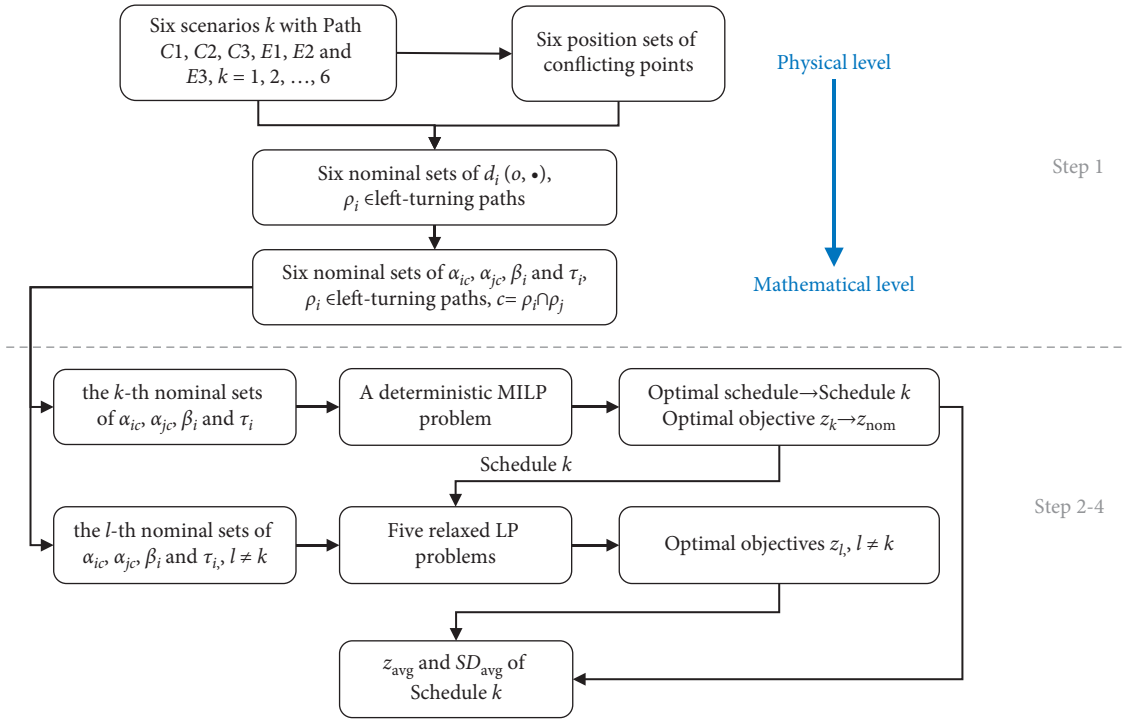


Figure 4: Procedure of the scenario-based approach.

value (z_k) is recorded as the nominal performance (z_{nom}) of this schedule.

Step 3. Using the optimal sequences of Schedule k in Step 2, the MILP problem is relaxed to be LP problems using parameter sets of alternative paths from all other scenarios l ($l \neq k$). The alternative schedules l and the optimized objective values (z_l) are obtained by solving those LP problems.

Step 4. The average performance and robustness metric of Schedule k are calculated. To be specific, the average performance of Schedule k (z_{avg}) is calculated as follows:

$$z_{avg} = \frac{z_k + \sum_l z_l}{p_{tot}}, \tag{20}$$

where p_{tot} is the total number of scenarios. The average performance (z_{avg}) is used to access the performance of a schedule over the entire variation interval.

The standard deviation (SD) is applied as the robustness metric for Schedule k. Like Jia and Ierapetritou [29], the SD is defined as

$$\text{SD}_{\text{avg}} = \sqrt{\sum_{k} \frac{\left(z_k - z_{\text{avg}}\right)^2}{\left(p_{\text{tot}} - 1\right)}}. \tag{21}$$

As expressed by equation (12), the objective z is the integration of all vehicles' arranged arrival time (t_i) at the intersection and travelling time (x_i) through the CZ in one optimization horizon. Lower z_{nom} represents the earlier, and lower z_{avg} represents the earlier average integrated passing time of vehicles over all scenarios with different left-turn paths, indicating the higher travelling efficiency of the intersection. Lower SD_{avg} of a schedule represents less fluctuation in the objective value under different left-turn paths. With the same vehicle passing sequences, this presents how close the passing arrangements (t_i and x_i) under different scenarios are to each other, which indicates the robustness of a schedule in terms of travelling efficiency. To evaluate the effects of the variation in left-turn paths on CIM, simulations are conducted under different traffic demand levels by the scenario-based approach using the six typical paths mentioned above. To eliminate the impacts of vehicle arrival times and vehicle numbers of different cases on the values of schedule performance metrics, we utilize travel delay per vehicle as the objective function of the MILP, which is modified as

$$z' = \frac{1}{s} \min \sum_{i \in \Omega} \left[t_i + x_i - \left(t_i(B) + e_i + \frac{d_i(o,f)}{\overline{U}_i} \right) \right], \forall i \in \Omega, \tag{22}$$

where s is the total number of vehicles in Ω.

Nominal performance (z'_{nom}), average performance (z'_{avg}), and robustness metric (SD'_{avg}) of six schedules are calculated to evaluate whether a schedule obtained by using a certain path outperforms the others. This selected path will be used for further evaluation.

Note that higher SD_{avg} values can indicate that the arranged arrival times vary seriously of at least some vehicles. The lower values, however, cannot refer to the less variation in the arranged arrival times of each vehicle. Therefore, the PET-based approach is proposed and then proposed to analyse the effects of path variation on individual vehicles and conflicts.

3.2. A PET-Based Approach. As demonstrated in equations (13) and (14), the potential conflict between two conflicting vehicles is resolved by determining the passing sequence and minimum clearance time of the conflict point based on the predefined paths. The path divergence between theoretical ones in the CIM models and actual ones in the schedule implementation would affect the actual locations and clearance time of the conflict points, raising safety concerns of failures in conflict separation, leading to collisions subsequently. To evaluate the traffic collision risks considering the variation in left-turn paths, Postencroachment Time (PET) is utilized in this study. Peesapati, Hunter, and Rodgers [31, 32] evaluated the effectiveness of PET as an indicator of safety assessment and as a surrogate measure for the propensity of crashes between left-turning vehicles and opposing through vehicles at four-armed signalized intersections.

As illustrated in Figure 5, PET is the difference between the time when the first vehicle (veh i) ends encroachment over the conflict point and the second vehicle (veh j) arrives the point, representing the proximity of conflicting vehicles to their actual crossing point and it has no requirement on the collision course [33, 34]. Only two timestamps are required to compute PET, and it has a definite boundary to differentiate a crash from a noncrash event. A PET value of 0 implies the occurrence of a collision, and the closer the value of PET is to 0, the higher the conflict risk is. In this study, two threshold values of PET play an important role in establishing its correlation with collisions as well as traffic safety. In the CIM model, a minimum clearance time g is set to ensure collision-free operation. Therefore, when the PET between a pair of conflicting vehicles is lower than g, it is considered as a serious conflict. The other threshold is set to filter out the nonconflict events from the analysis to eliminate the bias in the safety evaluation results [35], which usually takes 3 seconds [36].

Further simulations are conducted under different traffic demand levels where the schedules are obtained by using the path selected by the scenario-based approach, and left-turning vehicles are assumed to travel along one of the six typical paths randomly. All potential conflict events involving left-turning vehicles are collected, and two PETs are captured for each event, including theoretical PET calculated based on the nominal path used in the model, denoted by PET_t, and actual PET calculated based on the actual travel path, denoted by PET_a. Here, a potential conflict event refers to the event when two vehicles travelling on conflicting movements occupy the CZ simultaneously. Two sets of PETs will be compared and analysed to evaluate the effects of the variation in left-turn paths on intersection safety and travel efficiency.

4. Numerical Simulation and Result Analysis

Simulations are conducted in MATLAB R2021a, including random vehicle arrival generation and CIM operation. The MILP model and the relaxed LP problems are formulated and organized as the standard form in MATLAB, solved by the build-in Mixed Integer Programming (MIP) solver in GAMS [37].

In the simulations, the exemplar intersection is deadlocked when the traffic demand exceeds 900 veh/h/lane, and the vehicle arrives sparsely when the traffic demand falls below 500 veh/h/lane, in which the effects of variations in the left-turn path can hardly be identified clearly. As a result, the traffic volume on straight movements is set from 500 to

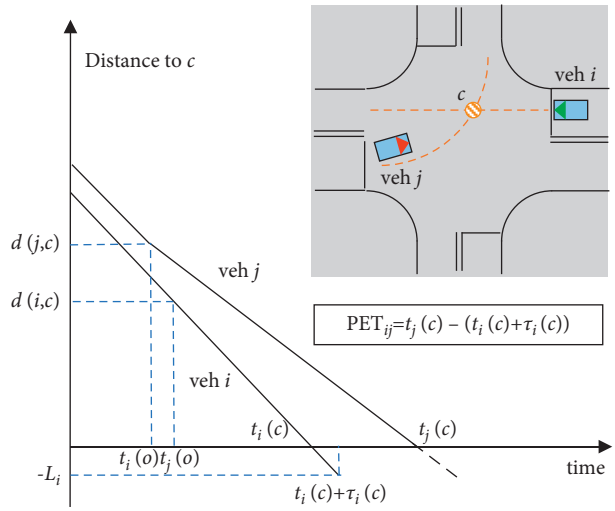


Figure 5: Traffic conflict diagram for PET.

900 veh/h/lane with a 100 veh/h/lane step. Left-turn and right-turn volumes are both set to be 20% of the straight traffic volume. For each condition, the simulation runs for 150 minutes. The minimum clearance time g is set to be 0.8 seconds, and the threshold to distinguish traffic conflict and nonconflict events is 3 seconds.

4.1. Travelling Efficiency and Robustness with Different Left-Turn Paths. The CIM process is conducted in a rolling-forward approach, and the indicators including nominal performance (z'_{nom}), average performance (z'_{avg}), and robustness metric (SD'_{avg}) are collected for each optimization horizon. Table 1 demonstrates the average values of nominal and average performance (z'_{nom} and z'_{avg}) of all cases under each traffic demand level, where "Sched." is short for "Schedule." The SD'_{avg} of six schedules under all traffic demand levels are illustrated in Figure 6.

As listed in Table 1, Schedule 3 and Schedule 5 generally have both higher nominal performances and average performances (lower average delay per vehicle) than others across all traffic demand levels, indicating higher travelling efficiency of vehicle passing schedules. Considering the SD'_{avg} illustrated in Figure 6, Schedule 5 provides better solutions than Schedule 3 in terms of robustness metric, indicating better accommodation to the variation in left-turn paths. Therefore, Path $E2$, as the nominal path, and the corresponding parameters will be used when modelling the CIM and formulating the MILP problem in the following simulation.

Schedule 1 and Schedule 4 have obvious worse robustness than others, as shown in Figure 5. The corresponding left-turn paths are Path $C1$ and Path $E1$, which have the largest turning radiuses among each turning pattern. The optimal schedules obtained with longer nominal paths reserve more time for left-turning vehicles passing through conflict points as well as the intersection, of which the passing sequences are more likely not to be optimal for scenarios with shorter nominal paths, leading to more fluctuation in the average delay among different schedules. On the other side, although the optimal schedules obtained with shorter nominal paths reserve less time for left-turning vehicles, the passing sequences may be still optimal for scenarios with longer nominal paths since the clearance time can be reduced as long as it satisfies the minimum clearance time.

4.2. Safety and Travelling Efficiency with the Path Divergence. In this section, PET_t is calculated from schedules obtained by using Path $E2$ as the nominal path. Figure 7 displays the distribution of potential conflict events in terms of their PET_a values against the increases compared to their PET_t. Some serious conflicts are observed, of which PET_a is lower than the minimum clearance time. To eliminate the effects of suboptimality of solutions from the solver, conflicts of which PET_a <0.8 and those of which PET_t <0.8 are extracted and compared in Table 2. Conflicts of which PET_t <0.8 are caused by the solver and its solving algorithms. The obvious increases in both the numbers of conflicts with PET_a <0.8 and their percentages across all traffic demand levels demonstrate higher collision risk at the intersection when the variation in the left-turn path exists but is not considered in the scheduling.

Another interesting change is the increase in the nonconflict events (PET >3) with rising traffic demand, as shown in Table 2. With the increased traffic volume, more potential conflicts involving the same vehicle need to be resolved simultaneously in one optimization horizon, and the solution might lead to higher PET values of some potential conflict events to ensure conflict separation for all vehicles.

After filtering out the nonconflict events, the distributions of conflicts under different traffic demand levels are demonstrated in Figure 8 to depict the fluctuations in conflict risk with the actual path divergence from the nominal path. The conflict numbers of which PET_a decreases compared to PET_t clearly surpass those of which PET_a increases under each traffic demand level. The overall effect of

Table 1: Travelling efficiency performance of different nominal left-turn paths.

Traffic demand (veh/h/lane)		Sched. 1 (Path $C1$)	Sched. 2 (Path $C2$)	Sched. 3 (Path $C3$)	Sched. 4 (Path $E1$)	Sched. 5 (Path $E2$)	Sched. 6 (Path $E3$)
500	z'_{nom}	0.488	0.487	0.481	0.487	0.485	0.484
	z'_{avg}	0.503	0.490	0.483	0.491	0.482	0.483
600	z'_{nom}	0.653	0.652	0.651	0.653	0.652	0.658
	z'_{avg}	0.682	0.661	0.653	0.667	0.655	0.653
700	z'_{nom}	0.871	0.877	0.876	0.874	0.878	0.879
	z'_{avg}	0.908	0.890	0.878	0.894	0.880	0.876
800	z'_{nom}	1.090	1.096	1.092	1.093	1.093	1.103
	z'_{avg}	1.129	1.109	1.096	1.113	1.098	1.098
900	z'_{nom}	1.924	1.928	1.927	1.927	1.929	1.941
	z'_{avg}	1.977	1.947	1.935	1.952	1.936	1.937

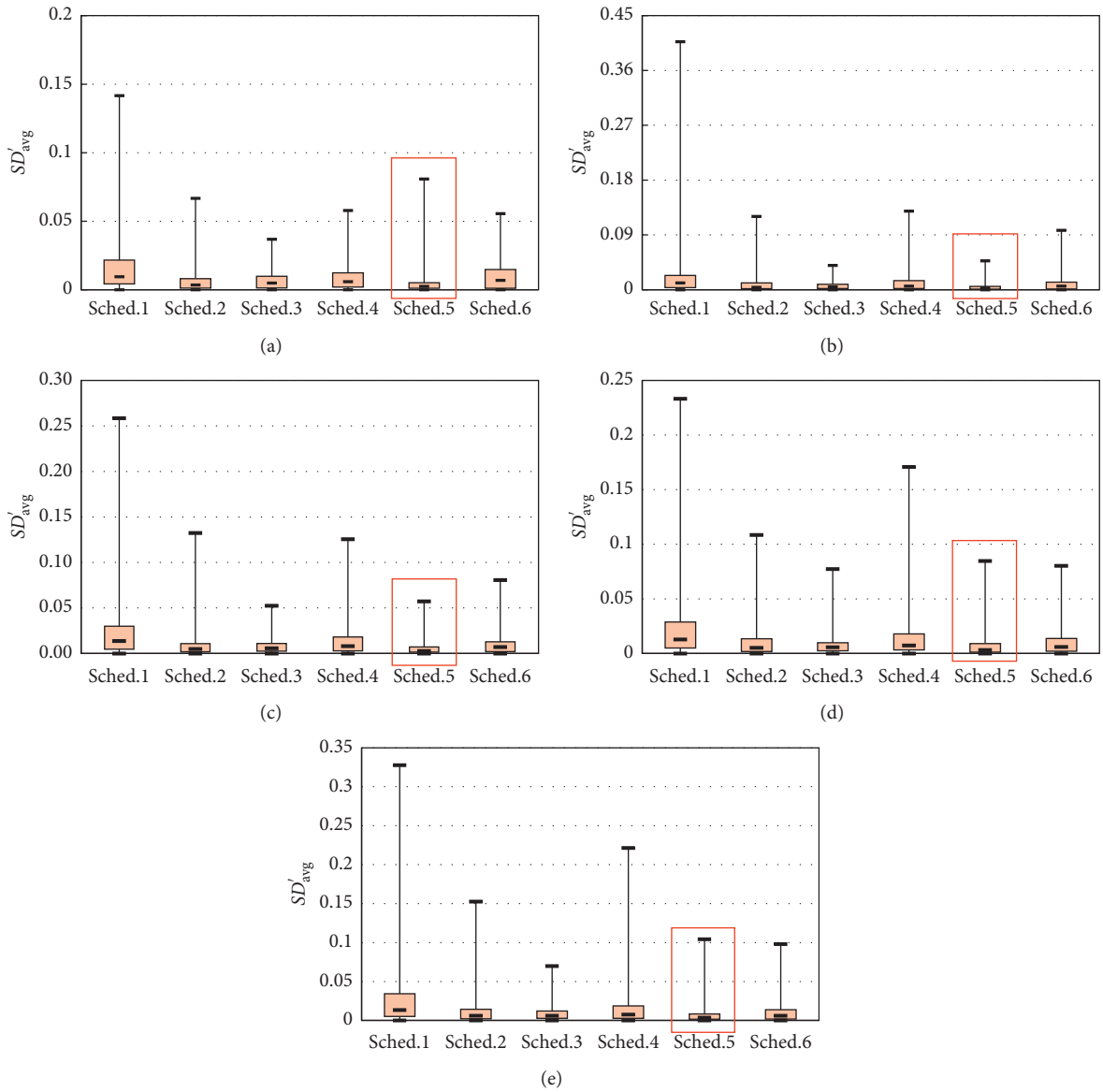


Figure 6: Schedule robustness of different nominal left-turn paths. (a) 500 veh/h/lane. (b) 600 veh/h/lane. (c) 700 veh/h/lane. (d) 800 veh/h/lane. (e) 900 veh/h/lane.

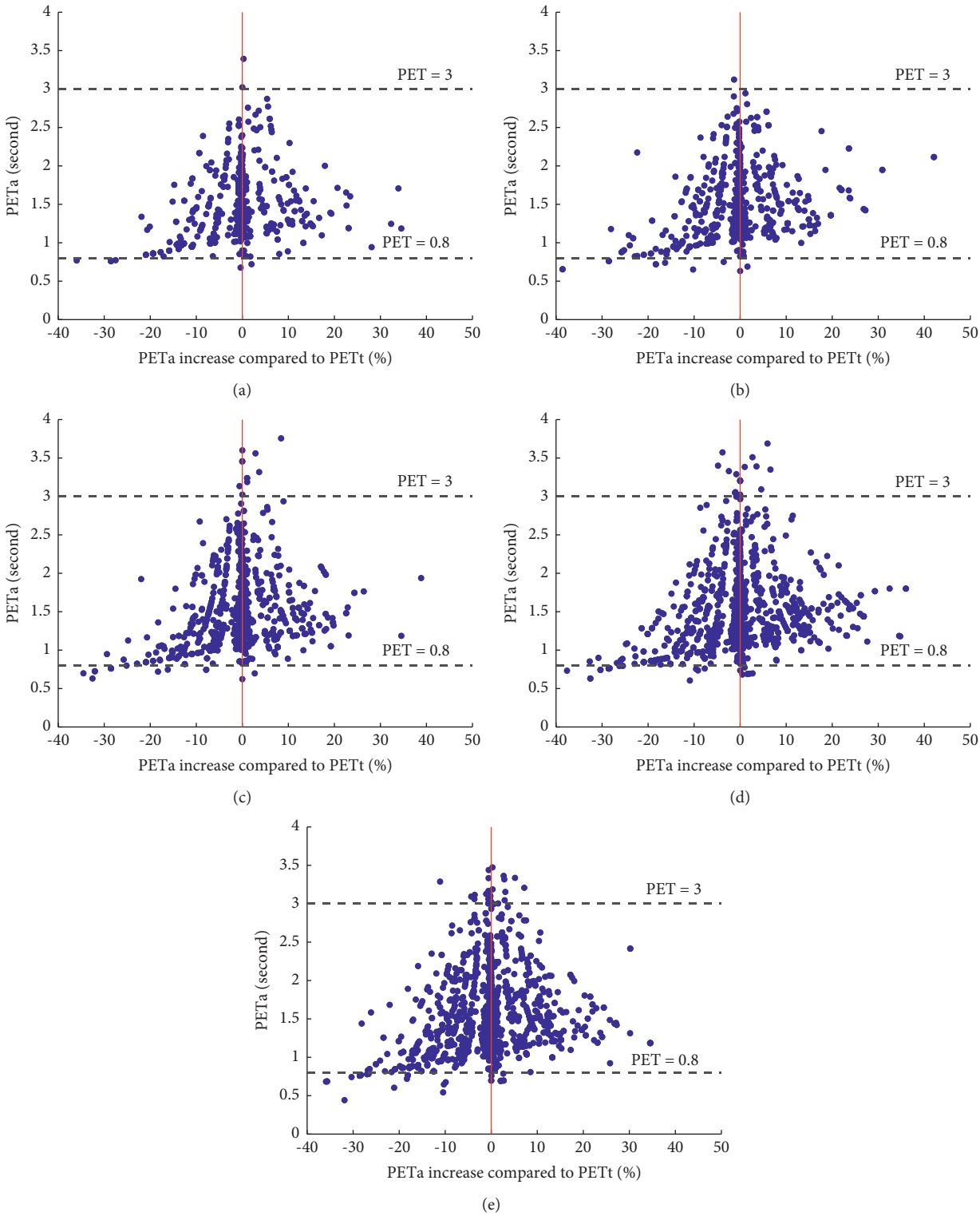


FIGURE 7: Relationships between PET_a and PET_a increase compared to PET_t. (a) 500 veh/h/lane. (b) 600 veh/h/lane. (c) 700 veh/h/lane. (d) 800 veh/h/lane. (e) 900 veh/h/lane.

actual path divergence on safety is negative since the decrease of PET_a indicates the increase in conflict/collision risk. Hence, the variations in the left-turn path should be seriously considered in the scheduling to improve safety.

Apart from the safety issue, the travelling efficiency is also jeopardized due to the path divergence. During CAVs travelling across the intersection according to the schedules, OBUs are also monitoring the movements of surrounding

Table 2: Conflicts of which PET_t and PET_a are beyond two thresholds.

Thresholds		500 veh/h/lane	600 veh/h/lane	700 veh/h/lane	800 veh/h/lane	900 veh/h/lane
PET_t <0.8 sec	Number	4	4	2	8	17
	Percentage	0.74	0.53	0.19	0.56	1.00
PET_a <0.8 sec	Number	9	10	13	27	29
	Percentage	1.65	1.34	1.23	1.92	1.71
PET_t >3 sec	Number	2	2	11	19	23
	Percentage	0.37	0.27	1.04	1.34	1.35
PET_a >3 sec	Number	2	2	11	16	26
	Percentage	0.37	0.27	1.04	1.14	1.54

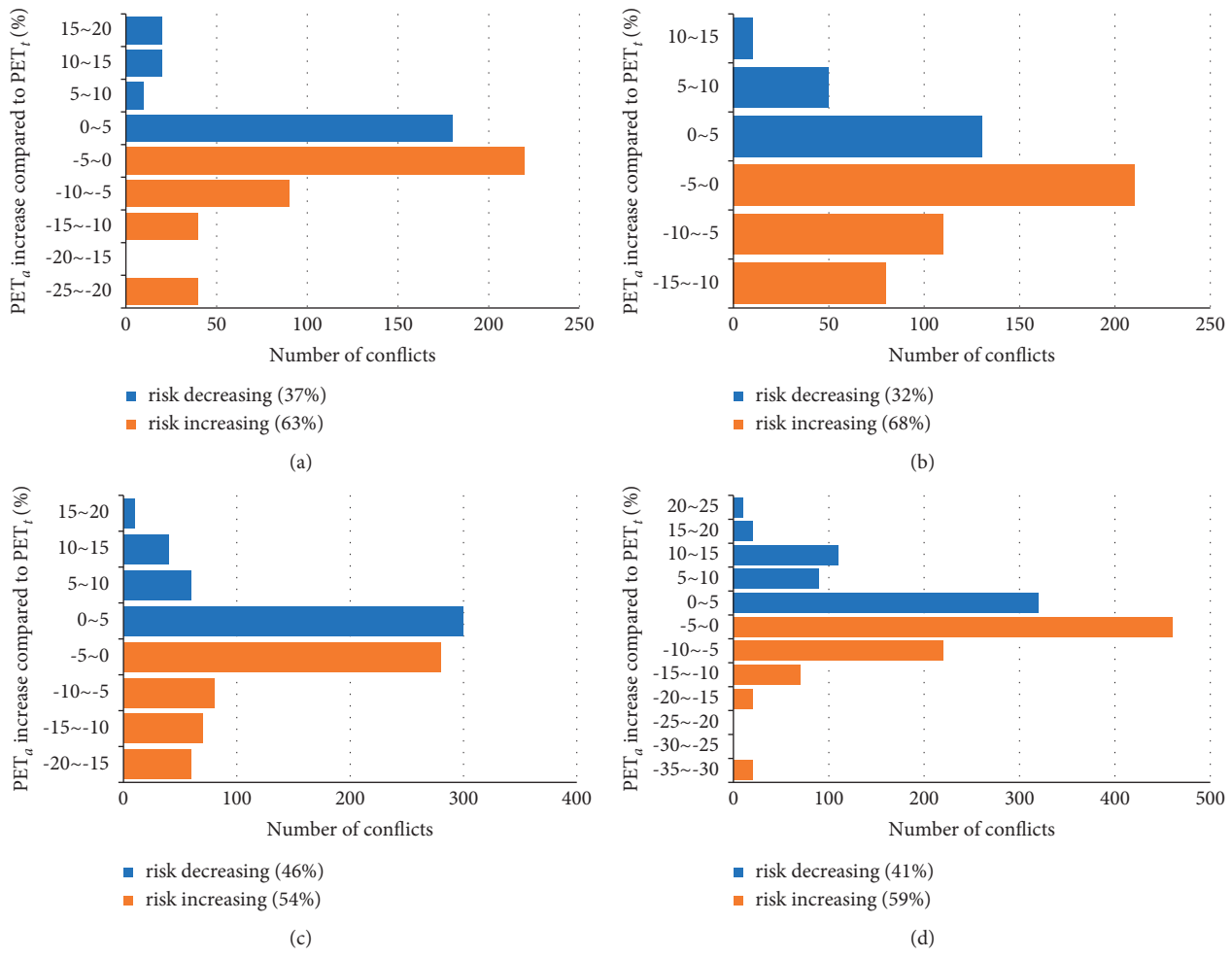
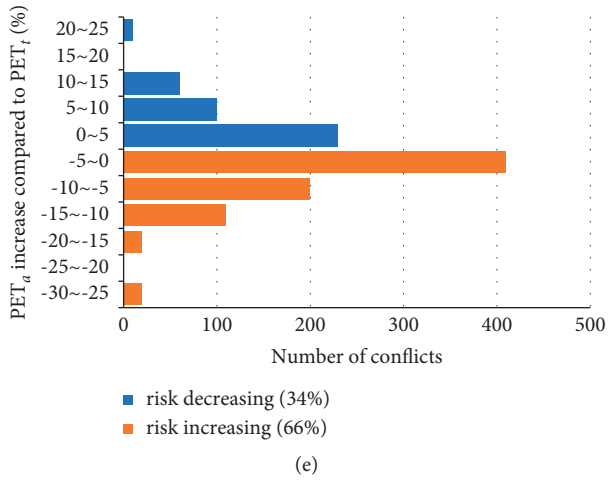


Figure 8: Continued.



(e)

Figure 8: Conflict risk changes due to left-turn path divergence. (a) 500 veh/h/lane. (b) 600 veh/h/lane. (c) 700 veh/h/lane. (d) 800 veh/h/lane. (e) 900 veh/h/lane.

Table 3: Estimated delay increase due to left-turn path divergence.

Traffic demand (veh/h/lane)	500	600	700	800	900
Total increased delay (sec)	16.19	19.41	26.18	34.52	40.06
Decreased intersection throughput (veh/h)	31.66	45.54	71.67	108.01	141.02

vehicles and adjusting vehicle states to avoid collision. When the actual PET is smaller than the theoretical one assumed by the IM due to the path divergence, CAVs might take risk-averse behaviors such as decelerating due to the increased conflict risk, leading to the drop in travelling speed as well as the intersection efficiency. When the actual PET is larger than the theoretical one, however, CAVs usually follow the schedules while maintaining their safety, and hence, travelling efficiency would not increase. The total delay increase caused by the path divergence is estimated and demonstrated in Table 3, as well as the intersection throughput decline translated from the difference between PET_a and PET_t. As observed, the delay increment and the throughput decrement both increase with the increase in traffic demand, indicating that the higher the traffic demand is, the larger effect the path divergence has on the intersection operation efficiency.

5. Conclusions and Future Directions

The variations in left-turn paths are usually neglected in previous research, which, however, might affect the safety and efficiency of CIM implementation. We conducted the initial study to analyse the effects by the simulation-based evaluation, including two approaches: the scenario-based approach of different left-turn paths and the PET-based approach with path divergence. Six typical left-turn paths are selected to represent the path variation interval, and the corresponding parameter space of mathematical models is discretized for simulations. In the former approach, the average performance and robustness of schedules obtained with different left-turn paths as nominal paths are evaluated and based on these indicators, using Path $E2$ as the nominal path is the best choice in terms of the optimality and robustness of solutions. On this basis, further simulation is conducted with Path $E2$ as the nominal path, and PETs of conflicts are calculated, respectively, with nominal path and actual path to evaluate the effects of path divergence on intersection safety and travelling efficiency. The results show that the number of serious conflicts increases significantly, as well as the overall conflict risk and vehicle delay due to the path divergence. The safety and efficiency deteriorations indicate that even with nominal path of highest robustness, the variations in left-turn paths still need more consideration in the current CIM methods.

The variation in vehicle paths might also affect the safety and efficiency of traffic infrastructure operation where both conflicting movements and turning movements exist, such as the merging roadway scenario. The CIM problem and the merging control problem are very similar in nature, and most of the approaches proposed for CIM can be easily adapted for merging coordination. The similar issue caused by the variation in turning paths of on-ramp vehicles should also be investigated, where the simulation-based evaluation method proposed in this research can be applied.

Work in progress is to further develop the proposed approach by modelling the continuous variation of vehicle paths and paths of irregular graphics to better understand the effects of real vehicle paths on the feasibility of current CIM algorithms. Besides, efforts are also put into combining vehicle collision-avoidance planning of actual turn paths and CIM optimal schedules.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the projects of the National Natural Science Foundation of China (nos. 61803083 and 51925801), "the Fundamental Research Funds for the Central Universities," and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX17_0140).

References

[1] K. Dresner and P. Stone, "A multiagent approach to autonomous intersection management," *Journal of Artificial Intelligence Research*, vol. 31, pp. 591–656, 2008.

[2] J. Lee and B. Park, "Development and evaluation of a cooperative vehicle intersection control algorithm under the connected vehicles environment," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 81–90, 2012.

[3] J. Lee, B. Park, K. Malakorn, and J. So, "Sustainability assessments of cooperative vehicle intersection control at an urban corridor," *Transportation Research Part C: Emerging Technologies*, vol. 32, pp. 193–206, 2013.

[4] Q. Jin, G. Wu, K. Boriboonsomsin, and M. Barth, "Platoon-based multi-agent intersection management for connected vehicle," in *Proceedings of the IEEE Conf. Intell. Transp. Syst. Proceedings, ITSC*, pp. 1462–1467, IEEE, The Hague, Netherlands, 6 October 2013.

[5] M. W. Levin and D. Rey, "Conflict-point formulation of intersection control for autonomous vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 85, no. September, pp. 528–547, 2017.

[6] S. A. Fayazi and A. Vahidi, "Mixed-integer linear programming for optimal scheduling of autonomous vehicle intersection crossing," *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 3, pp. 287–299, 2018.

[7] C. Yu, W. Sun, H. X. Liu, and X. Yang, "Managing connected and automated vehicles at isolated intersections: from reservation- to optimization-based methods," *Transportation Research Part B: Methodological*, vol. 122, pp. 416–435, 2019.

[8] Z. Li, Q. Wu, H. Yu et al., "Temporal-spatial dimension extension-based intersection control formulation for connected and autonomous vehicle systems," *Transportation Research Part C: Emerging Technologies*, vol. 104, no. April, pp. 234–248, 2019.

[9] H. Yu, P. Liu, Y. Fan, and G. Zhang, "Developing a decentralized signal control strategy considering link storage capacity," *Transportation Research Part C: Emerging Technologies*, vol. 124, Article ID 102971, 2021.

[10] W. Li, Y. Li, H. Deng, and L. Bao, "Planning of electric public transport system under battery swap mode," *Sustainability*, vol. 10, no. 7, pp. 2528–7, 2018.

[11] A. A. Malikopoulos, C. G. Cassandras, and Y. J. Zhang, "A decentralized energy-optimal control framework for connected automated vehicles at signal-free intersections," *Automatica*, vol. 93, pp. 244–256, 2018.

[12] Y. Wu, H. Chen, and F. Zhu, "DCL-AIM: decentralized coordination learning of autonomous intersection management for connected and automated vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 103, no. April, pp. 246–260, 2019.

[13] Y. Zhang and C. G. Cassandras, "Decentralized optimal control of Connected Automated Vehicles at signal-free intersections including comfort-constrained turns and safety guarantees," *Automatica*, vol. 109, Article ID 108563, 2019.

[14] C. Dong, H. Wang, Y. Li, W. Wang, and Z. Zhang, "Route control strategies for autonomous vehicles exiting to off-ramps," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 7, pp. 3104–3116, 2020.

[15] Y. Zheng, B. Ran, X. Qu, J. Zhang, and Y. Lin, "Cooperative lane changing strategies to improve traffic operation and safety nearby freeway off-ramps in a connected and automated vehicles environment," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 11, pp. 4605–4614, 2020.

[16] J. Wu, B. Kulcsár, S. Ahn, and X. Qu, "Emergency vehicle lane pre-clearing: from microscopic cooperation to routing decision making," *Transportation Research Part B: Methodological*, vol. 141, pp. 223–239, 2020.

[17] J. Wu, S. Ahn, Y. Zhou, P. Liu, and X. Qu, "The cooperative sorting strategy for connected and automated vehicle platoons," *Transportation Research Part C: Emerging Technologies*, vol. 123, Article ID 102986, 2021.

[18] T. Urbanik et al., *Signal Timing Manual*, Transportation Research Board, Washington, D.C., 2nd edition, 2015.

[19] Y. Guo, P. Liu, Y. Wu, and J. Chen, "Evaluating how right-turn treatments affect right-turn-on-red conflicts at signalized intersections," *Journal of Transportation Safety & Security*, vol. 12, no. 3, pp. 419–440, 2020.

[20] J. Wu et al., "Stationary condition based performance analysis of the contraflow left-turn lane design considering the influence of the upstream intersection," *Transportation Research Part C: Emerging Technologies*, vol. 123, Article ID 102919, 2021.

[21] K. Fitzpatrick, M. P. Pratt, and R. Avelar, "Speeds of right-turning vehicles at signalized intersections during green or yellow phase," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2675, no. 10, pp. 503–515, May 2021.

[22] Z. Ding, C. Sun, M. Zhou, Z. Liu, and C. Wu, "Intersection vehicle turning control for fully autonomous driving scenarios," *Sensors*, vol. 21, no. 12, pp. 3995–12, 2021.

[23] W. K. M. Alhajyaseen, M. Asano, H. Nakamura, and D. M. Tan, "Stochastic approach for modeling the effects of intersection geometry on turning vehicle paths," *Transportation Research Part C: Emerging Technologies*, vol. 32, pp. 179–192, 2013.

[24] Z. Pu, Z. Li, R. Ke, X. Hua, and Y. Wang, "Evaluating the nonlinear correlation between vertical curve features and crash frequency on highways using random forests," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, no. 10, Article ID 04020115, 2020.

[25] Z. Pu, Z. Cui, S. Wang, Q. Li, and Y. Wang, "Time-aware gated recurrent unit networks for forecasting road surface friction using historical data with missing values," *IET Intelligent Transport Systems*, vol. 14, no. 4, pp. 213–219, 2020.

[26] K. Gnanasaekaran, S. Kanya, B. Suresh, and S. K. Vasudevan, "Automated radius calculation of a turn for navigation and

safety enhancement in automobiles," *Advances in Intelligent Systems and Computing*, vol. 384, no. 1, pp. 327–334, 2016.

[27] W. Li, Z. Pu, Y. Li, and X. Ban, "Characterization of ride-splitting based on observed data: a case study of Chengdu, China," *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 330–353, 2019.

[28] M. W. Dawande and J. N. Hooker, "Inference-based sensitivity analysis for mixed integer linear programming," *Operations Research*, vol. 48, 1998.

[29] Z. Jia and M. G. Ierapetritou, "Short-term scheduling under uncertainty using MILP sensitivity analysis," *Industrial & Engineering Chemistry Research*, vol. 43, no. 14, pp. 3782–3791, 2004.

[30] Z. Ma, J. Xie, X. Qi, Y. Xu, and J. Sun, "Two-dimensional simulation of turning behavior in potential conflict area of mixed-flow intersections," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 5, pp. 412–428, 2017.

[31] L. N. Peesapati, M. P. Hunter, and M. O. Rodgers, "Evaluation of postencroachment time as surrogate for opposing left-turn crashes," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2386, no. 1, pp. 42–51, Jan. 2013.

[32] L. N. Peesapati, M. P. Hunter, and M. O. Rodgers, "Can post encroachment time substitute intersection characteristics in crash prediction models?" *Journal of Safety Research*, vol. 66, pp. 205–211, 2018.

[33] L. Zheng and T. Sayed, "From univariate to bivariate extreme value models: approaches to integrate traffic conflict indicators for crash estimation," *Transportation Research Part C: Emerging Technologies*, vol. 103, no. April, pp. 211–225, 2019.

[34] C. Wang, Y. Xie, H. Huang, and P. Liu, "A review of surrogate safety measures and their applications in connected and automated vehicles safety modeling," *Accident Analysis & Prevention*, vol. 157, Article ID 106157, 2021.

[35] S. Malkhamah, M. Tight, and F. Montgomery, "The development of an automatic method of safety monitoring at Pelican crossings," *Accident Analysis & Prevention*, vol. 37, no. 5, pp. 938–946, 2005.

[36] D. Gettman and L. Head, "Surrogate safety measures from traffic simulation models final report," vol. 126, 2003 Publ. No FHWA-RD-03-050.

[37] M. C. Ferris, *MATLAB and GAMS: Interfacing Optimization and Visualization Software*, MINDS@UW, 1999, https://minds.wisconsin.edu/handle/1793/64396.

Hindawi
Journal of Advanced Transportation
Volume 2021, Article ID 1191443, 19 pages
https://doi.org/10.1155/2021/1191443

Research Article

Optimal Electric Bus Scheduling under Travel Time Uncertainty: A Robust Model and Solution Method

Mengyan Jiang,[1] **Yi Zhang**,[1,2] **and Yi Zhang**[1,3]

[1]*Center of Environmental Science and New Energy Technology, Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Shenzhen 518055, China*
[2]*Department of Automation, Tsinghua National Laboratory for Information Science and Technology (TNList), Tsinghua University, Beijing 100084, China*
[3]*Institute of Future Human Habitats, Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China*

Correspondence should be addressed to Yi Zhang; zy1214@sz.tsinghua.edu.cn

Received 22 October 2021; Revised 18 November 2021; Accepted 24 November 2021; Published 20 December 2021

Academic Editor: Wenxiang Li

With the increasing adoption of electric buses (e-buses), e-bus scheduling problem has become an essential part of transit operation planning. As e-buses have a limited battery capacity, e-bus scheduling problem aims to assign vehicles to timetabled service trips on the bus routes considering their charging demand. Affected by the dynamic operation environment, the travel time and energy consumption of the e-buses often display considerable randomness, resulting in unexpected trip start delays and battery energy shortages. In this paper, we addressed the e-bus scheduling problem under travel time uncertainty by robust optimization approaches. We consider the cardinality constrained uncertainty set to formulate a robust multidepot EVSP model considering trip time uncertainty and partial recharging. The model is developed based on the dynamic programming equations that we formulated for trip chain robustness checking. A branch-and-price (BP) algorithm is devised to generate provably high-quality solutions for large-scale instances. In the BP algorithm, an efficient label setting algorithm is developed to solve the robust resource-constrained shortest path subproblem. Comprehensive numerical experiments are conducted based on the bus routes in Shenzhen to demonstrate the effectiveness of the suggested methodology. The robustness of the schedules was evaluated through Monte Carlo simulation. The results show that the trip start delay and battery energy shortage caused by the travel time uncertainty can be effectively reduced at the expense of an increase in the operational cost. A trade-off should be made between the reduction in infeasibility rate and increase in operational cost to choose a proper uncertainty budget.

1. Introduction

In recent years, an increasing number of electric buses (e-buses) have been introduced into the public transit systems because of their environmental and social benefits such as reducing on-road pollution, energy saving, and better onboard experience [1]. However, as the driving range of the e-buses per charge is limited and recharging the battery is time-consuming, the vehicles need extra time for battery recharging during the operation. To ensure operational efficiency, a charging plan for the e-bus fleet is required which specifies the optimal time, location, and amount to charge. In the operation planning phase, the electric vehicle scheduling problem (EVSP) aims to assign timetabled trips to the e-buses while making optimal charging plans to minimize the operational cost.

Two essential parameters taken as input to the EVSP are the expected travel time and energy consumption of the bus trips which have a great impact on the reliability of the schedule. In real-life cases, getting accurate estimations of the trip travel time is difficult because of the substantial variability of traffic condition, passenger demand, and driving conditions. Consequently, the energy consumption of a trip can also deviate from the estimated value considerably. To mitigate possible variations, on the operation planning side, the operators usually introduce a buffer time between consecutive trips to absorb small delays and a safe range for the battery state of charge (SoC). However, buffer

time is usually very short and not able to protect the schedule against trip time variations effectively. If some trips experience longer travel time than the expected value, the delay will propagate along the trip chain and influence the punctuality of the consecutive trips. The original charging plan can also become infeasible as the charging time and vehicle battery SoC are also influenced by the trip travel time.

With the aim to generate robust schedules against trip travel time variation and reduce the delays that cannot be absorbed by the buffer time, the stochastic vehicle scheduling model and dynamic rescheduling strategies are proposed in the literature [2–7]. Different from the exiting studies, we aim to tackle the travel time uncertainty in the EVSP by robust optimization methods. The optimized schedule can remain feasible regarding the trip start time and vehicle battery SoC when trip time varies in the budgeted uncertainty set. The main contribution of this study is as follows:

(i) We consider the cardinality constrained uncertainty set to formulate a robust multidepot EVSP (R-MD-EVSP) model considering trip time uncertainty and partial recharging. The R-MD-EVSP model is developed based on the dynamic programming (DP) equations that we formulated for individual trip chain robustness checking.

(ii) A branch-and-price (BP) algorithm is devised where the pricing subproblem, a robust resource-constrained shortest path problem, is solved by an efficient labeling algorithm based on the DP equations.

(iii) Comprehensive numerical experiments are conducted based on the cases in Shenzhen with hundreds of trips on multiple bus routes. The robustness of the robust solutions is verified through Monte Carlo simulations.

The remainder of the paper is organized as follows. In the next section, a literature review on the EVSP is provided. Section 3 proposes the DP equations for trip chain robustness checking and the robust EVSP formulation. The BP algorithm is described in Section 4. The results of the numerical experiments are presented in Section 5. We conclude the paper in Section 6.

2. Literature Review

In the operation planning stage of public transportation, the vehicle scheduling problem (VSP) aims at allocating transit vehicles to carry out timetabled trips with the objective to minimize the total number of vehicles used. Single-depot VSP is polynomial time solvable while multidepot VSP has proven to be NP-hard by Bertossi et al. [8]. With the wide adoption of e-buses in the transit system in recent years, EVSP arises under different scenarios regarding the charging technology, charging policy, and fleet composition. Li [9] and Yang et al. [10] developed models and algorithms for the EVSP under battery swapping modes. Considering the plug-in charging mode, Liu and Ceder [11] proposed a model based on the deficit function theory aimed at minimizing the number of vehicles and chargers. Mixed fleet EVSP was considered in [12–14]. Zhang et al. [15] addressed an EVSP considering the degradation of battery and nonlinear charging process. A tailored BP algorithm was devised to solve the problem. Their computational experiments showed that the optimized schedule can reduce the cost considerably which is mainly achieved by the substantial extension of battery life. Considering nonlinear charging process and multivehicle type, Zhang et al. [16] developed an MIP model for the EVSP with linear approximation of the nonlinear charging function. An ALNS heuristic was devised to solve the problem.

The above studies assume a full charging policy, while allowing for partial charging increases the operational flexibility and also the problem complexity. Considering partial charging, Wen et al. [17] developed a multidepot EVSP and an adaptive large neighborhood search heuristic to solve instances with customized bus trips. van Kooten Niekerk et al. [18] proposed two models for single-depot scenarios with a different level of detail resembling the actual charging processes. A column generation algorithm was developed to solve the problem. Kohání and Kohánia [19] proposed a linear MIP model for a single-depot EVSP considering the location of the charging stations and the assignment of chargers to the vehicles. The model was solved by the standard solver. Li et al. [20] addressed EVSP along with the charger deployment problem. An adaptive genetic algorithm is designed to solve the problem. Yıldırım and Yıldız [21] considered an optimal fleet composition and scheduling problem. An IP-column-generation algorithm is proposed to solve the pricing subproblem.

Charging scheduling problem aims to determine the time, amount, and specific charger for each e-bus to get recharged. In large-sized transit systems, due to the problem scale and complexity, charging scheduling often assumes that the e-bus operation schedule is given beforehand. Qin et al. [22] conducted a simulation analysis on the daily charging patterns and demand charges of a fleet of e-buses. An optimal charging strategy is identified to minimize demand charges. Wang et al. [23] developed a mixed-integer programming (MIP) model for the e-bus recharging scheduling based on a real-world case. The model aims to assign the charging station, chargers, and charging time to the e-buses to minimize the system operating costs, assuming given timetable and operation schedule. With the goal to minimize the electricity demand charges and energy charges, He et al. [24] proposed a modeling framework to determine the time to charge a vehicle and the actual charging power. Liu et al. [25] considered the e-bus charging scheduling problem under limited charging resources in the charging station. A column-generation-based algorithm is developed to solve the large-scale problem instances. The results show that the optimal charging plan can greatly reduce the charging cost compared with the uncontrolled charging.

Most of the studies take deterministic trip time as the model input and assume that the energy discharge is

proportional to the travel distance. To tackle the trip time uncertainty in the real world, on the input side, some methods were proposed to improve the accuracy of the trip time estimation based on historical running data [26, 27]. On the planning side, dynamic rescheduling strategies and stochastic VSP models were developed. Huisman et al. [3] introduced a dynamic vehicle scheduling approach to solve the VSP periodically with renewed estimation on the future trip time. He et al. [5] formulated a stochastic dynamic VSP and adopted an approximate dynamic programming approach where the objective function is approximated by a feed-forward neural network. Huisman et al. [3] developed a stochastic VSP model using typical disruption scenarios to minimize the expected total planned costs and costs caused by disruptions. Shen et al. [2] proposed a probabilistic network flow model assuming certain trip time distributions and developed a hybrid heuristic solution method. Considering the stochastic travel time, Tang et al. [6] proposed a stochastic model and a dynamic rescheduling paradigm for a single-depot EVSP with the full charging policy. A speed-energy-consumption relationship was incorporated in their models. Bie et al. [7] proposed a multiobjective stochastic e-bus scheduling model considering the variability of travel time and energy consumption. The numerical study showed that the optimized schedule can effectively protect against the accumulation of stochastic volatilities.

Compared with the EVSP, more studies have considered the travel time uncertainty in the vehicle routing problem (VRP). VRP arises in the logistics context which aims to plan for vehicle routes to serve customers at different locations. To tackle the uncertainty related to customer demand and vehicle travel time, stochastic and robust models as well as dynamic reoptimization approaches were proposed in the literature. The advantage of robust optimization is that it only requires defined bounds on the data rather than the underlying distribution while maintaining the tractability of the problem. Different kinds of uncertainty sets have been considered in robust VRP (RVRP) with the typical ones including the hypercube, budgeted uncertainty sets, and ellipsoidal sets. The cardinality constrained set proposed by Bertsimas and Sim [28] is widely used which specifies an upper bound of the number of customer demand or the travel links that can attain their maximum value.

Motivated by the travel time uncertainty in maritime transportation, Agra et al. [29] proposed two RVSP formulations. Dynamic programming recursive equations were proposed for path feasibility checking based on which path inequalities can be generated. They showed that it suffices to consider a subset of the extreme points of the uncertainty polytope and developed scenario reduction solution techniques. Salicru et al. [26] proposed a compact formulation for the RVRP under demand uncertainty with dynamic programming equations integrated. To solve the RVRP and its variants, exact algorithms were developed based on the branch-price-and-cut method by reformulating the robust problem into its deterministic counterpart [30, 31]. Local search-based heuristic frameworks with solution robustness checking procedure embedded in were devised in [32, 33] to solve large-sized instances. Pelletier et al. [34] considered a capacitated electric VRP with energy consumption uncertainty. They considered different kinds of uncertainty sets and proposed an LNS-based heuristic to solve the problem.

From the above discussion, we can see that although many studies have focused on the EVSP, few of them have considered the uncertainty related to the travel time and energy consumption in the problem. To the best of our knowledge, there has been no research to cope with the trip time uncertainty of the EVSP using the robust optimization method. Our study proposed a robust EVSP model and solution method considering the travel time uncertainty under budgeted uncertainty set.

3. Robust Multidepot EVSP

3.1. Problem Description. A bus route in the transit network is defined by two end stops and a series of intermediate stops. On a bus route, a service trip starts from one end stop at a scheduled time and ends at the other end stop to carry passengers. The timetable of the route includes all the trips that should be carried out in one day with their scheduled start time from the start stop and expected end time at the end stop. On the electrified bus routes, these trips are carried out by a fleet of e-buses. Each e-bus undertakes a sequence of trips in a day which is called a trip chain. An operation schedule consists of the trip chains for all the e-buses. For the convenience of operation, bus depots are established close to the end stops of the routes for vehicle parking, maintenance, and charging if charging facilities are established. Based on this setting, the R-MD-EVSP is described as follows. The list of notations is provided in Table 1.

Let $G = (V, A)$ be an acyclic direct graph, where V is the set of nodes and A is the set of arcs. Each timetabled trip is represented by a trip node in the graph. Denote $T \subseteq V$ as the set of trip nodes. Each trip node $i \in T$ is associated with a trip start time window $[a_i, b_i]$, where a_i and b_i are the scheduled start time and latest start time, respectively. The e-bus serving a trip should wait until the scheduled start time a_i if it arrives earlier at the trip start depot. Each trip node $i \in T$ is associated with two uncertain parameters: the trip time τ_i and trip energy consumption ε_i. τ_i is realized in the interval $[t_i - \widetilde{t}_i, t_i + \widehat{t}_i]$ where t_i is the nominal trip time; $\widehat{t}_i \geq 0$ and $\widetilde{t}_i \geq 0$ are the upper and lower deviations from t_i. ε_i is realized in the interval $[e_i - \widetilde{e}_i, e_i + \widehat{e}_i]$ where e_i is the nominal value; $\widehat{e}_i \geq 0$ and $\widetilde{e}_i \geq 0$ are the upper and lower deviations from e_i. Since the worst case will always be achieved at the right-hand side of the interval, we only consider the realization of τ_i in $[t_i, t_i + \widehat{t}_i]$ and ε_i in $[e_i, e_i + \widehat{e}_i]$ correspondingly.

Denote K as the set of depots. Each depot $k \in K$ is created with two nodes in the graph: operation start node o^k and operation end node d^k, indicating that the e-bus begins/ends the operation from/at depot k, respectively. For nodes o^k and d^k, $k \in K$, we assign a time window $[a_0, b_0]$, where a_o is the earliest operation start time and b_0 is the latest operation end time. This means that the e-buses can begin and end their operation at any time within time window $[a_0, b_0]$.

The arc set A includes three kinds of arcs: (i) the pull-out arc connects a depot node o^k, $k \in K$, and a trip node $i \in T$, representing that a vehicle begins operation from depot k to

Table 1: Definitions of the sets and parameters for the R-MD-EVSP model.

Sets	
K	Set of depots with index k
V_t	Set of trip nodes
V	Set of nodes
A	Set of arcs
Parameters	
a_0	Earliest operation start time
b_0	Latest operation end time
a_i	Scheduled departure time of trip node i
b_i	Latest departure time of trip node i
τ_i	Trip time of trip node i, an uncertain parameter
t_i	Nominal trip time of trip node i
$\widehat{t}_i$	Upper trip time deviation from t_i
$\widetilde{t}_i$	Lower trip time deviation from t_i
δ_i^t	Trip time deviation of trip node i, $\delta_i^t \in [0, \widehat{t}_i]$
ε_i	Trip energy consumption of trip node i, an uncertain parameter
e_i	Nominal energy consumption of trip node i
$\widehat{e}_i$	Upper energy consumption deviation from e_i
$\widetilde{e}_i$	Lower energy consumption deviation from e_i
δ_i^e	Energy consumption deviation of trip node i, $\delta_i^e \in [0, \widehat{e}_i]$
s_{ij}	Empty travel time from node i to node j
e_{ij}	Energy consumption of the empty travel from node i to node j
c_{ij}	Cost of arc $(i, j) \in A$
c^d	Departure delay cost per minute
c^e	Unit charging cost
t^p	Charging preparation time
r_s	Energy charging rate
r_u	Energy consumption rate
Γ^t	Uncertainty budget, maximum number of trips that can take their upper bound trip time, indexed by γ
u^{max}	Upper bound of the vehicle battery SoC
u^{min}	Lower bound of the vehicle battery SoC
$u_{i\gamma}^{min}$	Lower bound of variable $Y_{i\gamma}$, $\gamma \leq \Gamma^t$

carry out the first trip i; (ii) the pull-in arc connects a trip node $i \in T$ and an end node d^k, $k \in K$, representing that a vehicle finishes its last trip i and ends the operation, returning to depot k; and (iii) the trip connection arc connects two trip nodes, representing that two trips are carried out consecutively. Each arc $(i, j) \in A$ is associated with a nonservice travel time s_{ij} and energy consumption e_{ij}. To ensure that the trip start time window is respected, trip nodes i and j, $i, j \in T$, are connected only if the time compatible condition $a_i + t_i + s_{ij} \leq b_j$ is satisfied. The cost of an arc c_{ij} includes the empty travel cost and the vehicle usage cost f for the pull-out arcs. Denote g_{ij}, $(i, j) \in A$, as the unit charging cost in the time interval between the end of trip node i and the start of trip node j; g_{ij} is calculated as the weighted average value of the time-of-use tariff in the time interval.

Figure 1 gives an example of the graph for an instance with two depots and four trips. Each depot is created with a start node o^k and an end node d^k, $k = 1, 2$. Each trip is represented by a trip node v_i associated with a departure time window and a scheduled trip time, $i = 1, 2, 3, 4$.

We assume that the e-buses begin the operation with a fully charged battery and get recharged at the destination depot after finishing a trip if needed. The charging amount is proportional to the charging time with a rate of r_s (kWh/min). Denote t^s as the unit charging time and t^p as the charging preparation time. The e-buses should charge integer times of t^s.

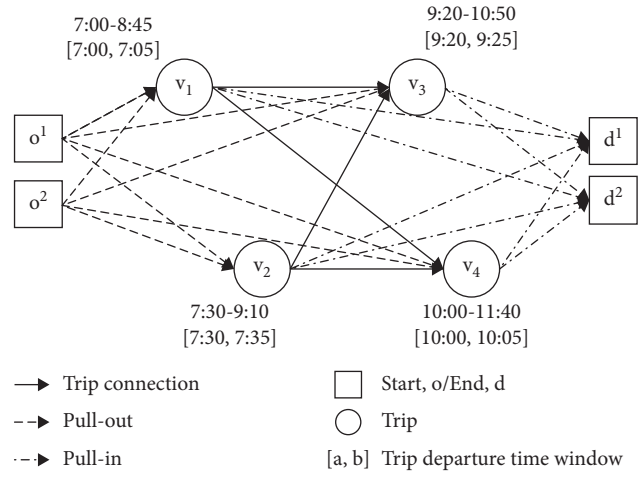


Figure 1: An illustrative example of the graph for the robust multidepot EVSP.

The historical running data of 150 bus routes in five months in Shenzhen were used to analyze the relationship between the trip time and trip energy consumption. Most of them display similar relationship. Figure 2 shows the

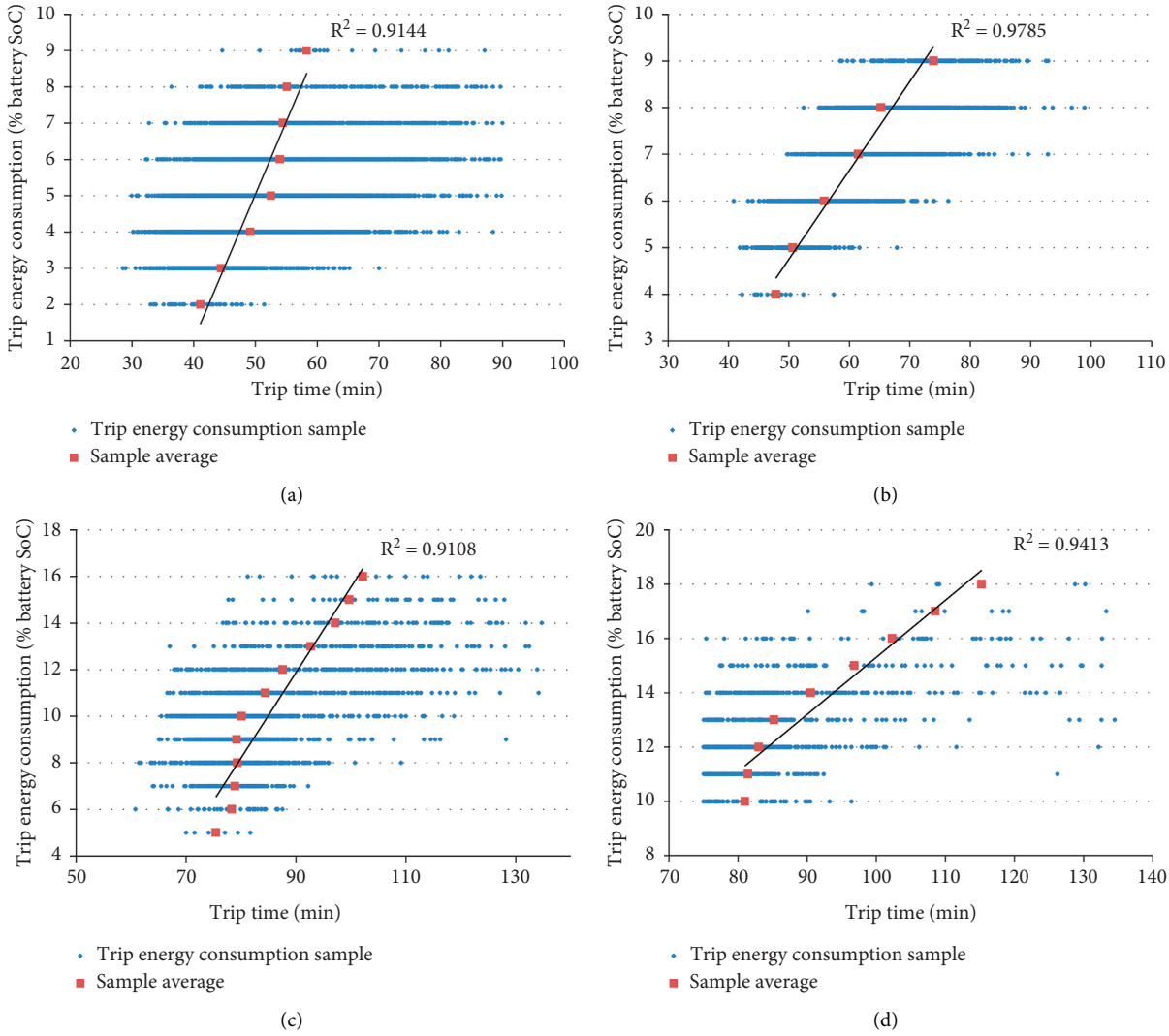


Figure 2: Relationship between the trip time and energy consumption for the selected bus routes. (a) Route 1. (b) Route 2. (c) Route 3. (d) Route 4.

relationship between the average trip time and energy consumption of the selected four bus routes. The solid lines are the linear regression results of the sample average data with the R-square values given next to it. The R-square values are both above 0.9, meaning that the trip energy consumption can be described as a linear function of the trip time, although the data of Route 3 and Route 4 display some nonlinear trends. The trip time and trip energy consumption may display different relationships in different cities and routes. Based on the cases of Shenzhen, we assume that for a trip i, the energy consumption deviation $\delta_i^e \in [0, \widehat{e}_i]$ is proportional to the trip time deviation $\delta_i^t \in [0, \widehat{t}_i]$, i.e., $\delta_i^e = \alpha \delta_i^t r_u$ where $\alpha \geq 1$ is an augment constant and r_u is the energy consumption rate (kWh/min). Note that the robust EVSP model developed in this study does not depend on the assumption of a linear relationship between trip time and trip energy consumption. We assume that the energy consumption of a trip i varies within interval $[e_i, e_i + \widehat{e}_i]$; as such, we only need to give the values of e_i and $\widehat{e}_i$ as the model inputs. Therefore, any empirical method that can give estimation on the values of e_i and $\widehat{e}_i$ works.

We adopt a cardinality constrained uncertainty set U as discussed in [24] to control the level of trip time uncertainty. U is defined as $U = \left\{\tau \in \mathbb{R}^{|T|} | \tau_i = t_i + \theta_i \widehat{t}_i, 0 \leq \theta_i \leq 1, \forall i \in T, \sum_{i \in T} \theta_i \leq \Gamma^t\right\}$ where the cumulative uncertainty of the random variable θ_i is bounded by the budget Γ^t. We choose this uncertainty set because it is commonly used in the RVRP, the solution structure of which is similar to our problem: the trip chain assigned to an e-bus is equivalent to the customer route assigned to a vehicle in the VRP, where a bus trip can be regarded as a customer.

Based on the above settings, the problem looks for the minimum cost schedule that satisfies the following constraints: (i) each trip is carried out exactly once; (ii) the departure time window of each trip is respected when the trip time varies within the budgeted set U; (iii) each vehicle begins and ends its operation at the same depot; and (iv) the vehicle battery SoC is always kept within the safety range $[u^{\min}, u^{\max}]$ when the trip

time varies within the budgeted set U. Following the robust optimization paradigm, a trip chain is robust feasible if it satisfies conditions (ii)–(iv) for all possible trip time realizations. A feasible schedule of the problem is given as set of feasible trip chains that together also satisfies condition (i).

3.2. Deterministic Problem Formulation. Let X_{ijk} be the binary variables that take value 1 if a vehicle housed at terminal k traverses arc $(i, j) \in A$. At each trip node $i \in T$, denote Z_i as the trip start time and Y_i as the lowest battery SoC of the vehicle at the trip start time. The lower bound of Y_i is denoted as $u_i^{min} = u^{min} + e_i$. At each operation start (end) node o^k (d^k), $k \in K$, denote Z_i as the operation start (end) time and Y_i as the lowest battery SoC of the vehicle at the operation start (end) time. On each arc $(i, j) \in A$, $i, j \in T$, denote W_{ij} as the amount of energy to be charged after completing trip i before starting trip j. According to our assumption, the charging place is at the ending depot of trip node i. R_{ij} are the auxiliary variables defined corresponding to Z_i for constraint linearization. Denote E_{ij} as the binary variables that are equal to 1 if the e-buses can be charged on arc $(i, j) \in A$. Denote U_{ij} as the integer variables that specify the number of time units an e-bus spends charging on arc $(i, j) \in A$. The deterministic problem is formulated as follows (MD-EVSP):

$$\min \sum_{k\in K}\sum_{(i,j)\in A} c_{ij}X_{ijk} + \sum_{(i,j)\in A} g_{ij}W_{ij}, \tag{1}$$

$$\text{s.t.} \sum_{k\in K}\sum_{j:\ (i,j)\in A} X_{ijk} = 1, \quad i \in T, \tag{2}$$

$$\sum_{j:\ (i,j)\in A} X_{ijk} - \sum_{j:\ (j,i)\in A} X_{jik} = 0, \quad i \in T,\ k \in K, \tag{3}$$

$$\sum_{\beta\in K\setminus\{k\}}\sum_{i\in V} X_{o^\beta ik} = 0, \quad k \in K, \tag{4}$$

$$\sum_{\beta\in K\setminus\{k\}}\sum_{i\in V} X_{id^\beta k} = 0, \quad k \in K, \tag{5}$$

$$a_i \le Z_i \le b_i, \quad i \in V, \tag{6}$$

$$Z_j - Z_i - t_i - s_{ij} + M\left(1 - \sum_{k\in K} X_{ijk}\right) \ge 0,\ (i, j) \in A, \tag{7}$$

$$Y_i - e_i + W_{ij} - e_{ij} + M\left(1 - \sum_{k\in K} X_{ijk}\right) \ge Y_j, \quad (i, j) \in A, \tag{8}$$

$$E_{ij} - \sum_{k\in K} X_{ijk} \le 0,\ (i, j) \in A, \tag{9}$$

$$-\left(a_j - Z_i - t_i - s_{ij} - t^p\right) - M\left(1 - E_{ij}\right) \le 0,\ (i, j) \in A, \tag{10}$$

$$W_{ij} + R_{ij}r_s - \left(a_j - t_i - s_{ij} - t^p\right)r_s E_{ij} \le 0,\ (i, j) \in A, \tag{11}$$

$$W_{ij} + Y_i - e_i - u^{max} \le 0,\ (i, j) \in A, \tag{12}$$

$$W_{ij} - r_s t^s U_{ij} = 0,\ (i, j) \in A, \tag{13}$$

$$R_{ij} - b_i E_{ij} \le 0,\ (i, j) \in A, \tag{14}$$

$$a_i E_{ij} - R_{ij} \le 0,\ (i, j) \in A, \tag{15}$$

$$R_{ij} - Z_i + a_i\left(1 - E_{ij}\right) \le 0,\ (i, j) \in A, \tag{16}$$

$$Z_i - b_i\left(1 - E_{ij}\right) - R_{ij} \le 0,\ (i, j) \in A, \tag{17}$$

$$u_i^{\min} \le Y_i \le u^{\max}, \quad i \in V, \tag{18}$$

$$W_{ij} \ge 0,\ (i, j) \in A, \tag{19}$$

$$X_{ijk} \in \{0, 1\}, \quad (i, j) \in A,\ k \in K, \tag{20}$$

$$E_{ij} \in \{0, 1\}, \quad (i, j) \in A, \tag{21}$$

$$R_{ij} \ge 0, \quad (i, j) \in A, \tag{22}$$

$$U_{ij} \in Z, \quad (i, j) \in A. \tag{23}$$

Objective (1) minimizes the sum of vehicle usage, empty travel, and charging cost. Constraint (2) ensures that each trip is carried out only once. Constraint (3) is the flow conservation constraint. Constraints (4) and (5) require that each vehicle begin and end its operation at its base depot. Constraint (6) limits the departure time of each trip to be within the departure time window. Constraint (7) ensures the time consistency of two successive trips. Constraint (8) ensures the vehicle battery energy consistency of two successive trips. Constraint (9) requires that charging on arc (i, j) is performed only if trip j is succeeded by trip i. Constraint (10) restricts that charging on arc (i, j) can be performed only if time is available. Constraint (11) defines the upper bound of W_{ij} on arc $(i, j) \in A$ according to the time available. Originally, constraint (11) is formulated as follows:

$$W_{ij} + \left(Z_i - a_j - t_i - s_{ij} - t^p\right)r_s E_{ij} \le 0,\ (i, j) \in A. \tag{24}$$

To linearize quadratic terms $Z_i E_{ij}$, auxiliary variables R_{ij} are introduced where $R_{ij} = Z_i E_{ij}$ is satisfied by constraints (14)–(17). Constraint (12) requires the battery SoC after charging not to exceed $u^{\max}$. Constraint (13) ensures that the charging time can only be integer times of the minimum charging time unit. Constraint (18) keeps the vehicle battery SoC within the safety range. Constraints (19)–(23) define the domains of the variables.

3.3. DP Equations to Define the Robustness of a Trip Chain. In this section, we introduce the DP equations to define the robustness of a trip chain under uncertainty set U. In graph

G, a trip chain is represented by a path $P = \{o^k, v_1, v_2, \ldots, v_n, d^k\}$ starting from an operation start node o^k to an operation end node d^k, visiting a sequence of trip nodes $v_i, i = 1, 2, \ldots, n$. In solving the RVRP, Agra et al. [29] showed that only the travel time vector $\tau \in ext(U)$ needs to be considered in the problem formulation to ensure the robustness feasibility. $ext(U)$ is the set that contains all the extreme points of set U. Because of the structure of set U, they defined DP recursive equations to define the robustness of a vehicle route. As our R-MD-EVSP problem shares a similar solution structure with the RVRP, we can define the robustness of a trip chain by the DP recursive equations (25) and (26) as follows.

Denote $Z_{i\gamma}$ as the earliest possible start time of node v_i and $Y_{i\gamma}$ as the lowest battery SoC of the vehicle upon the start of v_i; let $W_{ij\gamma}$ and $\widetilde{W}_{ij\gamma}$ be the charging amount on arc (i, j), i.e., after the end of trip node v_i and before the start of trip node v_j, when $\gamma \le \Gamma^t$ trips from v_1 to v_i are taking their maximum trip time, among which trip node v_i takes the nominal trip time and maximum trip time, respectively. Denote $u_{i\gamma}^{min}$ as the lower bound of $Y_{i\gamma}$: if $\gamma < \Gamma^t$, $u_{i\gamma}^{min} = u^{min} + e_i + \hat{e}_i$; otherwise, $u_{i\gamma}^{min} = u^{min} + e_i$.

$Z_{i\gamma}$ and $Y_{i\gamma}$ are defined by the recursive functions (25) and (26), respectively. A trip chain is robust if $Z_{i\gamma} \in [a_i, b_i]$ and $Y_{i\gamma} \in [u_{i\gamma}^{\min}, u^{\max}]$, for $\forall v_i \in P$ and $\gamma \le \Gamma^t$. In fact, it suffices to check $Z_{i\Gamma^t}$ and $Y_{i\Gamma^t}$ for $v_i \in P$.

$$Z_{j\gamma} = \begin{cases} a_0 \\ \max\left\{a_j, Z_{i\gamma} + t_i + t_{ij}\right\} \\ \max\left\{a_j, Z_{i\gamma} + t_i + \dfrac{W_{ij\gamma}}{r_s} + t_{ij}, Z_{i,\gamma-1} + t_i + \hat{t}_i + \dfrac{\widetilde{W}_{ij\gamma}}{r_s} + t_{ij}\right\} \\ j = 0 \\ \gamma = 0 \\ \text{otherwise} \end{cases}, \tag{25}$$

$$Y_{j\gamma} = \begin{cases} u^{\max} & j = o \\ \min\left\{u^{\max}, Y_{i\gamma} - e_i + W_{ij\gamma}\right\} - e_{ij} & \gamma = 0 \\ \min\left\{\min\left\{u^{\max}, Y_{i\gamma} - e_i + W_{ij\gamma}\right\} - e_{ij}, \right. & \text{otherwise} \\ \left. \min\left\{u^{\max}, Y_{i,\gamma-1} - e_i - \hat{e}_i + \widetilde{W}_{ij\gamma}\right\} - e_{ij}\right\} & \end{cases} \tag{26}$$

3.3.1. The Optimal Total Charging Amount and Charging Policy. Denote H as the optimal total charging amount of a trip chain. H is the maximum total charging amount required when the trip times varies within uncertainty set U which is calculated as follows: let $b = \min\{n, \Gamma^t\}$ be the maximum number of trips of path P that can attain their maximum trip time. Choose the b trips with the maximum worst case trip time and take their maximum trip times. The rest of the trips take their nominal trip times. Then, H is given by $H = u^{\min} + E - u^{\max}$ where E is the total energy consumption of P. The charging policy for an e-bus carrying out trip chain P is as follows: the e-buses can get recharged during the layover time between two successive trips for any flexible amount that is integer time of the unit charging time t^s until the total charging amount reaches H.

3.4. Robust Problem Formulation. We introduce an MIP formulation for the R-MD-EVSP under cardinality uncertainty set U by incorporating the DP equations for trip chain robustness checking into the deterministic model.

At each node $i \in V$, given γ trips have taken their maximum trip time, $\gamma = 0, 1, 2, \ldots, \Gamma^t$, variables $Z_{i\gamma}$, $Y_{i\gamma}$, $W_{ij\gamma}$, $\widetilde{W}_{ij\gamma}$ and parameter $u_{i\gamma}^{min}$ are consistent with those defined in Section 3.3. Given that γ trips on the trip chain from the first trip to trip j (trip j not included) are taking their maximum trip time, denote $E_{ij\gamma}$ and $\widetilde{E}_{ij\gamma}$ as the binary variables that are equal to 1 if vehicle charging can be carried out on arc $(i, j) \in A$ when trip i takes its nominal and maximum trip time, respectively; denote $U_{ij\gamma}$ and $\widetilde{U}_{ij\gamma}$ as the integer variables that specify the number of time units an e-bus spends charging on arc $(i, j) \in A$ when trip i takes its nominal and maximum trip time, respectively. Denote $R_{ij\gamma}$ and $\widetilde{R}_{ij\gamma}$ as the auxiliary variables defined to linearize the quadratic terms in the constraints.

In the robust model, the robustness of a trip chain is ensured by linearizing of the DP equations (25) and (26) and adding them into the problem constraints. The following constraints ((27)–(30)) are the linearized forms of the DP equations (25) and (26).

(i) Time consistency constraints:

$$Z_{j\gamma} - Z_{i\gamma} - t_i - s_{ij} + M\left(1 - \sum_{k\in K} X_{ijk}\right) \geq 0, \quad (i,j) \in A_p,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{27}$$

$$Z_{j\gamma} - Z_{i,\gamma-1} - t_i - \widehat{t}_i - s_{ij} + M\left(1 - \sum_{k\in K} X_{ijk}\right) \geq 0, \quad (i,j) \in A_p,\ \gamma = 1,2,\ldots,\Gamma^t. \tag{28}$$

Constraints (27) and (28) are the robust constraints corresponding to constraint (7). It is developed based on the linearization of equation (25), which means the time consistency of two successive trips should be satisfied for $\forall\gamma = 0,1,2,\ldots,\Gamma^t$.

(ii) Battery SoC consistency constraints:

$$Y_{i\gamma} - e_i + W_{ij\gamma} - e_{ij} + M\left(1 - \sum_{k\in K} X_{ijk}\right) \geq Y_{j\gamma}, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{29}$$

$$Y_{i,\gamma-1} - e_i - \widehat{e}_i + \widetilde{W}_{ij\gamma} - e_{ij} + M\left(1 - \sum_{k\in K} X_{ijk}\right) \geq Y_{j\gamma}, \quad (i,j) \in A,\ \gamma = 1,2,\ldots,\Gamma^t. \tag{30}$$

Constraints (29) and (30) are the robust constraints corresponding to constraint (8). It is developed based on the linearization of equation (26), which means the battery SoC consistency of two successive trips should be satisfied for $\forall\gamma = 0,1,2,\ldots,\Gamma^t$.

Other constraints are defined corresponding to the deterministic models with the variables substituted by the ones defined in the robust model. The R-MD-EVSP model is formulated as follows:

$$\min \sum_{k\in K}\sum_{(i,j)\in A} c_{ij}X_{ijk} + \sum_{\gamma=0}^{\Gamma^t}\sum_{(i,j)\in A} g_{ij}W_{ij\gamma} + \sum_{\gamma=1}^{\Gamma^t}\sum_{(i,j)\in A} g_{ij}\widetilde{W}_{ij\gamma}, \tag{31}$$

s.t. constraints (2)–(5) and constraints (27)–(30).

$$a_i \leq Z_{i\gamma} \leq b_i, \quad i \in V,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{32}$$

$$E_{ij\gamma} - \sum_{k\in K} X_{ijk} \leq 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{33}$$

$$\widetilde{E}_{ij\gamma} - \sum_{k\in K} X_{ijk} \leq 0, \quad (i,j) \in A,\ \gamma = 1,2,\ldots,\Gamma^t, \tag{34}$$

$$-\left(a_j - Z_{i\gamma} - t_i - s_{ij} - t^p\right) - M\left(1 - E_{ij\gamma}\right) \leq 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{35}$$

$$-\left(a_j - Z_{i,\gamma-1} - t_i - \widehat{t}_i - s_{ij} - t^p\right) - M\left(1 - \widetilde{E}_{ij\gamma}\right) \leq 0, \quad (i,j) \in A,\ \gamma = 1,2,\ldots,\Gamma^t, \tag{36}$$

$$W_{ij\gamma} + R_{ij\gamma}r_s - \left(a_j - t_i - s_{ij} - t^p\right)r_s E_{ij\gamma} \leq 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{37}$$

$$\widetilde{W}_{ij\gamma} + \widetilde{R}_{ij\gamma}r_s - \left(a_j - t_i - \widehat{t}_i - s_{ij} - t^p\right)r_s \widetilde{E}_{ij\gamma} \leq 0, \quad (i,j) \in A,\ \gamma = 1,2,\ldots,\Gamma^t, \tag{38}$$

$$W_{ij\gamma} + Y_{i\gamma} - e_i - u^{\max} \leq 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{39}$$

$$\widetilde{W}_{ij\gamma} + Y_{i,\gamma-1} - e_i - \widehat{e}_i - u^{\max} \leq 0, \quad (i,j) \in A,\ \gamma = 1,2,\ldots,\Gamma^t, \tag{40}$$

$$W_{ij} - r_s t^s U_{ij\gamma} = 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{41}$$

$$W_{ij} - r_s t^s \widetilde{U}_{ij\gamma} = 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{42}$$

$$R_{ij\gamma} - b_i E_{ij\gamma} \le 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{43}$$

$$a_i E_{ij\gamma} - R_{ij\gamma} \le 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{44}$$

$$R_{ij\gamma} - Z_{i\gamma} + a_i\left(1 - E_{ij\gamma}\right) \le 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{45}$$

$$Z_{i\gamma} - b_i\left(1 - E_{ij\gamma}\right) - R_{ij\gamma} \le 0, \quad (i,j) \in A_p,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{46}$$

$$\widetilde{R}_{ij\gamma} - b_i \widetilde{E}_{ij\gamma} \le 0, \quad (i,j) \in A,\ \gamma = 1,2,\ldots,\Gamma^t, \tag{47}$$

$$a_i \widetilde{E}_{ij\gamma} - \widetilde{R}_{ij\gamma} \le 0, \quad (i,j) \in A,\ \gamma = 1,2,\ldots,\Gamma^t, \tag{48}$$

$$\widetilde{R}_{ij\gamma} - Z_{i,\gamma-1} + a_i\left(1 - \widetilde{E}_{ij\gamma}\right) \le 0, \quad (i,j) \in A,\ \gamma = 1,2,\ldots,\Gamma^t, \tag{49}$$

$$Z_{i,\gamma-1} - b_i\left(1 - \widetilde{E}_{ij\gamma}\right) - \widetilde{R}_{ij\gamma} \le 0, \quad (i,j) \in A,\ \gamma = 1,2,\ldots,\Gamma^t, \tag{50}$$

$$u_{i\gamma}^{\min} \le Y_{i\gamma} \le u^{\max}, \quad i \in V,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{51}$$

$$W_{ij\gamma} \ge 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{52}$$

$$\widetilde{W}_{ij\gamma} \ge 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{53}$$

$$X_{ijk} \in \{0,1\}, \quad (i,j) \in A,\ k \in K, \tag{54}$$

$$E_{ij\gamma} \in \{0,1\}, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{55}$$

$$R_{ij\gamma} \ge 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{56}$$

$$U_{ij\gamma} \in Z, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{57}$$

$$\widetilde{E}_{ij\gamma} \in \{0,1\}, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{58}$$

$$\widetilde{R}_{ij\gamma} \ge 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{59}$$

$$\widetilde{U}_{ij\gamma} \in Z, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t. \tag{60}$$

Objective (31) consists of the vehicle usage, empty travel, and the sum of charging cost in all possible cases regarding the realizations of the trip times. The worst-case charging cost cannot be expressed explicitly; therefore, it is not included in the objective but calculated after the optimal solution is obtained. Constraint (32) limits the departure time of each trip to be within the departure time window. Constraints (33) and (34) require that charging on arc (i, j) is performed only if trip j is succeeded by trip i. Constraints (35) and (36) restrict that charging on arc (i, j) can be performed if time is available. Constraints (37) and (38) define the upper bound of W_{ij} on arc $(i, j) \in A$ according to the available time. Originally, constraints (37) and (38) are formulated as follows:

$$W_{ij\gamma} + \left(Z_{i\gamma} - a_j - t_i - s_{ij} - t^p\right) r_s E_{ij\gamma} \le 0, \quad (i,j) \in A,\ \gamma = 0,1,2,\ldots,\Gamma^t, \tag{61}$$

$$\widetilde{W}_{ij\gamma} + \left(Z_{i,\gamma-1} - a_j - t_i - \widehat{t}_i - s_{ij} - t^p\right) r_s \widetilde{E}_{ij\gamma} \le 0, \quad (i,j) \in A,\ \gamma = 1,2,\ldots,\Gamma^t. \tag{62}$$

To linearize quadratic terms $Z_{i\gamma}E_{ij\gamma}$ and $Z_{i,\gamma-1}\widetilde{E}_{ij\gamma}$, auxiliary variables $R_{ij\gamma}$ and $\widetilde{R}_{ij\gamma}$ are introduced where $R_{ij\gamma} = Z_{i\gamma}E_{ij\gamma}$ is satisfied by constraints (43)–(46) and $\widetilde{R}_{ij\gamma} = Z_{i,\gamma-1}\widetilde{E}_{ij\gamma}$ is satisfied by constraints (47)–(50).

Constraints (39) and (40) require the battery SoC after charging not to exceed u^{max}. Constraints (41) and (42) ensure that the charging time can only be integer times of the minimum charging time unit. Constraint (49) keeps the vehicle battery SoC within the safety range. Constraints (52)–(60) define the domains of the variables.

4. Branch-and-Price Method

In this section, we introduce a BP algorithm to solve the R-MD-EVSP based on the set partitioning formulation. Let Ω be the set of feasible trip chains. Each trip chain $p \in \Omega$ is defined by c_p, the total operation cost, and $a_{pi}, i \in T$, a binary parameter that is equal to 1 if trip i is carried out in p. Denote θ_p, $p \in \Omega$, as a binary variable that takes value 1 if trip chain p is selected to be part of the solution. The set partitioning formulation of the R-MD-EVSP is as follows:

$$\min \sum_{p\in\Omega} c_p\theta_p, \tag{63}$$

$$\text{s.t.} \sum_{k\in K} a_{pi}\theta_p = 1, \quad i \in T, \tag{64}$$

$$\theta_p \in \{0,1\}, \quad p \in \Omega. \tag{65}$$

Objective (63) minimizes the total operation cost of the schedule. Constraint (64) ensures that each trip is carried out once by a vehicle. As set Ω includes a large number of feasible trip chains, it is impractical to solve the model directly. Therefore, we developed a column generation (CG) algorithm to solve the linear relaxation of models (63)–(65) called master problem (MP). The CG algorithm starts by solving the linear relaxation of models (63)–(65) with an initial set of feasible trip chains $\overline{\Omega}$, called the restricted master problem (RMP). In each iteration, a pricing problem is solved to generate columns with negative reduced cost to be added to the RMP. Let π_i, $i \in T$, be the dual variables associated with constraint (64). Let $\overline{c}_p$ be the reduced cost of trip chain $p \in \Omega$ with respect to π_i, $i \in T$, i.e., $\overline{c}_p = c_p - \sum_{i\in T} a_{pi}\pi_i$. The pricing problem is defined as follows:

$$\min \sum_{p\in\Omega} \overline{c}_p\theta_p. \tag{66}$$

A feasible trip chain $p \in \Omega$ corresponds to an o-d path in G. The problem aims at finding a feasible trip chain with the minimum reduced cost. We modify the cost of each arc $(i, j) \in A$ by a modified cost $\overline{c}_{ij} = c_{ij} - \pi_i$ where $\pi_i = 0$ for $i \in \{o, d\}$. The reduced cost of an o-d path p equals the sum of $\overline{c}_{ij}$, $(i, j) \in p$. The problem is a resource-constrained elementary shortest path problem (RCESPP), and we solve it by developing a label setting algorithm [35]. The CG alternates between the optimization of the RMP and the pricing problems until no more columns with negative reduced cost are generated, implying that the MP has been solved to optimality.

4.1. Robustness Checking of a Trip Chain. In solving the pricing problem, we need to generate robust o-d paths in graph G. Based on the DP equations (25) and (26) defined in Section 3.3, the robustness of path P can be checked by the following procedure.

The Robustness Checking Procedure. On an arc (v_i, v_j) of path P, for each $\gamma \le \Gamma^t$, determine the values of $W_{ij\gamma}$ and $\widetilde{W}_{ij\gamma}$ by equations (67) and (68). Then, the values of $Z_{j\gamma}$ and $Y_{j\gamma}$ are determined by equations (25) and (26), respectively. Path P is robustness feasible if $Z_{i\gamma} \in [a_i, b_i]$ and $Y_{i\gamma} \in [u_{i\gamma}^{\min}, u^{\max}]$ for $v_i \in P$ and $\gamma \le \Gamma^t$.

$$W_{ij\gamma} = \max\left\{0,\ \min\left\{u^{max}, \left(a_j - Z_{i\gamma} - t_i - t_{ij} - t^p\right)r_s\right\}\right\}, \tag{67}$$

$$\widetilde{W}_{ij\gamma} = \max\left\{0,\ \min\left\{u^{max}, \left(a_j - Z_{i,\gamma-1} - t_i - \widehat{t}_i - t_{ij} - t^p\right)r_s\right\}\right\}. \tag{68}$$

4.2. Label Setting for the Pricing Problem. In this section, we describe the label setting algorithm for generating robust o-d paths based on the robustness checking approach proposed in Section 4.1. In solving the RCESPP on graph G, the labels are used to represent the partial paths starting from the start node and the resources accumulated along the path. New labels are generated by extending the existing labels on graph G following the resource extension functions (REFs). New labels are checked for resource feasibility and infeasible labels are discarded. At each node, dominance rule is then applied to eliminate the labels which cannot lead to a path with a better reduced cost than the dominate labels.

Denote $L_i = (C_i, R_{i\gamma}^t, R_{i\gamma}^e)$, $\gamma \le \Gamma^t$, as a label representing a partial path from the start node o to a node $i \in V$ with three kinds of resource defined as follows: C_i is the reduced cost of the path, $R_{i\gamma}^t$ is the earliest departure time of node i, and $R_{i\gamma}^e$ is the lowest battery SoC of the vehicle upon the start of trip i, when γ trips from node o to node i are attaining their maximum trip time.

In the VRP, when the customer time windows are wide, two customers can be served in different orders so that the graph is cyclic. As such, in solving the pricing problem, at a given node i in the graph, an elementary resource is defined for each customer to record the number of times the customer has been visited on the partial path $o, 1, 2, \ldots, i$ [31]. In the R-MD-EVSP, because the departure time window of the trips is small, any two trips cannot be served in different orders and graph G is acyclic. Therefore, in solving the pricing problem, the elementary resource is not needed.

Based on the recursive equations (25) and (26) and the feasibility checking approach introduced in Section 4.1, the forward extension of a label L_i to a label L_j along an arc $(i, j) \in A$ is performed by REFs (69)–(71). As the feasibility checking approach always ensures $R_{j\gamma}^e \in [u_{j\gamma}^{\min}, u^{\max}]$, the

resulting label L_j is robust feasible if and only if $R^t_{j\gamma} \in [a_j, b_j]$, $\gamma = 0, 1, 2, \ldots, \Gamma^t$.

$$C_j = C_i + \dot{c}_{ij} - \pi_i, \tag{69}$$

$$R^e_{j\gamma} = \min\left\{u^{\max}, R^e_{i\gamma} - e_i + W_{ij\gamma} - e_{ij}, R^e_{i,\gamma-1} - e_i - \widehat{e}_i + \widetilde{W}_{ij\gamma} - e_{ij}\right\}, \tag{70}$$

$$R^t_{j\gamma} = \max\left\{a_i, R^t_{i\gamma} + t_i + \frac{W_{ij\gamma}}{r_s} + t_{ij}, R^t_{i\gamma-1} + t_i + \widehat{t}_i + \frac{\widetilde{W}_{ij\gamma}}{r_s} + t_{ij}\right\}. \tag{71}$$

Variables $W_{ij\gamma}$ and $\widetilde{W}_{ij\gamma}$ are consistent with those defined in Section 3.3. $\dot{c}_{ij}$ is the estimated cost on arc $(i, j) \in A$ consisting of the cost of vehicle usage, empty travel, trip start delay, and charging cost. $\dot{c}_{ij}$ is computed as follows. Let $P_j = \{o, 1, 2, \ldots, i, j\}$ be a partial path from node o to node j. Denote H_j as the maximum total charging amount required for P_j when the trip times varies within uncertainty set U. $\dot{c}_{ij} = c_{ij} + g_{ij}(H_j - H_i)$ where c_{ij} is the vehicle usage and empty travel cost assigned to arc (i, j); $g_{ij}(H_j - H_i)$ is the increased charging cost.

In the label setting algorithm, dominance rule plays a critical role in reducing the number of extensions required as the labeling process goes on. In Proposition 1, we introduce a dominance rule for our label setting algorithm to improve the computational efficiency.

Proposition 1. *Let L^1_i and L^2_i be the two labels associated with the paths ending at node i. L^1_i dominates L^2_i if (i) $C^1_i \le C^2_i$; (ii) $R^{e1}_{i\gamma} \ge R^{e2}_{i\gamma}$, $\gamma = 0, 1, 2, \ldots, \Gamma^t$; and (iii) $R^{t1}_{i\gamma} \le R^{t2}_{i\gamma}$, $\gamma = 0, 1, 2, \ldots, \Gamma^t$.*

4.3. Heuristic Pricing. The label setting algorithm can be time-consuming, especially in solving large-scale instances. As we know, it is not necessary to solve the pricing problem to its optimal and return a column with the most negative reduced cost at each iteration of the CG. The computational efficiency of the GC procedure can be improved by only finding suboptimal solutions in the pricing problem at the expense of an increase in the number of iterations. As such, we adopt heuristic decisions before invoking the exact pricing algorithm. Instead of storing all the nondominant columns in the label setting, we maintain only a prespecified number of them to improve the algorithm efficiency. When a new nondominant label is added to the list and the total number of labels has reached the limit, the label with the largest objective value in the list will be discarded. In this way, fewer labels are extended at each node, speeding up the process of negative path detection. Note that this heuristic may lead to more CG iterations and the number of labels maintained should be carefully set through preliminary experiment.

4.4. Branching. An initial solution is formulated by assigning each e-bus with one trip. After each CG procedure ends, to obtain the integer solution, branches are created in the branch-and-bound tree by branching on the total flow of the arcs $(x_{ij})_{(i,j)\in A}$. If not all the $(x_{ij})_{(i,j)\in A}$ are integers, we choose to branch on x_{ij} that has the highest fractional value, creating two nodes, one with $x_{ij} = 1$ and the other with $x_{ij} = 0$. Otherwise, an integer solution is obtained. A depth-first search strategy is used to explore the branch-and-bound tree. In order to obtain integer solutions more efficiently, after each CG procedure ends, x_{ij} values in the RMP that are larger than 0.99 are set to 1, while x_{ij} values that are less than 0.01 are set to 0. The preliminary experiments show that this setting works well on speeding up the computational time and has minor impact on the solution quality.

4.5. Obtaining Optimal Charging Plan. After the optimal schedule is obtained by the BP algorithm, we further generate the optimal charging plan for each individual trip chain, considering the time-of-use tariff. The problem can be defined as a robust fixed-route electric vehicle refueling problem (R-FRVRP) which aims to determine the optimal charging time and amount for an e-bus on a given trip chain.

5. Numerical Experiments

We conducted numerical experiments based on the case of bus routes in Shenzhen. We first conducted experiments on single-route scheduling and then on multiroute scheduling where e-buses are allowed to carry out trips on different routes. In the following cases, plug-in DC chargers are established in the depots, the majority of which have the maximum charging power of 100 kW. BYD e-buses with a battery capacity of 260 kWh, length of 10.5 m, weight of 18000 kg, and maximum driving distance of 220 km operate on the routes. The characteristics of the bus routes and charging facilities are presented in Table 2. According to the historical running data, the charging rate is set to be 1.6 kWh per minute, and the battery energy consumption rate is estimated to be 0.6 kWh per minute. As defined in Section 3.1, the energy consumption of a trip depends on the trip time following the equation $\delta^e_i = \alpha\delta^t_i r_u$. Here we set $\alpha = 1$. The charging preparation time t^p and minimum charging time t^s are set as 2 min and 5 min, respectively. Parameter α takes the value of 0.5. The daily usage cost of an e-bus c^f is set as 1000 CNY. Figure 3 shows the time-of-use tariff in Shenzhen, which is used to set the values of g_{ij}, $(i, j) \in A$.

Table 2: Characteristics of the bus routes in the single and multiroute scheduling cases.

Route	Length (km)	Number of stops	Operation period	Headway (min)	Ave. trip time (min)	Number of trips	End depot I	End depot II
M133	37.5	44	6 : 10–22 : 30	5–12	110	276	1 (16)	2 (8)
E7	39.7	15	7 : 00–22 : 00	10–60	70	84	3 (20)	4 (5)
E11	75.5	18	6 : 00–21 : 30	6–10	120	194	3	5 (10)
E14	41.9	22	7 : 30–22 : 00	15–50	110	70	3	6 (10)
E22	57.1	19	6 : 30–20 : 00	60	120	34	3	7 (5)
42	17.3	23	6 : 30–22 : 30	10–15	45	132	8 (10)	9 (8)
43	17.6	25	6 : 30–22 : 30	10–15	60	130	3	9
81	31.6	44	6 : 30–21 : 30	5–15	80	204	3	8

Note. End depots are numbered from 1 to 9 as follows: 1, Changlingdong depot; 2, Shekou bus depot; 3, Shenzhen North Railway Station; 4, Longgangyiwu bus depot; 5, Nanaoxinda bus depot; 6, Tantou West; 7, Hangzi bus depot; 8, Moon Bay; 9, Window of the World. The values in the brackets next to the depot index are the number of charging piles established at the depots.

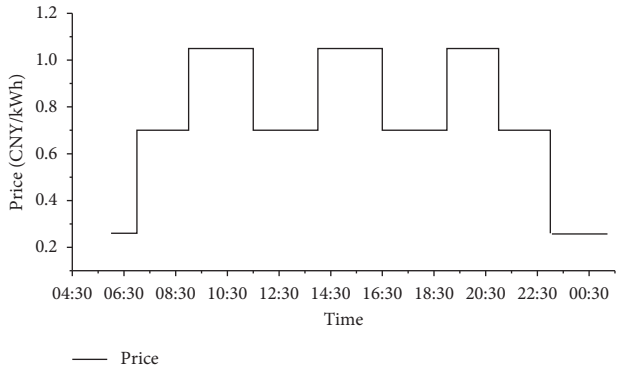


Figure 3: Time-of-use tariff in Shenzhen.

The MIP model was solved by the standard optimization solver Cplex 12.6, and the BP framework was coded in Java using Cplex 12.6 to solve the RMP. All the experiments were run on a PC with Windows 10, Intel Core i5-8250U, 1.80 GHz, and 8 GB RAM.

5.1. Single-Route Scheduling. We first analyze the performance of the MIP model and BP algorithm based on Route M133 in Shenzhen. A display of bus Route M133 is shown in Figure 4.

In generating a timetable, estimation of the scheduled trip times is critical. We adopted the method suggested in [27] to generate homogeneous running time (HRT) periods, the trip time within each of which follows the same distribution. We generate the HRT periods for Route M133 based on the trip time data of 100 weekdays from August 1 to December 31 in 2017. Figure 5 presents the distribution of the trip time of Route M133 during different periods of the day. Within each HRT period, the scheduled trip time is set according to a rule-of-thumb: the average trip time plus the standard deviation of the trip time. Figure 6 shows the scheduled trip time and maximum deviation from the scheduled trip time within each running time period. We generate the timetable including 276 trips. Except for the original timetable, we also create timetables with the number of scheduled trips ranging from 64 to 154 by changing the headway requirements. These timetables account for the small-scale instances.

5.1.1. Results. The performance of the MIP model and BP algorithm with uncertainty budget $\Gamma^t = 0, 1, 2, 3$ is presented in Table 3. The deterministic solution corresponds to the case with $\Gamma^t = 0$. The table shows the objective of the MIP model obtained by Cplex and BP algorithm (Obj), number of vehicles used (#V) in the solution obtained by the BP, optimal gap of the MIP and BP, and the computational time. When the MIP gap is higher than 90%, we report the root relaxation solution obtained by Cplex in the bracket. If the root relaxation is not solved by Cplex within 3600 s, an em dash ("—") is used.

The results indicate that Cplex is not able to solve the model to near optimal in a short time. Some instances are not solved by Cplex in 3600 s; however, the BP algorithm is able to generate high-quality solutions with small BP gaps within a reasonable computational time. The objective obtained by the BP algorithm is consistently lower than that obtained by the MIP model. For a given instance, the total operational cost, including the cost of vehicle usage and charging, becomes higher with the increase of the uncertainty budget Γ^t. This is due to the more vehicles put into use and a larger amount of planned daytime charging amount. When the level of protection Γ^t increases, the planned daytime charging amount becomes higher if the number of vehicles used stays the same; the daytime charging amount can be reduced when the number of vehicles used increases.

5.1.2. Impact of the Trip Start Buffer Time. We investigated the impact of trip start buffer time on the optimal schedule. Each case is named as "number of trips—budget value." The trip start delay time t^d varies among 0, 3, 5, and 8 minutes. $t^d = 0$ means that trip start delay is not permitted. As shown in Figure 7, the number of e-buses used is reduced with the increase of t^d in cases "96-0," "112-0," and "276-0." The results indicate that allowing for trip start buffer time can reduce the number of e-buses put into operation in some cases. Trip start buffer time can increase time flexibility of the schedule and saves the operational cost in some scenarios. With the increase of the uncertainty budget, the planned charging amount increases, protecting the schedule against the variation of energy consumption.

Figure 4: A display of bus Route M133 in Shenzhen.

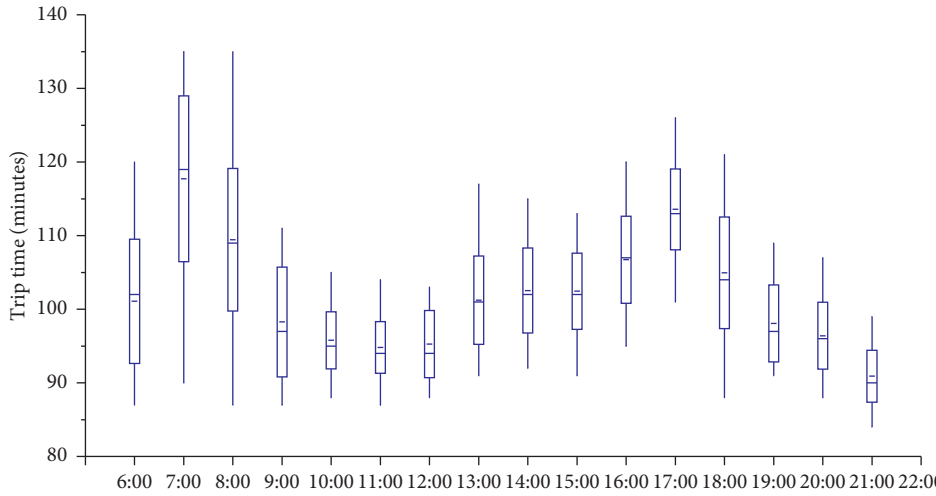


Figure 5: Trip time distribution of Route M133 in a day based on the historical running data.

5.2. Multiroute Scheduling. We carried out numerical experiments on two cases with multiple bus routes and depots. The multiroute scheduling is applied in the cases. The characteristics and layouts of the transit system in the case are shown in Table 2 and Figure 8. Case I includes five depots and four express bus routes: Route Nos. E7, E11, E14, and E22, that all have one end stop at the Shenzhen North Railway Station, formulating a hub-and-spoke topology. The timetable of the four routes includes a total number of 382 trips. Case II includes three depots and three bus routes: Route Nos. 42, 42, and 81, that share end stops with each other, formulating a circle topology. All the depots are equipped with charging facilities. The timetable of case I and case II includes a total number of 382 and 466 trips.

5.2.1. Results. We obtained the optimal schedule for cases I and II by the BP algorithm. The computational results are shown in Table 4. The table shows the objective of the MIP model obtained by the BP algorithm (Obj), number of vehicles used (#V), total charging amount, computational time, and BP gap with uncertainty budget $\Gamma^t = 0, 1, 2, 3$.

5.3. Schedule Robustness. We investigated the schedule robustness through Monte Carlo simulation based on the randomly generated trip time data. The simulation is performed by generating $N^{ran\,d} = 10000$ random realizations of trip time from normal distribution $N(t_i^a, sd_i)$ for all trip nodes $i \in T$, where t_i^a and sd_i are the mean and standard deviation of the trip time. Each random realization represents one day's e-bus schedule.

Table 5 presents the statistics on the deterministic ($\Gamma^t = 0$) and robust ($\Gamma^t = 1, 2, 3$) solutions for single-route instances. The number of vehicles used (#V), infeasibility rate of the trip chain (IR-TC), schedule (IR-S), and price of robustness (PoR) are reported. Denote the total number of infeasible trip chains as N_{tc}^{IF}. If the schedule contains infeasible trip chains, the schedule is regarded as infeasible. Denote the total number of infeasible schedules as N_s^{IF}. IR-TC is calculated as $N_{tc}^{IF}/(N^{ran\,d} * N^{veh})$ where N^{veh} is the number of e-buses used; IR-S is calculated as $N_s^{IF}/N^{ran\,d}$. PoR is the percentage of increase in the operational cost of the robust solutions compared with that of the deterministic solution.

About 0.94% to 9.81% of the simulated e-bus trip chains experience delays using the deterministic solutions. With $\Gamma^t = 1$, the infeasibility rate drops to almost zero while the total operational cost increases by 1.25% to 16.14%. We calculated the penalty of the trips chains incurred by the violation of the trip start time window and battery SoC safety

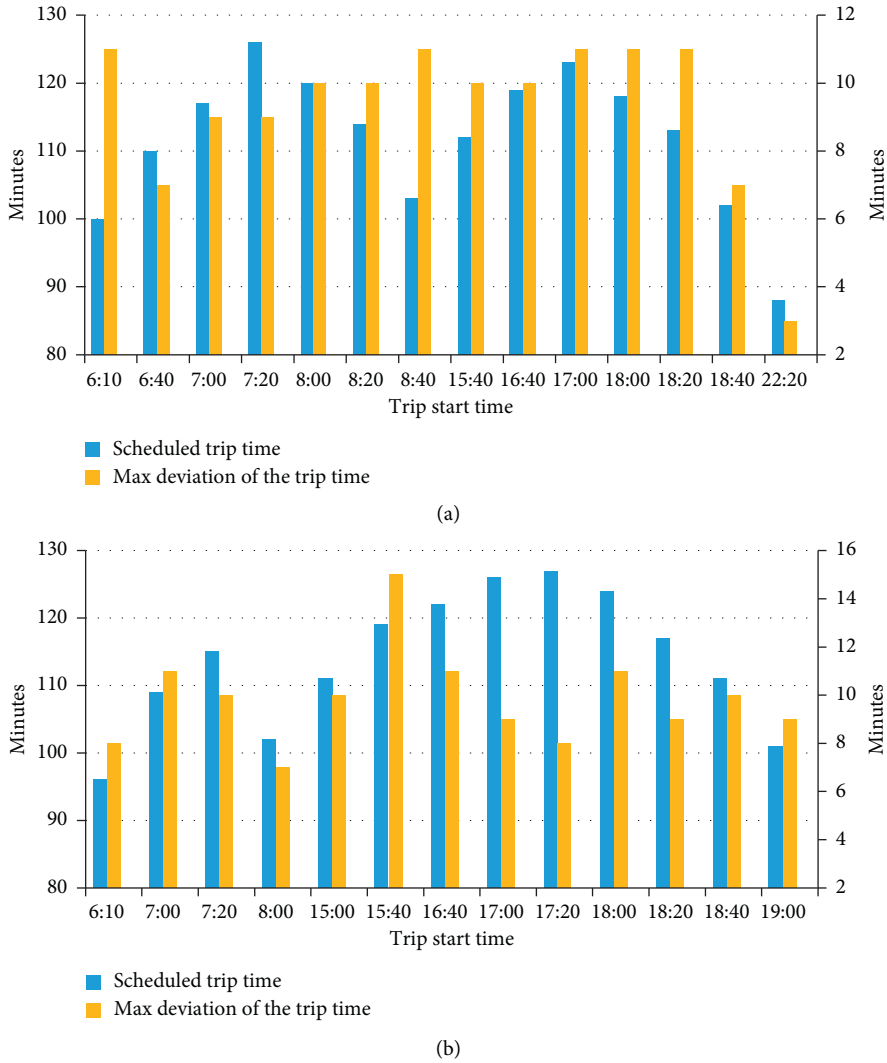


Figure 6: HRT periods of Route M133 generated based on the historical running data. (a) Outbound. (b) Inbound.

Table 3: Computational results of the MIP model and BP on single-route instances.

Number of trips	Γ^t	Obj		#V	Gap (%)		Run time (s)	
		MIP	BP	BP	MIP	BP	Cplex	BP
62	0	16916.5	16912.8	16	5.4	0	3600	0
	1	18107.3	16400.0	16	11.64	0	3600	1
	2	19876.8	16800.0	16	19.50	0	3600	1
	3	21763.3	17000.0	17	26.48	0.12	3600	1
96	0	22900.7	22676.7	21	8.30	0.14	3600	8
	1	25664.7	25504.5	22	18.18	0.21	3600	5
	2	30395.0	27279.2	22	30.91	0.21	3600	6
	3	35098.3	33331.9	22	40.17	0.26	3600	8
112	0	26294.3	25711.1	23	12.53	0.42	3600	10
	1	30508.7	26714.7	24	24.61	0.60	3600	7
	2	36842.2	30202.2	24	37.57	0.58	3600	8
	3	51072.5	33858.5	24	54.97	0.71	3600	10
154	0	38082.3	36578.7	33	13.35	0.20	3600	24
	1	69344.3	37041.8	33	52.41	0.22	3600	34
	2	(33000)	42180.0	33	—	0.68	—	56
	3	(33000)	45606.0	33	—	0.84	—	13
276	0	(49000)	56443.3	49	—	0.82	—	129
	1	(49000)	67309.1	49	—	0.86	—	169
	2	—	82386.7	49	—	1.02	—	191
	3	—	99134.4	49	—	0.92	—	113

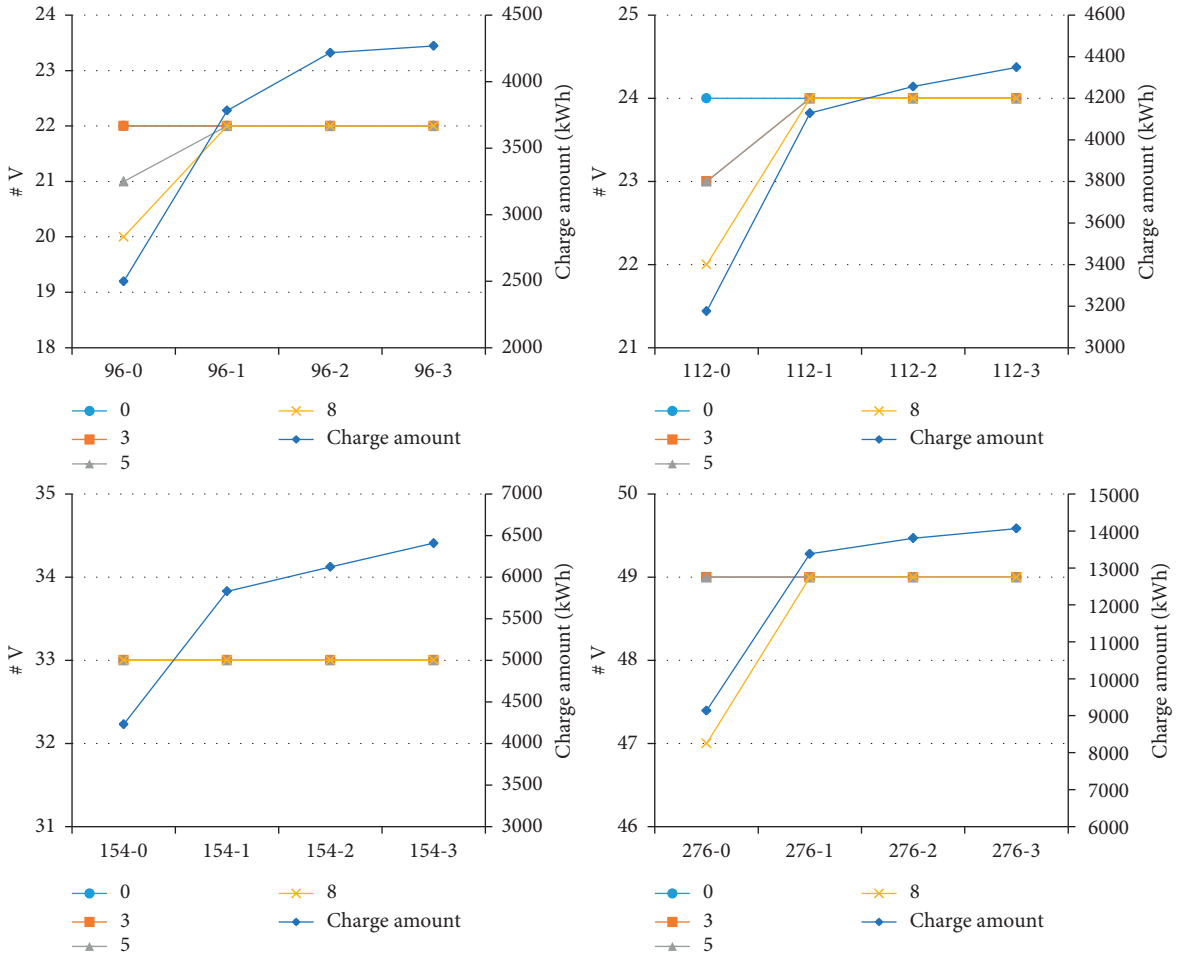


Figure 7: The influence of departure delay time on the number of vehicles used under different values of uncertainty budget.

(a)

Figure 8: Continued.

(b)

Figure 8: A display of the route layout of (a) case I and (b) case II.

Table 4: Computational results of the BP on multiroute instances.

Case	Γ^t	Obj.	#V	Charging amount (kWh)	Run time (s)	BP gap (%)
Case I	0	81638.3	73	11141.3	392	1.02
	1	92313.4	76	16313.4	398	0.96
	2	114236.0	76	17522.4	676	1.10
	3	132200.5	76	18247.8	590	1.32
Case II	0	69279.0	64	6895.6	6248	1.89
	1	76556.2	64	11178.4	8979 8261	1.84
	2	86540.8	64	13533.2	8261	0.12
	3	98607.3	64	13444.1	6350	1.06

Table 5: Robustness of the single-route schedules.

Number of trips	Γ^t	Infeasibility rate (%)		PoR (%)
		Trip chain	Schedule	
96	0	0.94	19.27	—
	1	0.00	0.00	9.52
	2	0.00	0.00	15.41
	3	0.00	0.00	30.77
112	0	7.20	87.31	—
	1	0.00	0.05	3.76
	2	0.00	0.00	14.87
	3	0.00	0.00	24.06
154	0	5.65	92.08	—
	1	0.01	0.26	1.25
	2	0.02	0.70	13.28
	3	0.03	1.04	19.79
276	0	9.81	99.86	—
	1	0.00	0.00	16.14
	2	0.00	0.00	31.49
	3	0.00	0.00	43.06

range. Figure 9 shows the changes of the objective and penalty per 100 trip chains under different values of Γ^t. With the increase of Γ^t, the objective increase while the penalty decreases. This result coincides with that in Table 3, indicating that the schedules with $\Gamma^t = 0$ and $\Gamma^t = 1$ are Pareto optimal solutions.

Table 6 presents the robustness evaluation results of the multiroute scheduling cases. When $\Gamma^t = 0$, about 4.09% and

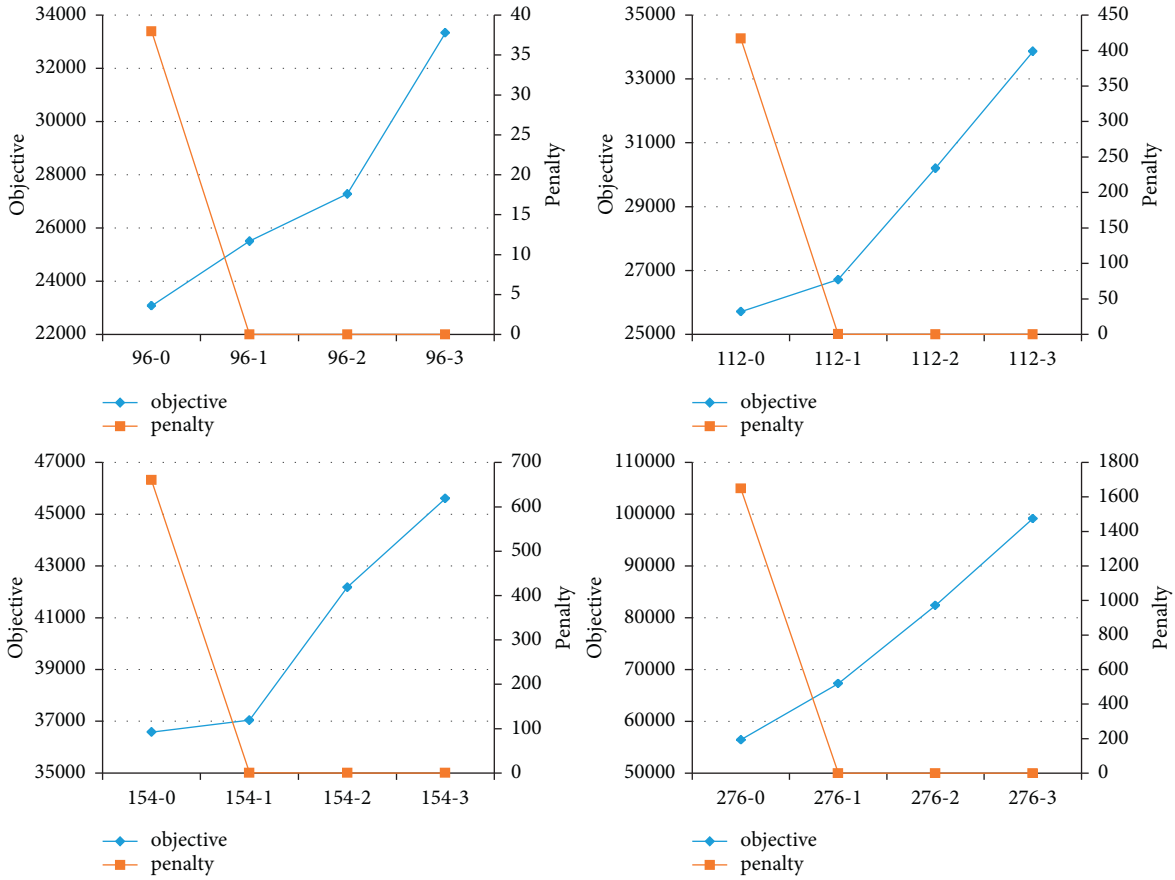


Figure 9: Objective and penalty under different values of Γ^t in single-route cases.

Table 6: Robustness of the multiroute schedules.

Case	Γ^t	Infeasibility rate (%)		PoR (%)
		Trip chain	Schedule	
Case I	0	4.09	98.84	—
	1	0.00	0.00	10.49
	2	0.00	0.00	19.19
	3	0.00	0.00	13.59
Case II	0	3.12	95.36	—
	1	0.00	0.00	8.48
	2	0.00	0.00	11.54
	3	0.00	0.00	12.24

3.12% of the simulated trip chains experience delay in cases I and II, respectively. When $\Gamma^t = 1$, the infeasibility rate drops to zero while the objective increases by 10.49% and 8.48% in cases I and II, respectively. We also calculated the penalty of the trips chains incurred by the violation of the trip start time window and battery SoC safety range. Figure 10 shows the changes of the objective and penalty per 100 trip chains under different values of Γ^t. With the increase of Γ^t, the objective increases while the penalty decreases. The results indicate that the schedules with $\Gamma^t = 0$ and $\Gamma^t = 1$ are Pareto optimal solutions.

Based on the above evaluation results, we can make the following conclusions:

(i) The infeasibility rate of the deterministic solution is significantly higher than that of the robust solutions in the simulation.

(ii) The schedule robustness can be greatly improved at the expense of an increase in the total operational cost.

(iii) Trade-off between the infeasibility rate and PoR is needed to choose a proper value of uncertainty budget when generating robust schedules.

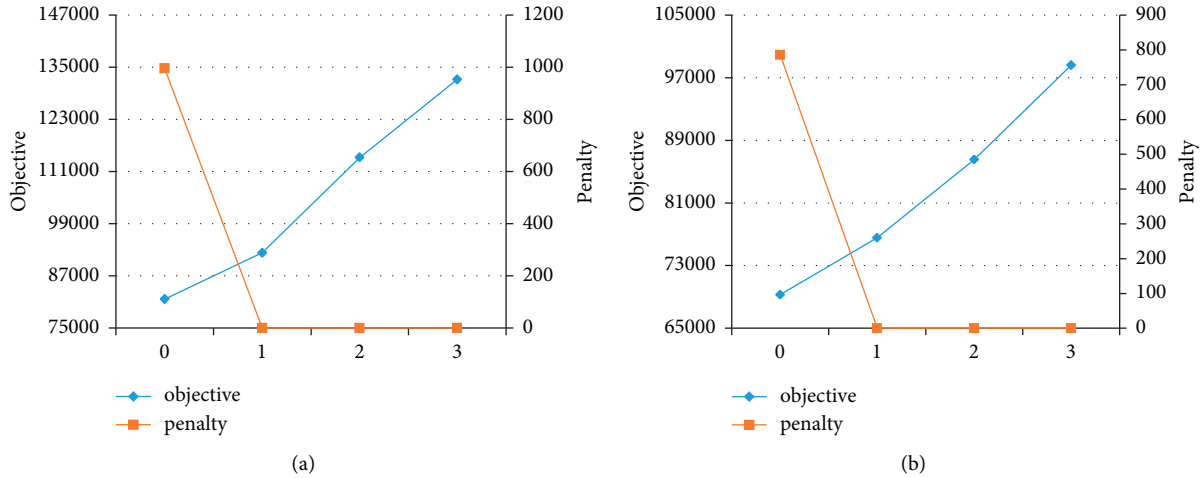


Figure 10: Objective and penalty under different values of Γ^t in multiroute cases. (a) Case I. (b) Case II.

6. Conclusion

In this paper, an R-MD-EVSP model and BP algorithm are introduced as a novel approach for e-bus scheduling under travel time uncertainty. Partial charging and trip start time window are taken into consideration. The cardinality constrained uncertainty set was adopted to control the level of uncertainty. The proposed methodology was tested on the real-world transit operation cases in Shenzhen including single-route and multiroute cases. The size of the instances varies from 62 to 466 timetabled trips. The results showed that the BP algorithm is able to generate provably high-quality solutions efficiently for large-scale instances. With the increase of the uncertainty budget, the planned charging amount increases to protect the schedule against the variation of energy consumption when the number of e-buses used stays the same; the charging amount can be reduced when the number of vehicles used increases. Additionally, a sensitivity analysis of the trip start buffer time was carried out. The analysis showed that allowing for trip start buffer time can increase time flexibility of the schedule and save the operational cost in some scenarios. The robustness of the schedules was evaluated through Monte Carlo simulation based on the randomly generated trip time data. The results showed that the infeasibility rate of the deterministic solution is significantly higher than that of the robust solutions. Applying the robust schedules can effectively protect against trip time and energy consumption variation at the expense of an increase in the operational cost. Therefore, transit operators need to trade off between the infeasibility rate and PoR to choose a proper value of uncertainty budget when adopting robust schedules.

In future research studies, more realistic battery charging and energy consumption process should to be considered in the robust e-bus scheduling model. Besides, heuristic methods are needed to solve the problem instances of larger scale. Furthermore, different kinds of uncertainty sets can be considered to control the trip time variation. Trip time and energy consumption can be described by independent uncertainty sets, accounting for more general e-bus operation cases.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported in part by the Basic Research Program of Shenzhen Science and Technology Innovation Committee (JCYJ20180307123910003), Scientific Research Start-Up Funds of Tsinghua Shenzhen International Graduate School (QD2021007N), and National Natural Science Foundation of China (61673233).

References

[1] J.-Q. Li, "Battery-electric transit bus developments and operations: a review," *International Journal of Sustainable Transportation*, vol. 10, no. 3, pp. 157–169, 2016.

[2] Y. Shen, J. Xu, and J. Li, "A probabilistic model for vehicle scheduling based on stochastic trip times," *Transportation Research Part B: Methodological*, vol. 85, pp. 19–31, 2016.

[3] D. Huisman, R. Freling, and A. P. M. Wagelmans, "A robust solution approach to the dynamic vehicle scheduling problem," *Transportation Science*, vol. 38, no. 4, pp. 447–458, 2004.

[4] M. Naumann, L. Suhl, and S. Kramkowski, "A stochastic programming approach for robust vehicle scheduling in public bus transport," *Procedia - Social and Behavioral Sciences*, vol. 20, pp. 826–835, 2011.

[5] F. He, J. Yang, and M. Li, "Vehicle scheduling under stochastic trip times: an approximate dynamic programming approach," *Transportation Research Part C: Emerging Technologies*, vol. 96, pp. 144–159, 2018.

[6] X. Tang, X. Lin, and F. He, "Robust scheduling strategies of electric buses under stochastic traffic conditions," *Transportation Research Part C: Emerging Technologies*, vol. 105, pp. 163–182, 2019.

[7] Y. Bie, J. Ji, X. Wang, and X. Qu, "Optimization of Electric Bus Scheduling Considering Stochastic Volatilities in Trip Travel Time and Energy Consumption," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 12, pp. 1530–1548, 2021.

[8] A. Bertossi, P. Carraresi, and G. Gallo, "On some matching problems arising in vehicle scheduling models," *Networks*, vol. 17, no. 3, pp. 271–281, 1989.

[9] J.-Q. Li, "Transit bus scheduling with limited energy," *Transportation Science*, vol. 48, no. 4, pp. 521–539, 2014.

[10] Y. Yang, M. Guan, and J. Ma, "Battery electric transit bus scheduling problem based on column generation approach," *Journal of Transportation Systems Engineering and Information Technology*, vol. 16, pp. 198–204, 2016.

[11] T. Liu and A. Ceder, "Battery-electric transit vehicle scheduling with optimal number of stationary chargers," *Transportation Research Part C: Emerging Technologies*, vol. 114, pp. 118–139, 2020.

[12] L. Li, H. K. Lo, and F. Xiao, "Mixed bus fleet scheduling under range and refueling constraints," *Transportation Research Part C: Emerging Technologies*, vol. 104, pp. 443–462, 2019.

[13] M. Rinaldi, E. Picarelli, A. D'Ariano, and F. Viti, "Mixed-fleet single-terminal bus scheduling problem: modelling, solution scheme and potential applications," *Omega*, vol. 96, Article ID 102070, 2020.

[14] E. Yao, T. Liu, T. Lu, and Y. Yang, "Optimization of electric vehicle scheduling with multiple vehicle types in public transport," *Sustainable Cities and Society*, vol. 52, Article ID 101862, 2020.

[15] L. Zhang, S. Wang, and X. Qu, "Optimal electric bus fleet scheduling considering battery degradation and non-linear charging profile," *Transportation Research Part E: Logistics and Transportation Review*, vol. 154, Article ID 102445, 2021.

[16] A. Zhang, T. Li, R. Tu et al., "The effect of nonlinear charging function and line change constraints on electric bus scheduling," *Promet - Traffic & Transportation*, vol. 33, no. 4, pp. 527–538, 2021.

[17] M. Wen, E. Linde, S. Ropke, P. Mirchandani, and A. Larsen, "An adaptive large neighborhood search heuristic for the Electric Vehicle Scheduling Problem," *Computers & Operations Research*, vol. 76, pp. 73–83, 2016.

[18] M. E. van Kooten Niekerk, J. M. van den Akker, and J. A. Hoogeveen, "Scheduling electric vehicles," *Public Transport*, vol. 9, no. 1-2, pp. 155–176, 2017.

[19] M. Kohání and M. Kohánia, "Exact approach to the electric bus fleet scheduling," *Transportation Research Procedia*, vol. 40, pp. 1380–1387, 2019.

[20] X. Li, T. Wang, L. Li et al., "Joint optimization of regular charging electric bus transit network schedule and stationary charger deployment considering partial charging policy and time-of-use electricity prices," *Journal of Advanced Transportation*, vol. 2020, pp. 1–16, 2020.

[21] Ş. Yıldırım and B. Yıldız, "Electric bus fleet composition and scheduling," *Transportation Research Part C: Emerging Technologies*, vol. 129, Article ID 103197, 2021.

[22] N. Qin, A. Gusrialdi, R. Paul Brooker, and A. T. Raissi, "Numerical analysis of electric bus fast charging strategies for demand charge reduction," *Transportation Research Part A: Policy and Practice*, vol. 94, pp. 386–396, 2016.

[23] Y. Wang, Y. Huang, J. Xu, and N. Barclay, "Optimal recharging scheduling for urban electric buses: a case study in Davis," *Transportation Research Part E: Logistics and Transportation Review*, vol. 100, pp. 115–132, 2017.

[24] Y. He, Z. Liu, and Z. Song, "Optimal charging scheduling and management for a fast-charging battery electric bus system," *Transportation Research Part E: Logistics and Transportation Review*, vol. 142, Article ID 102056, 2020.

[25] K. Liu, H. Gao, Z. Liang, M. Zhao, and C. Li, "Optimal charging strategy for large-scale electric buses considering resource constraints," *Transportation Research Part D: Transport and Environment*, vol. 99, Article ID 103009, 2021.

[26] M. Salicrú, C. Fleurent, and J. M. Armengol, "Timetable-based operation in urban transport: run-time optimisation and improvements in the operating process," *Transportation Research Part A: Policy and Practice*, vol. 45, pp. 721–740, 2011.

[27] J. Xu and Y. Shen, "Setting scheduled trip time based on AVL data," *Journal of Transportation Systems Engineering and Information Technology*, vol. 12, no. 5, pp. 40–45, 2012.

[28] D. Bertsimas and M. Sim, "Robust discrete optimization and network flows," *Mathematical Programming*, vol. 98, no. 1, pp. 49–71, 2003.

[29] A. Agra, M. Christiansen, R. Figueiredo, L. M. Hvattum, M. Poss, and C. Requejo, "The robust vehicle routing problem with time windows," *Computers & Operations Research*, vol. 40, no. 3, pp. 856–866, 2013.

[30] P. Munari, A. Moreno, J. De La Vega, D. Alem, J. Gondzio, and R. Morabito, "The robust vehicle routing problem with time windows: compact formulation and branch-price-and-cut method," *Transportation Science*, vol. 53, no. 4, pp. 1043–1066, 2019.

[31] D. Lu and F. Gzara, "The robust vehicle routing problem with time windows: solution by branch and price and cut," *European Journal of Operational Research*, vol. 275, no. 3, pp. 925–938, 2019.

[32] S. Braaten, O. Gjønnes, L. M. Hvattum, and G. Tirado, "Heuristics for the robust vehicle routing problem with time windows," *Expert Systems with Applications*, vol. 77, pp. 136–147, 2017.

[33] C. Hu, J. Lu, X. Liu, and G. Zhang, "Robust vehicle routing problem with hard time windows under demand and travel time uncertainty," *Computers & Operations Research*, vol. 94, pp. 139–153, 2018.

[34] S. Pelletier, O. Jabali, and G. Laporte, "The electric vehicle routing problem with energy consumption uncertainty," *Transportation Research Part B: Methodological*, vol. 126, pp. 225–255, 2019.

[35] S. Irnich and G. Desaulniers, "Shortest path problems with resource constraints," in *Column Generation*, G. Desaulniers, J. Desrosiers, and M. M. Solomon, Eds., Springer, Boston, MA, USA, pp. 33–65, 2005.

Hindawi
Journal of Advanced Transportation
Volume 2021, Article ID 9445070, 12 pages
https://doi.org/10.1155/2021/9445070

Research Article

New Generation of Smart Highway: Framework and Insights

Chenglong Liu,[1,2] **Yuchuan Du**,[1,2] **Yiheng Ge**,[2] **Difei Wu**,[1] **Cong Zhao**,[1] **and Yishun Li**[1]

[1]*The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai 201804, China*
[2]*Shanghai Engineering Research Center of Urban Infrastructure Renewal, Shanghai 200032, China*

Correspondence should be addressed to Yuchuan Du; ycdu@tongji.edu.cn

Received 17 September 2021; Revised 13 October 2021; Accepted 2 November 2021; Published 16 December 2021

Academic Editor: Linchuan Yang

The new generation of smart highway (NGSH) has become an irresistible global trend to improve transport efficiency and safety. The exploration of the features and framework for NGSH can guide us to upgrade the current highway system. This paper summarizes the fundamental features of the NGSH from the perspective of the interactive evolution of automobile industry and road transport. In line with the popularity of automated and connected vehicles, the primary technical features of the NGSH are proposed as (I) complete elements sensing, (II) cyber-physical systems, (III) cooperative vehicle-infrastructure applications, and (IV) 5th generation mobile communication technology. The corresponding physical framework and data flow are introduced, in which three data attributes (data accuracy, dimensionality, and freshness) are highlighted to describe the data requirements for various scenarios. The development path of the NGSH is further discussed in terms of the different vehicle automation levels. The characteristics of five levels of NGSH are identified from R1 to R5. Different combinations of NGSH level and vehicle automation level lead to distinct system functions. Several urgent problems in the current stage are pointed out in terms of system compatibility, standard specification, and information security. This paper provides new insights for sustainable and reproducible highway reformation, drawing some implications for NGSH design.

1. Introduction

Highway is a fundamental component of the road transport system, which connects major cities along its route crossing all provinces and reaching nearly all of their capital cities [1]. It is a strong support for national economics. The highway in China has reached over 150,000 km in 2020, ranking first in the world [2]. With the rapid growth of highway's coverage, how to use advanced sensing, computing, and communication technologies to improve its service capability has been a pivotal issue for the next generation of smart highway (NGSH). China's fast-paced development has resulted in it being one of the most flexible nations in terms of technological integration into the NGSH [3]. In 2014, the Ministry of Transport of China released the "Four Transportation" strategy and proposed the concept of "smart transportation" [4]. In the following years, the conception of smart transportation has constantly been enriched and refined, generating many new technical forms, such as smart highway, smart pavement, and smart infrastructures [5]. Various regions and cities in China, such as Zhejiang, Shanghai, and Shandong, have carried out a large number of exploratory pilot projects to construct the smart highway system and achieve fruitful results. However, some of them are case-specific and function-oriented, making them quite hard to be replicated in other areas. The framework of the NGSH directly affects its portability and effectiveness, which is also of considerable significance to the evolution of the current road transport system.

The concept of the smart highway was proposed internationally as early as the 1940s [6]. In 1992, the United States launched the Automated Highway System (AHS) program and built a vehicle-vehicle (V2V) and vehicle-infrastructure (V2I) communication module to reduce traffic congestion and improve driving safety [7]. In the early 21st century, the Smart Highway project in Virginia was an important milestone for intelligent transportation system in America. It integrated multiple technologies such as facility performance monitoring, energy recovery, and automated driving to create

a 2.2-mile smart expressway [8]. The smart corridor is one of the most representative smart highway programs in the US conducted on the I-80 highway in California [9]. This system integrated a variety of sensing and active transportation technologies through electronic identification and dedicated short-range communication (DSRC) to improve driving safety and transit time reliability [10]. Since the 1990s, Japan has carried out many research and projects of intelligent transportation systems [11]. The SmartWay public test released in 2006 has become a worldwide ITS benchmark. The system contains vehicle information and communication system (VICS), electronic toll collection system (ETC), advanced safety vehicle system (ASV), and other functional modules to improve the user's travel experience [12]. Tokyo Motor Show 2015 exhibited the ETC 2.0 that takes V2I as the core to enhance tolling collection and information exchange between vehicles and facilities [13]. Japan has formed unified standards throughout the country and employed ETC and VICS systems. In the middle of the last century, Europe started the research on smart road infrastructure, which is mainly oriented by sustainability, eco-transport and safety, including the famous "Euraka" plan [14] and EasyWay project [15]. In 2013, Drive Me, a Gothenburg experimental road, provided roadside sensing data to vehicles in real-time for functions such as automatic parking and adaptive cruise [16]. In 2016, Norway's E8 highway applied vehicle-road collaboration to create a unique scenario application for autonomous truck driving [17]. Most countries utilized roadside cameras, millimeter-wave radars, etc. to achieve a large-scale perception of the road environment and combined different communication methods for information interaction between vehicles and infrastructures. The objectives vary from automated driving, accident alert, eco-based driving, safety enhancement, etc. Table 1 compares the technical features of typical smart highway projects throughout the world.

China built multiple smart highways for automated/semiautomated driving and cooperative vehicle-infrastructure systems (CVIS). For example, Zhejiang integrates technologies such as free-flow charging, wireless charging, and partially automated driving to create a "Superhighway" of 161 kilometers. Its driving speed is expected to exceed 120 kilometers per hour [18, 19]. Hunan builds 113 kilometers of "5G +" smart highway to achieve precise sensing of each section of the road and each vehicle to create a dynamic digital high-definition map for L3-L5 automated driving [20]. Based on the BeiDou Navigation Satellite System (BDS), Guangxi province constructs a smart highway system with functions of vehicle guidance, safety monitoring, and digitization of facilities for drivers. Shandong province launched the first domestic fully autonomous smart highway, which uses advanced laser radars, microwave radars, and other technologies to conduct macro-control of vehicles [20]. This program can significantly improve traffic efficiency and safety for fully connected automated vehicles, but the testing can only be applied at the designated training ground in the current stage instead of open environment. Although China has made tremendous achievements in the construction of smart highways, it has not yet produced a unified and instructive definition and framework of the NGSH. The lack of unified standards and design has affected the large-scale promotion of the NGSH nationwide.

As a new type of technical form of the road transport, the NGSH will become a reliable measure to solve traffic congestion and road crashes. The clarity of its fundamental architecture can effectively guide the design of the NGSH and serve as a complement to the current transportation system. The remainder of this paper is organized as follows: section 2 presents the interactive development between the automobile industry and road transport, which derives the connotation of the NGSH; section 3 elaborates the corresponding technical framework from the perspectives of physical architecture and data attributes; section 4 describes its development path and defines different levels of the NGSH; finally, conclusion part presents the findings of this study.

2. Understandings of the New Generation of Smart Highway

2.1. Technical Features of the NGSH. The core function of the road is the carrier of road users. According to the evolution history of the road transport system (see Figure 1), it shows an interactive development trend with the automobile industry. The first electric vehicle was invented in 1881, marking that the automobile has become a new tool for human mobility [21]. The Benz Patent-Motorwagen, built in 1885, is widely regarded as the world's first oil-fueled automobile, which is propelled by an internal combustion engine. Since then, the automobiles have gradually realized volume production [22]. However, the inadequate performance of road facilities has limited the efficiency of the vehicles. Therefore, the tarmac roads generated in 1901 provided a solid foundation for the development of automotive technology, leading to an improvement of vehicles' driving speed [23]. With the rapid expansion of human's social and travel distances, the automobiles require a safer and more efficient driving environment. There comes the highway system, which has the dedicated right-of-way and confined driving space. The world's first highway was built in 1932, and then it was widely adopted internationally [24]. Since the 21st century, the automated driving and Internet of vehicles have put forward new requirements for the existing highway system. The interactions between vehicles and roads are strengthened, leading the previous "weak tie" connection to become the "strong tie" coupling. The information exchange between them has been an indispensable part in the next generation of a road transport system. Therefore, the technical features of the NGSH need to fully consider the development status and functional requirements of the automotive industry.

User demand is the primary driving force for upgrading the highway system. At the beginning of the 20th century, hardening the pavement is the most significant measure to improve the serviceability of the pavement. By hardening the pavement, the traffic capacity of a single-lane increased to 600–800 pcu/h, and the driving speed can reach up to 40–60 km/h [25]. As the economic activities vigorously grew, the scope of people's activities was further expanded, producing lots of city-to-city travel demand. Therefore, the engineers came up with an "exclusive" highway system in

Table 1: Typical smart highway projects and technical features.

Area	Project	Communication	Key functions	Technical features
US	AHS [7]	DSRC	Collision avoidance Dynamic navigation Vehicle auxiliary control	V2V/V2I communication Automated roadside system
	Virginia smart road [8]	High-bandwidth fiber network	Energy recovery Pavement perception Intelligent lighting	Smart infrastructures Centralized communications
	I-80 SMART [9]	Optical cable	Adaptive ramp control Dynamic event broadcast Bus priority driving	Real-time traffic information Electronic signs, ramp meters
Japan	SmartWay [12]	DSRC	ETC Traffic broadcast Route optimization Over-the-horizon sight	Roadside sensor network ITS-on-board-unit VICS
Europe	E8-aurora [17]	C–V2X	Mobility as a service Automated driving in extreme weathers	Digital road facilities Intelligent asset management
	E8- borealis [17]	C–V2X	Truck autopilot	Digital road facilities C-ITS, V2V, V2I, V2X Radar and camera fusion
	Goteborg drive me [16]	C–V2X	Automated driving Automated parking	Internet of vehicles Infrastructure specifications Sensus navigation
China	Smart highway, Shandong [18]	5G LTE-V	Automated driving Truck platoon Full-element awareness	Semi-closed test field High precision positioning Internet of things Fiber grating
	Super highway, Zhejiang [19]	5G LTE-V DSRC	Free-flow tolling system Wireless charging Automated driving Data-driven maintenance	Energy recovery pavement Cloud & edge computing High precision positioning 5G-based VIIF system

*AHS: automated highway system; DSRC: dedicated short-range communications; V2V: vehicle-to-vehicle; V2I: vehicle-to-infrastructure; V2X: vehicle-to-everything; ETC: electronic toll collection system; VICS: vehicle information and communication system; C–V2X: cellular vehicle-to-everything; C-ITS: cooperative intelligent transport systems; LTE-V: long-term evolution-vehicle; 5G: 5th generation mobile network; VIIF: vehicle and infrastructure information fusion.

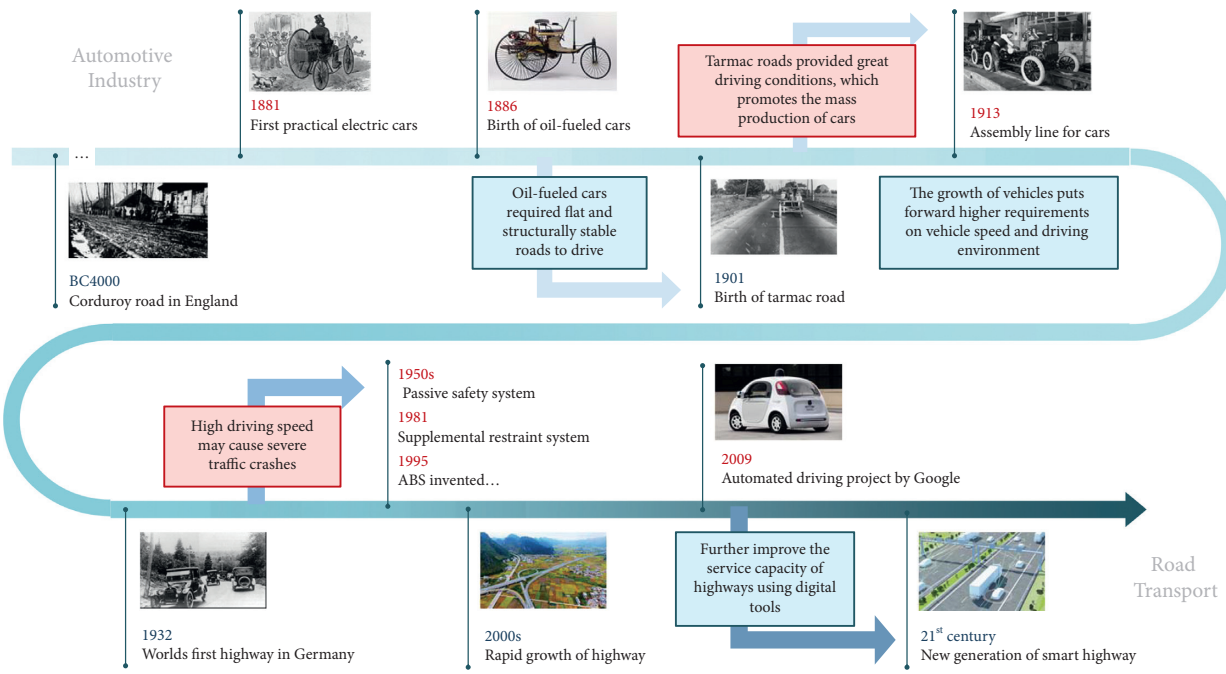


Figure 1: Interactive development of automobile industry and road transport.

1932, which only serves as a way open for the use of motor vehicles. The highway system expanded the single-lane capacity to 1800–2400 pcu/h and raised the vehicle speed to 80–120 km/h [26]. Given the gradual saturation of land resources, there are limited lands to construct new roads. The precise and fast sensing technology and low-latency communication assist vehicles in reducing their headway distance and ensure the stability of vehicle control under the existing infrastructure conditions. According to the simulation and field tests, the traffic capacity can leap up to 3000–3500 pcu/h, and the driving speed would exceed 120 km/h [27]. The basic idea of the NGSH is to combine the multiple, distributed, and precise sensing, computation, and communication technologies to improve the highway's capacity and driving speed, which also serves digital infrastructure management and safety enhancement (see Figure 2).

2.2. Fundamental Components of the NGSH System. The application of NGSH varies in terms of road users and managers. For users, the service functions are designed towards different automation levels of vehicles. As for L2 or less, the NGSH mainly provides reliable environmental information and corresponding service support, such as free-flow charging, hazard weather warning, and route and speed guidance [28]. For semiautomated driving vehicles (L3~L4), the NGSH utilizes its advantages of large-scale integrated sensing to enhance the vehicles' perception and provides more dedicated information to ensure driving safety. For fully automated vehicles (L5), the NGSH controls the vehicle fleet and platooning from the macro-perspective to improve the efficiency of the whole highway, as illustrated in Figure 3.

For managers, the NGSH monitors the status of static facilities and dynamic vehicles in real-time to carry out lifecycle maintenance management and optimization, which improves the response capabilities of emergencies and guarantees the stability of system operation. Based on the above requirements, the components of NGSH are roughly divided into four categories: (I) complete elements sensing, (II) cyber-physical systems (CPS), (III) cooperative vehicle-infrastructure applications, and (IV) high-speed communication (see Figure 4).

Complete elements sensing mainly applies advanced distributed sensing technologies such as video, millimeter-wave radar, and high-precision positioning system to achieve full element awareness of transportation system, which serves as the underlying structure of the NGSH. For example, vehicle reidentification, trajectory reconstruction, and velocity detection are applied to describe dynamic objects. Distress detection, roughness perception, marking recognition, etc. are employed to evaluate the quasistatic objects, such as pavement and roadside facilities. Environmental information such as hazard weather, water accumulation, and poor visibility is also recorded with a short interval. The NGSH fuses and correlates all these sensing data to construct a dynamic complete element sensing network.

Cyber physical system (CPS) is a computer system, in which a mechanism is controlled or monitored by computer-based algorithms [29]. CPS is similar to the digital twin platform [30], which reflects the dynamic operation information of the elements throughout their life cycle, and can spatially represent the sensing data through the global identification tags and data fusion model. CPS has the characteristics of the Internet of Things (IoT), but the CPS presents a higher combination and coordination between physical and computational elements [31]. The CPS platform is a living digital simulation that brings all the sensing data together and updates it from multiple sources to depict its physical counterpart dynamically. The CPS is the core platform of the NGSH.

CVIS applications refer to the use of complete sensing data and CPS platform to realize the application of automated driving and lifecycle infrastructure management, including CVIS-based fully automated driving, vehicle platooning control, sustainable maintenance scheduling, and rapid emergency response. According to different development stages of the NGSH, the systematic applications may vary from a single module to multiple superimposed functions.

Communication media refers to the use of 5G, DSRC, C–V2X, and other communication techniques to transfer data flow among the smart highway system's core components. The 5 G communication is expected to be a suitable tool for the large-scale vehicle-to-road information exchange [32], leading to the real-time calculation and processing of the integrated sensing data. In turn, the fused results and control orders from CPS platform can be sent to the vehicles by 5G in real-time.

3. Technical Framework of the New Generation of Smart Highway

The technical framework of the NGSH is designed based on its physical entities and types of application. The former explains the basic elements, physical composition, data mode, and information flow of the NGSH system. The latter enumerates the application scenarios.

3.1. Basic Elements of the NGSH. The first step to develop the technical framework of the NGSH is to determine its basic elements. Current studies prefer to determine the basic elements based on the application scenarios or the type of information. However, such definitions are not appropriate for the top-level framework design of the NGSH due to the complex items and function in the NGSH system. Thus, this study adopted a new method for determining the basic elements of the NGSH from the aspect of data freshness. Through summarizing the data requirements of each application scenario in the NGSH, the basic elements of the NGSH are determined as the four types of elements: highly dynamic elements, transient dynamic elements, transient static elements, and permanent static elements, as shown in Figure 5.

(1) Highly dynamic elements mainly include different types of traffic states, including the vehicles, roadblock, trajectories, and control commands.

(2) Transient dynamic elements mainly include the road network conditions, including the work zones, traffic congestions, weather conditions, road toll information, and signal control information.

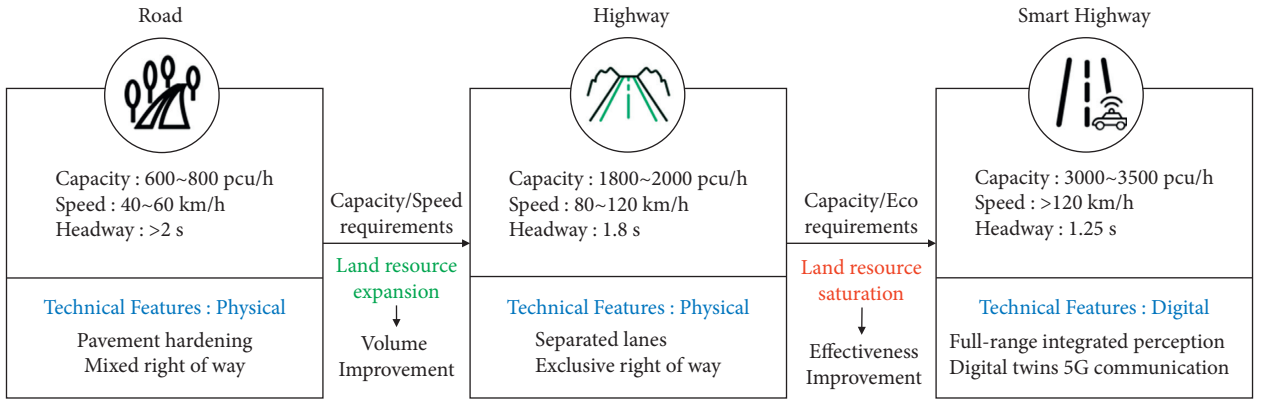


Figure 2: Process of upgrading the technical form of the road transport system.

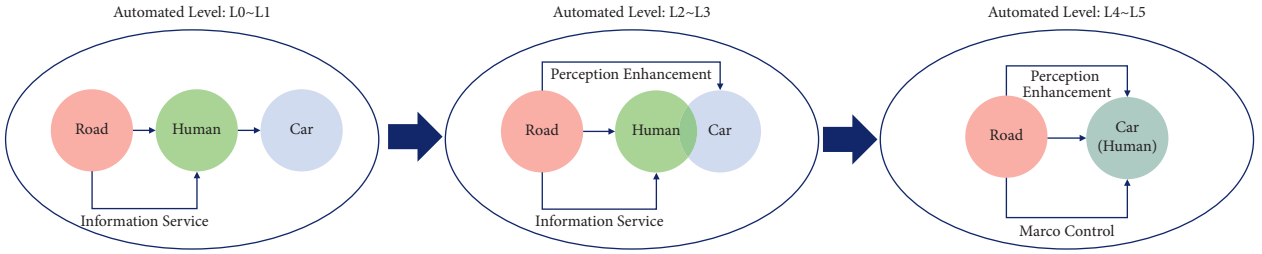


Figure 3: Functions of NGSH for different level of automation.

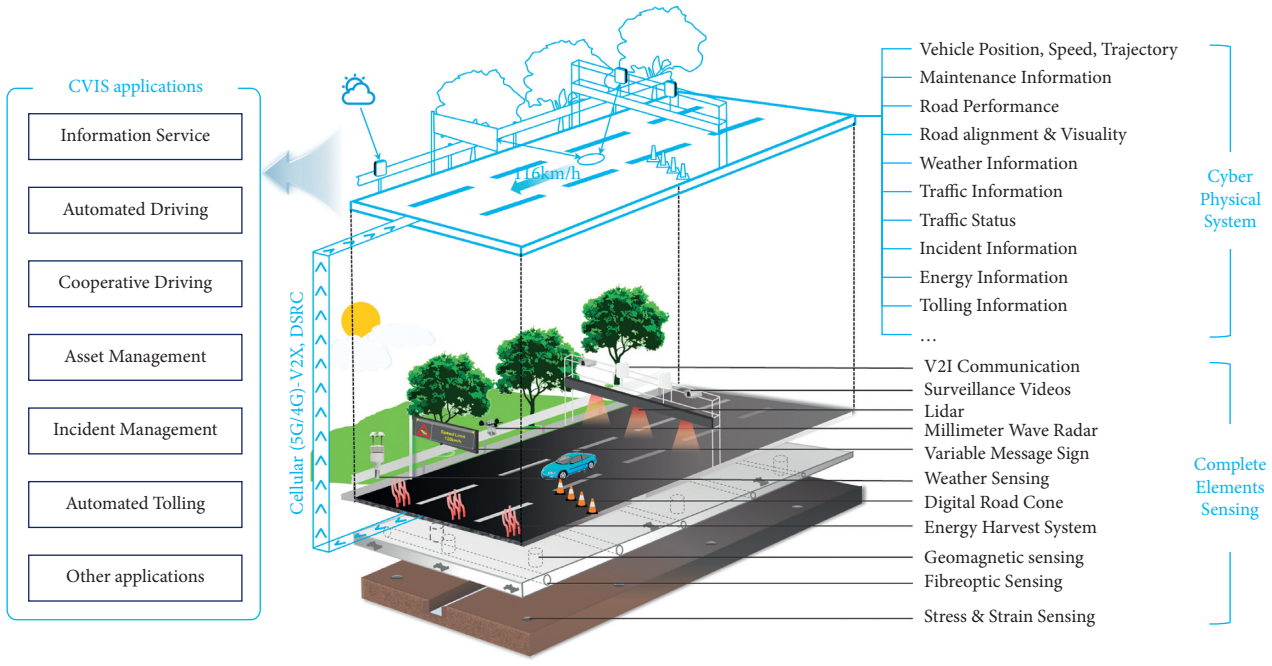


Figure 4: Main components of NGSH.

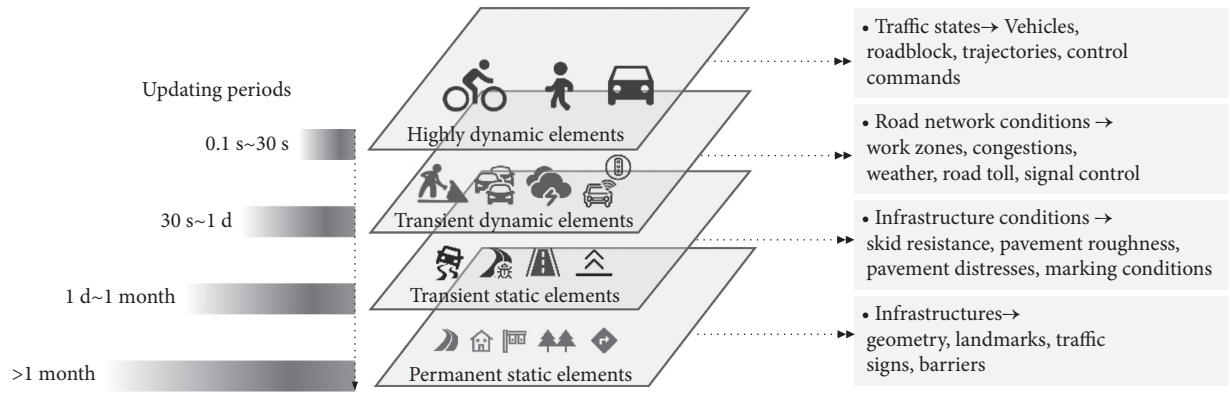


Figure 5: Basic elements of NGSH.

(3) Transient static elements are mainly the road network infrastructure conditions, including the pavement skid resistance, pavement roughness and unevenness, pavement distresses, and marking conditions.

(4) Permanents static elements refer to the features that are stable and almost unchanging, including the road geometry, landmarks, traffic signs, and barriers.

3.2. Physical Architecture of the NGSH. Architecture of an intelligent system can be developed by two methods: object-oriented method and process-oriented method. Compared with process-oriented method, object-oriented method focuses mainly on the objects in the system, with few characterizations on the process between objects, which is more scalable, reusable, and adaptable. Thus, this study adopts the object-oriented method to develop the physical architecture of the NGSH.

Figure 6 illustrates the physical architecture of the NGSH. The four objects of the architecture are firstly determined based on the basic elements. The attribute information, the perception methods, the processing, and analyzing methods are encapsulated in the element determinations and then make up the perception layer. For each object, different perception methods are adopted for data collection and information gathering.

The perceived data are then uploaded to the CPS layer by means of emerging communication techniques such as the 5G communication. The CPS platform is composed of the physical elements including the database, the basic platform, the computation center, and the functional platform. The CPS platform relies on the high-resolution digital base map. The physical elements of NGSH are mapped and managed in the CPS platform. The basic database is the management center of the massive multisource heterogeneous data such as digital infrastructure information, operational conditions, and environment data, the core of which is to ensure the synchronization of the data in physical and virtual worlds. Various types of data are presented in the basic platform according to data protocol. The computation center contains multiple virtual models, decision models, and analysis models, providing sufficient data resources to support the application systems.

The top layer of NGSH architecture is the application systems, which contains multiple services including the CVIS, intelligent infrastructure management and maintenance, information service, and intelligent traffic management. Compared to the other two layers, the architecture of the application system is more flexible and can be constituted by one or several functions.

The foundation of the physical framework is the complete sensing network and CPS management platform. The two parts must be compatible and replicable, which is conducive to promoting the productization and commercialization of the NGSH. It is also noted that the operation of one single application may require multiple information from CPS layer and perception layer. For instance, the automated vehicle in the CVIS requires both road geometry information (permanent static object), skid resistance information (transient static object), signal control information (transient dynamic object), and surrounding vehicle information (highly dynamic object) for data fusion and processing to generate control strategies.

3.3. The Data Requirements of the NGSH. The physical architecture defines the basic elements and major components in the NGSH system. For different application scenarios or systems, the data requirements are distinct. They must be specified firstly, because they determine the selection of perception method and the technical requirements of the perception devices. Therefore, we discuss the data requirements from the aspects of data dimension and accuracy, referring to the data richness and data precision. Table 2 summarizes the data requirements corresponding to different application scenarios.

3.4. Application Architecture of the NGSH. From the perspective of the applications, the users in NGSH involve private vehicles, commercial vehicles, traffic operation authorities and asset management departments, etc. Figure 7 describes the applications for multiple types of users. Private drivers can efficiently capture the data including the traffic condition, infrastructure status, and environment conditions to improve their driving safety and efficiency [33]. For commercial driving, the vehicles can be controlled remotely

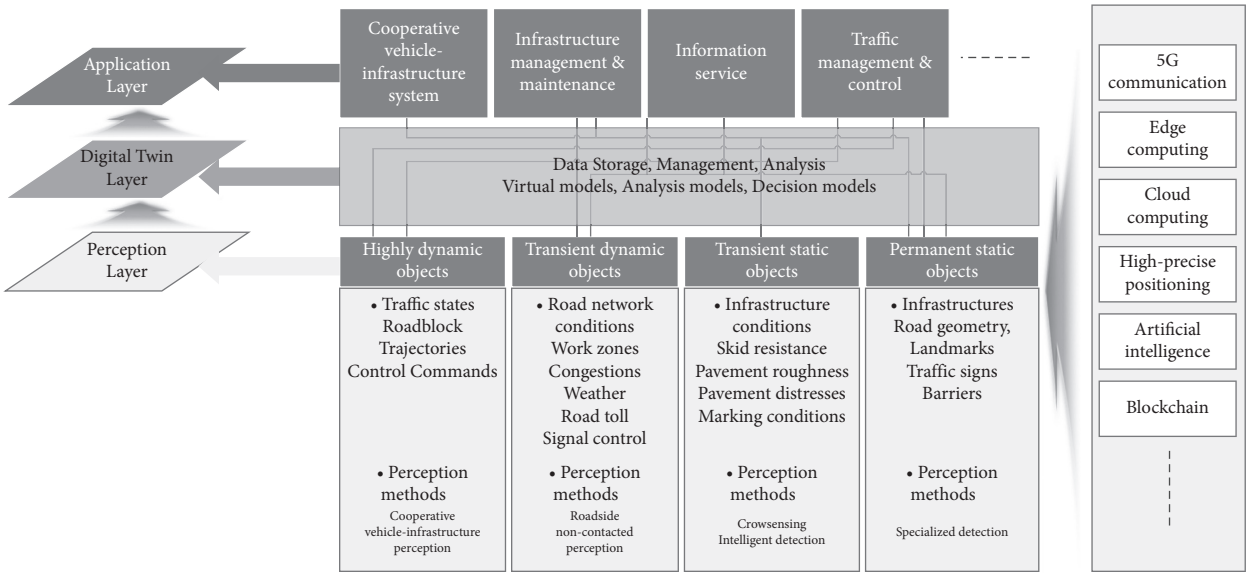


Figure 6: Physical architecture of the NGSH.

Table 2: Data requirements of the NGSH for different applications.

Applications	Highly dynamic data		Transient dynamic data		Transient static data		Permanent static data	
CVIS	Accuracy	●●●●●	Accuracy	●●●●●	Accuracy	●●●●	Accuracy	●●●●
	Dimension	●●●●●	Dimension	●●●●●	Dimension	●●●	Dimension	●●
Asset management & maintenance	Accuracy	●	Accuracy	●●●	Accuracy	●●●	Accuracy	●●●
	Dimension	●	Dimension	●●●	Dimension	●●●●●	Dimension	●●●
Information service	Accuracy	●●●●●	Accuracy	●●●●	Accuracy	●●●	Accuracy	●●
	Dimension	●●●●●	Dimension	●●●●	Dimension	●●●	Dimension	●●
Traffic management & control	Accuracy	●●●●●	Accuracy	●●●●	Accuracy	●●	Accuracy	●●
	Dimension	●●●●●	Dimension	●●●	Dimension	●●	Dimension	●●

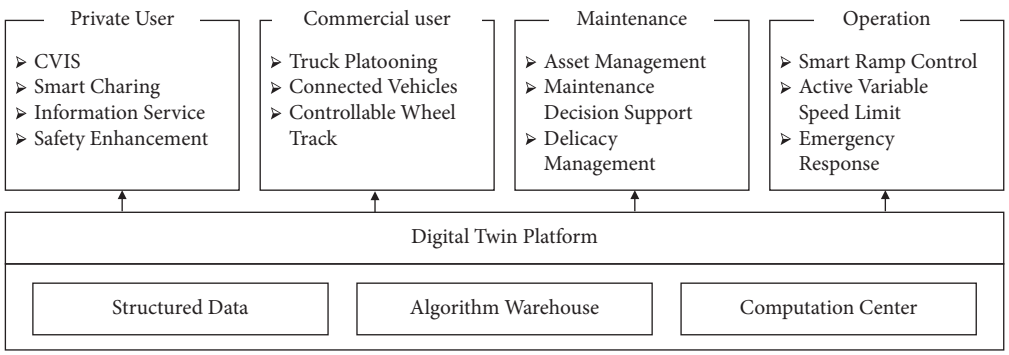


Figure 7: Applications architecture of the NGSH.

to achieve the automated driving and vehicle platooning based on the real-time traffic and environment data. Based on the infrastructure and traffic operation data, the NGSH empowered the ability of the intelligent asset management and lifecycle maintenance.

4. Development Path of New Generation of Smart Highway

4.1. Interactive Development of NGSH and Automated Vehicles. The development of NGSH is a process of mutual promotion of intelligent vehicles and smart roads, which needs to be built together with the emerging technologies. The effectiveness depends on both the vehicle's automation level and the highway's smart level.

Similar to the classification of automation level of automated vehicles (AVs), NGSH can also be divided into different smart levels according to their digital abilities. Carreras et al. established a road infrastructure support levels for AVs (ISA levels) [34]. These levels can be assigned to parts of the network in order to give AVs guidance for the coming highway automation era. The core idea about this classification scheme is to differentiate the infrastructure considering their digital functions, such as digital information, cooperative perception, and cooperative driving. The NGSH as a whole is composed of many smart infrastructures. The NGSH serves not only AVs, but also traffic operators. Therefore, the smart levels of NGSH are determined by both their data features and support functions, as shown in Table 3. Five smart levels are defined from conventional highway (road smart level 1, R1) to fully smart highway (road smart level 5, R5). The details of data attributes can refer to Table 2, and the function attributes are roughly categorized into three aspects according to ISA levels.

From the user's perspective, the combination of the vehicles and highways with different automated and smart levels can realize diverse functions, as shown in Figure 8. Without vehicle-to-road communication, even the highest-level roads cannot provide information services for L0 vehicles. However, for L5 AVs, they can achieve autonomous driving without any external infrastructure support, so they can be applied on R1~R5 highways. Similarly, cooperative driving cannot take effect without highly dynamic data. For R4 roads and L4 vehicles, vehicles perceive most the dynamic data under normal conditions, and roads can provide auxiliary sensing data of regions of interest. They achieve the complete elements awareness together, making cooperative driving possible. With the upgrade of NGSH, smart infrastructures of the highway will provide vehicles with stronger perception capabilities, empowering them with the functions of high-level autonomous driving. In some scenarios, the combinations of different levels of AVs and NGSH achieve a similar effect. For example, both 'L4 vehicles + R3 highway' and 'L2 vehicles + R5 highway' can reach autonomous driving. Therefore, the smart level of the NGSH should be rationally designed according to the automation level of the major operating vehicles in practice.

From the operator's perspective, different smart levels of NGSH improve operational capabilities to varying degrees. We concentrate on three major tasks (asset management, safety enhancement, and efficiency improvement) for identifying the effectiveness of different levels of NGSH (see Figure 9). For low-smart level highways, most operation works are conducted manually, which are time-consuming and labor-intensive. For highways with highly smart level, data can be collected with distributed sensors, and tasks can be completed by automated equipment. As most operation works do not need to know the highly dynamic data of each moving vehicle, R4 and R5 highways basically have the same functions. With the increase of the smart level of NGSH, significant economic and environmental benefits will also be generated. By reducing delay time, improving traffic efficiency, and optimizing vehicle trajectory, pollution emissions, and fuel consumption can be effectively reduced. The classification method proposed in this paper is mainly oriented to the smart and digitalization level of highways. A detailed discussion of its economic and environmental impact will be carried out in future work.

4.2. Key Issues in Current Stage for Developing NGSH. The NGSH reshapes the existing highway system by using digital and intelligent technology to improve its serviceability, such as traffic capacity, driving speed, and driving safety. The NGSH explores the potential of the current physical facilities. However, there are still many critical issues in practice, which are summarized as follows:

(1) At present, some NGSH functions are too "ahead of the time," resulting in poor compatibility with the current system. For example, some scenarios are dedicatedly designed for L4/L5 AVs. However, considering the limitations of laws and regulations, highways can only be used for the testing of autonomous driving. The deployment of too many roadside facilities for AVs hardly brings out the value of the NGSH at this stage. It is recommended to set the proper construction contents according to the regional features and predominant users. The architecture of the NGSH has to consider its sustainability to guarantee that the proposed smart highway is applicable for different development stages.

(2) The existing NGSH design and construction mainly concentrate on a single scenario and lacks overall consideration. Many projects are designed for a specific objective (i.e., autonomous driving, lifecycle maintenance, and information publication), which leads to the utility of the complete element sensing not being brought into play. It is necessary to fully integrate existing resources to create a synthetical system, which can be utilized for various scenarios.

(3) The construction of the NGSH lacks uniform data specifications and standards, resulting in poor migration and replication of the mature applications. Since the construction of NGSH is still in the exploration stage, there are still large differences in

Table 3: Smart levels of the new generation of smart highway.

Level	Name	Date feature				Function features		
R1	Conventional highway	Permanent static data	Transient static data	Transient dynamic data	Highway dynamic data	Information provision	Perception enhancement	Cooperation driving
R2	Preliminary smart highway							
R3	Partial smart highway	●	●	◦		●	◦	
R4	Conditional smart highway	●	●	●	◦		●	●
R5	Full smart highway	●	●	●	●	●	●	●
Level	Name	Description						
R1	Conditional smart highway	Conventional highway without any digital facilities						
R2	Preliminary smart highway	Highway with some precollected static data (road alignment, marks, signs, etc.), which provides basic map functions.						
R3	Partial smart highway	R2 highway + the ability of collecting the transient static data (pavement performance, maintenance, activities, etc.) and limited transient dynamic data, which provides some primary environment information for vehicles.						
R4	Conditional smart highway	R3 highway+the ability of sensing the transient dynamic data (congestion, signal, etc.) and limited highly dynamic data, which enhances the vehicles perception towards road infrastructures.						
R5	Full smart highway	R4 highway+the ability of sensing the highly dynamic data (trajectory, vehicle movements, etc.) Which provides complete elements awareness and cooperative driving. R5 highway can control and optimize the traffic flow from a macro perspective						

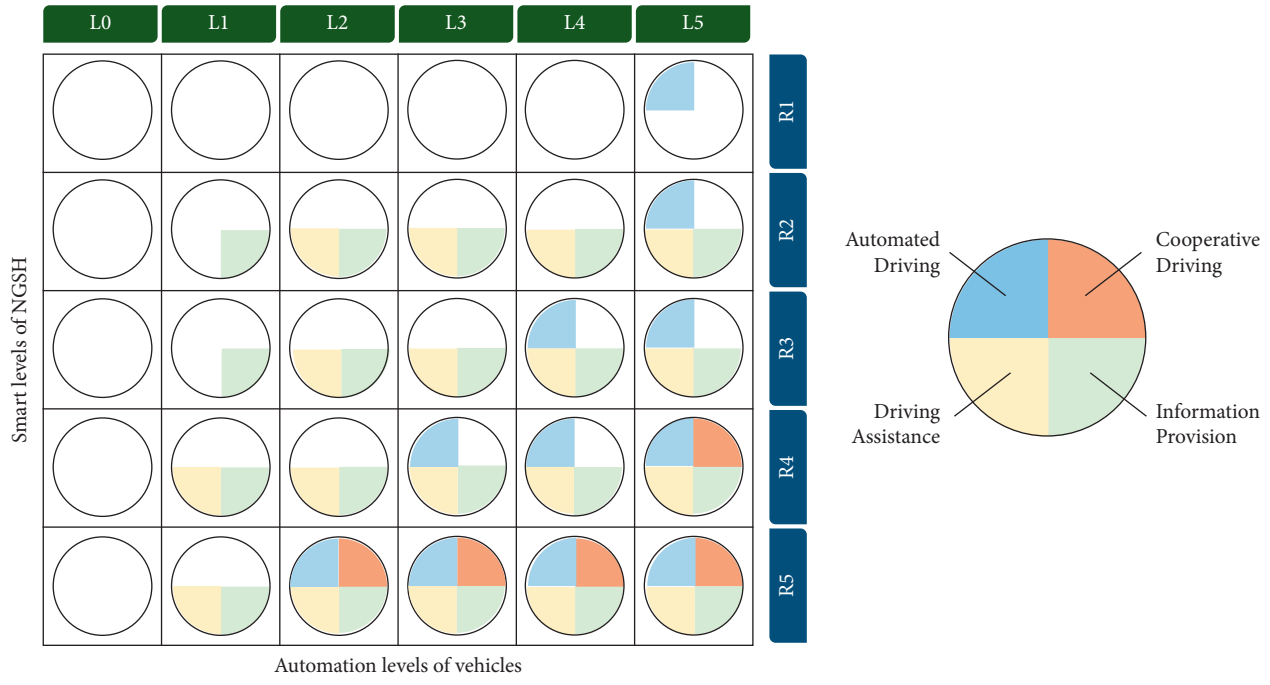


Figure 8: Combined effect of vehicle's automated level and highway's smart level.

data types, formats, protocols, and communication methods. Therefore, relevant standards should be set to ensure that the NGSH applications with different architectures can interoperate with each other.

(4) Although current NGSH functions, such as CVIS, over-the-horizon perception, and security enhancement have achieved productive research results, there are certain problems in communication stability, data security, and privacy

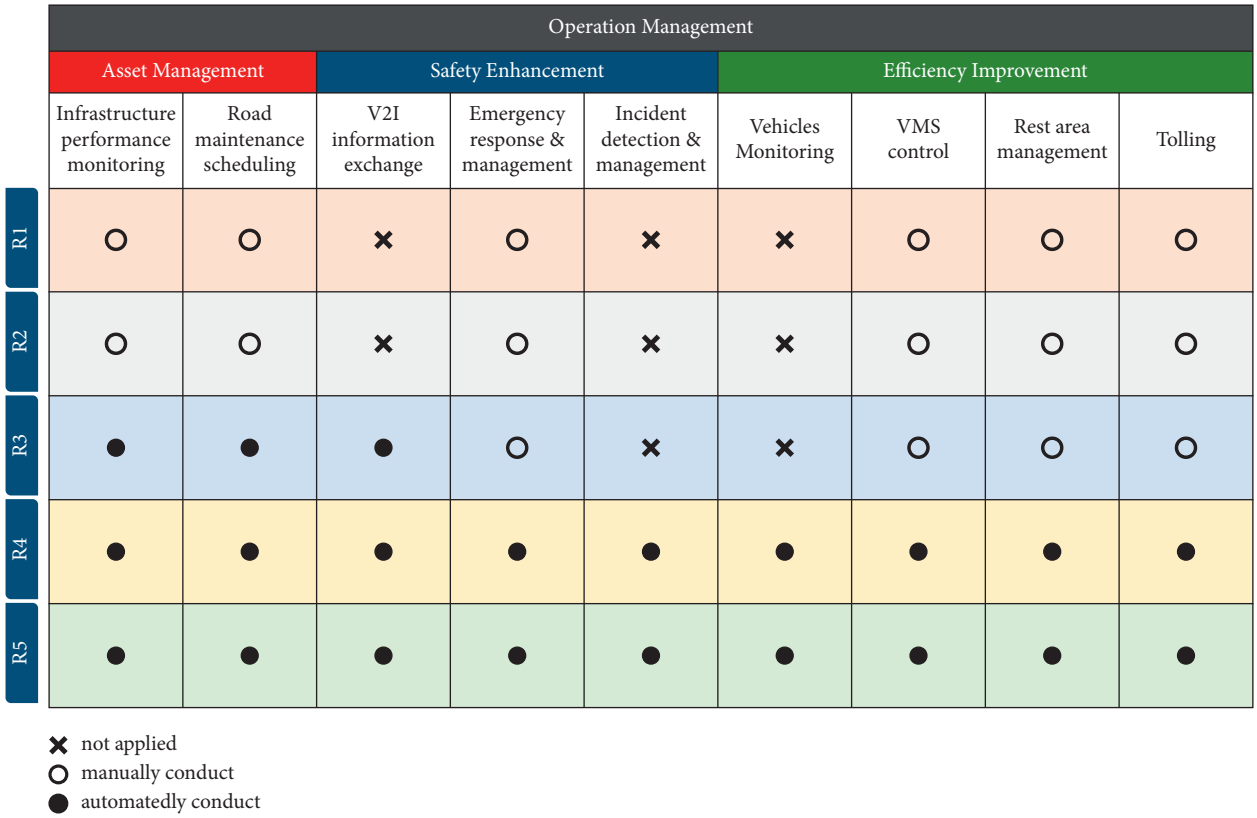


Figure 9: Operation management classification under different smart levels of NGSH.

protection, especially for the case of multiterminal interactive environment. Therefore, the study in the next stage needs to improve the robustness of the NGSH to external attacks and protect the privacy of users.

5. Conclusion

This paper proposes the framework of the new generation of the smart highway (NGSH), which is a cyber-physical system and integrates various computing, communication, and control technologies. According to the coupling development of road transport and the automotive industry, we summarize the fundamental technical features of the NGSH. The following findings and implications are drawn:

(1) The NGSH consists of four main components: complete element sensing, CPS platform, CVIS applications, and telecommunications. The core feature of NGSH is the interactive collaboration between automated vehicles and smart infrastructure.

(2) The physical architecture of the NGSH defines the basic elements and major components in the NGSH system. The data requirements for different application scenarios have been specified in terms of accuracy, dimension, and freshness.

(3) The smart levels of the NGSH are divided according to the data time-sensitive features and function features. Five smart levels are defined from conventional highway (R1) to full smart highway (R5).

(4) According to the update frequency, the CPS platform is composed of permanent static data (i.e., road geometry, landmarks), transient static data (i.e., road quality, visuality), transient dynamic data (i.e., congestion, weather), and highly dynamic data (i.e., moving vehicles, pedestrians).

(5) The CVIS applications are divided into information provision, perception enhancement, and cooperative driving, which is oriented to vehicles with different automation levels.

(6) The combined effect of the vehicle's automated level and highway's smart level is discussed, which reveals different stages of the NGSH and automated vehicle system. Meanwhile, the key issues in the current stage are illustrated from the perspective of the sustainability and economical applicability of the NGSH.

Data Availability

All of the data related to this paper are available for peer researchers to validate.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was jointly supported by Scientific Research Project of Shanghai Science and Technology Commission (20DZ2251900, 20DZ1202104, and 21DZ1200601) and National Natural Science Foundation of China (NSFC52108411 and NSFC 52008309). The authors would like to acknowledge the support provided by Jing Cao, Zihang Weng for providing relevant research data.

References

[1] P. H. Wright and R. J. Paquette, *Highway Engineering*, John Wiley and Sons Ltd, New York, NY, USA, 2nd edition, 1987.

[2] Y. Gao, J. Wenwen, and H. Chen, "Influence of highway on regional economy: a case from qingdao yinchuan expressway route," in *Proceedings of the E3S Web of Conferences*, vol. 253, Changsha, China, April 2021.

[3] Y. Yang, Z. Yuan, J. Chen, and M. Guo, "Assessment of osculating value method based on entropy weight to transportation energy conservation and emission reduction," *Environmental Engineering & Management Journal (EEMJ)*, vol. 16, no. 10, 2017.

[4] Y. Ge, X. Liu, L. Tang, and D. M. West, "Smart transportation in china and the united states," *Center for Technology Innovation*, Brookings Institution, Washington, DC, USA, 2017.

[5] D. Guo, W. Zhou, A. Sha, and R. Bai, "Application of Uncertainty Analytic Hierarchy Process Method for Asphalt Pavement Construction Quality Control in China," *Transportation research record*, vol. 2098, no. 1, pp. 43–50, 2009.

[6] R. Whelan, *Smart Highways, Smart Cars*, Artech House Publishers, Norwood, MA, USA, 1995.

[7] W. B. Stevens, "The automated highway system program: a progress report," *IFAC Proceedings Volumes*, vol. 29, no. 1, pp. 8180–8188, 1996.

[8] I. L. Al-Qadi, A. Loulizi, M. Elseifi, and S. Lahouar, "The Virginia Smart Road: the impact of pavement instrumentation on understanding pavement performance," *Journal of the Association of Asphalt Paving Technologists*, vol. 73, no. 3, pp. 427–465, 2004.

[9] A. Pande, *Evaluation of Adaptive Ramp Metering on I-80 in the San Francisco Bay Area*, The University of Texas at Arlington, Arlington, TX, USA, 2018.

[10] A. Saroj, S. Roy, A. Guin, M. Hunter, and R. Fujimoto, "Smart City Real-Time Data-Driven Transportation Simulation," in *Proceedings of the 2018 Winter Simulation Conference (WSC)*, pp. 857–868, Gothenburg, Sweden, December 2018.

[11] S. An, B.-H. Lee, and D.-R. Shin, "A survey of intelligent transportation systems," in *Proceedings of the 2011 Third International Conference on Computational Intelligence, Communication Systems and Networks*, pp. 332–337, Bali, Indonesia, July 2011.

[12] A. Fujimoto, S. Koichi, O. Michiya et al., "Toward realization of Smartway in Japan," in *Proceedings of the 15th World Congress on Intelligent Transport Systems and ITS America's 2008 Annual Meeting*, New York, NY, USA, 2008.

[13] R. Imai, T. Matsushima, and S. Kanai, "Proposal of automatic correction method of incorrect origin-destination in etc 2.0 probe data," *Journal of Japan Society of Civil Engineers, Ser. F3 (Civil Engineering Informatics)*, vol. 74, 2018.

[14] C. Hildebrandt, T. Bandyszak, A. Petrovska, N. Laxman, E. Cioroaica, and S. Törsleff, "EURECA: epistemic uncertainty classification scheme for runtime information exchange in collaborative system groups," *SICS Software-Intensive Cyber-Physical Systems*, vol. 34, no. 4, pp. 177–190, 2019.

[15] F. R. Soriano, V. R. Tomás, and M. Pla-Castells, "Deploying harmonized ITS services in the framework of EasyWay project: traffic Management Plan for corridors and networks," in *Proceedings of the in 2012 6th Euro American Conference on Telematics and Information Systems (EATIS)*, pp. 1–7, Valencia, Spain, May 2012.

[16] T. Victor, M. Rothoff, E. Coelingh, A. Ödblom, and K. Burgdorf, "When autonomous vehicles are introduced on a larger scale in the road transport system: the Drive Me project," in *Automated Driving*Springer, New York, NY, USA, 2017.

[17] J. Erhart, A. Carreras, X. Daura et al., "Novel approaches for analysing and testing the effect of autonomous vehicles on the traffic flow," in *Proceedings of the 26th ITS World Congress (ITSCWC 2019)*, Singapore, October 2019.

[18] Y. Du, B. Qin, C. Zhao, Y. Zhu, J. Cao, and Y. Ji, "A novel spatio-temporal synchronization method of roadside asynchronous MMW radar-camera for sensor fusion," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[19] S. Fu, L. Wenhong, J. Ge, and Y. Qu, "Review on the application of freeway CVIS communication technology," in," *MATEC Web of Conferences*, vol. 325, p. 01006, 2020.

[20] L. Sun, H. Zhao, H. Tu, and Y. Tian, "The smart road: practice and concept," *Engineering*, vol. 4, no. 4, pp. 436-437, 2018.

[21] E. H. Wakefield, "History of the electric automobile-hybrid electric vehicles," vol. 187, SAE International, Warrendale, PA, USA, 1998.

[22] H. Niemann, *The Mercedes-Benz History: Automobile Legends and Stories since 1886*, Mercedes-Benz Classique Car Library, Stuttgart, Germany, 2007.

[23] R. Snedden, *Roads*, The Rosen Publishing Group, Inc, New York, NY, USA, 2016.

[24] R. P. Roess and E. S. Prassas, *The Highway Capacity Manual: A Conceptual and Research History*, Springer, New York, NY, USA, 2014.

[25] M. G. Lay, *Ways of the World: A History of the World's Roads and of the Vehicles that Used Them*, Rutgers university press, New Brunswick, NJ, USA, 1992.

[26] W. K. Kittelson and R. P. Roess, "Highway capacity analysis after highway capacity manual 2000," *Transportation Research Record*, vol. 1776, no. 1, pp. 10–16, 2001.

[27] L. Zhao and J. Sun, "Simulation framework for vehicle platooning and car-following behaviors under connected-vehicle environment," *Procedia-Social and Behavioral Sciences*, vol. 96, pp. 914–924, 2013.

[28] S. O.-R. A. D. Committee, "S. A. E. J3016. taxonomy and definitions for terms related to driving automation systems for On-road motor vehicles," Technical Report, SAE International, Warrendale, PA, USA, 2016.

[29] R. Baheti and H. Gill, "Cyber-physical systems," *The Impact of Control Technology*, vol. 12, no. 1, pp. 161–166, 2011.

[30] M. Batty, *Digital Twins*, SAGE Publications Sage UK, London, England, 2018.

[31] C.-R. Rad, O. Hancu, I.-A. Takacs, and G. Olteanu, "Smart monitoring of potato crop: a cyber-physical system architecture model in the field of precision agriculture," *Agriculture and Agricultural Science Procedia*, vol. 6, pp. 73–79, 2015.

[32] N. Al-Falahy and O. Y. Alani, "Technologies for 5G networks: challenges and opportunities," *IT Professional*, vol. 19, no. 1, pp. 12–20, 2017.

[33] Y. Yang, Z. Z. Yuan, D. Y. Sun, and X. L. Wen, "Analysis of the factors influencing highway crash risk in different regional types based on improved Apriori algorithm," *Advances in Transportation Studies*, vol. 49, pp. 165–178, 2019.

[34] A. Carreras, X. Daura, J. Erhart, and S. Ruehrup, "Road infrastructure support levels for automated driving," in *Proceedings of the ITS World CongressAt: Copenhagen*, EU-TP1488, Copenhagen, Denmark, December 2018.

Hindawi
Journal of Advanced Transportation
Volume 2021, Article ID 3207025, 19 pages
https://doi.org/10.1155/2021/3207025

Research Article

Optimizing Customized Transit Service considering Stochastic Bus Arrival Time

Qian Sun,[1] **Steven Chien**,[1,2] **Dawei Hu**,[1] **and Xiqiong Chen**[1]

[1]*School of Transportation Engineering, Chang'an University, Xi'an 710064, China*
[2]*John A. Reif Jr. Department of Civil and Environmental Engineering, New Jersey Institute of Technology, Newark, NJ 07102, USA*

Correspondence should be addressed to Dawei Hu; dwhu@chd.edu.cn

Received 2 September 2021; Revised 12 October 2021; Accepted 6 November 2021; Published 10 December 2021

Academic Editor: Lei Wang

The introduction of customized bus (CB) service intends to expand and elevate existing transit service, which offers an efficient and sustainable alternative to serve commuters. A probabilistic model is proposed to optimize CB service with mixed vehicle sizes in an urban setting considering stochastic bus arrival time and spatiotemporal demand, which minimizes total cost subject to bus capacity and time window constraints. The studied optimization problem is combinatorial with many decision variables including vehicle assignment, bus routes, timetables, and fleet size. A heuristic algorithm is developed, which integrates a hybrid genetic algorithm (HGA) and adaptive destroy-and-repair (ADAR) method. The efficiency of HGA-ADAR is demonstrated through numerical comparisons to the solutions obtained by LINGO and HGA. Numerical instances are carried out, and the results suggested that the probabilistic model considering stochastic bus arrival time is valuable and can dramatically reduce the total cost and early and late arrival penalties. A case study is conducted in which the proposed model is applied to optimize a real-world CB service in Xi'an, China. The relationship between decision variables and model parameters is explored. The impacts of time window and variance of bus arrival time, which significantly affect service reliability, are analysed.

1. Introduction

Public transportation agencies are experimenting with on-demand and shared technologies to augment traditional fixed-route bus services, such as flexible-route bus services [1], demand responsive transit (DRT) [2], ridesharing, microtransit [3], and customized bus (CB) service [4]. These new, flexible transit solutions have tremendous potential to expand agencies' service areas, attract new riders, fill transportation gaps, and provide more effective and sustainable ways to reach low-density communities and other traditionally hard-to-serve situations.

As an emerging transportation mode, CB is similar to microtransit, which offers another option between the pricey convenience of taxis and slow, cheaper public transit. CB service is focused on providing commuters with fast and comfortable commuting buses in morning and evening rush hours. It allows commuters to make reservations for trips by submitting the information related to spatiotemporal demand (i.e., origination and destination (OD) and desired arrival time) and arranges buses to serve the shared commuters. The service plan is updated on a short-term basis subject to the passenger demand change. To ensure service quality, CB does not take walk-in riders without reservation [5].

The major challenges that need to be addressed by operators are the following problems: how many buses to be used and how to determine bus types, assignments, routes, and timetables to serve passengers. To comprehensively assess the feasibility and cost-effectiveness of CB service, the operators need to take a holistic approach that can jointly determine vehicle assignment, routes, timetables, and fleet size with the spatiotemporal passenger demand.

In addition, bus travel time on the route as well as bus arrival time at the stop is affected by recurring and non-recurring congestion in urban settings. Optimizing service planning with deterministic travel time might underestimate the cost and result in an unexpected level of service.

Moreover, the CB customers would expect to arrive at destinations within the acceptable time windows, which makes it challenging for optimizing bus routing, scheduling, and timetabling.

This paper aims to optimize the CB service with a mixed bus fleet to minimize total cost, considering stochastic bus arrival time, which is known to be an NP-hard problem. A hybrid heuristic algorithm that integrates the features of genetic algorithm (GA), simulated annealing (SA), and adaptive destroy-and-repair (ADAR) is proposed to efficiently search for the approximate optimal solution. The HGA hybridizing GA and SA [5] performs well in global exploration but spends much time converging to the optimal solution. In contrast, local search methods, such as ADAR, can find the optimal solution quickly. So, we develop an HGA-ADAR algorithm by incorporating ADAR into HGA to improve its performance to reach the optimal solution in the region of convergence.

In summary, the main contributions of this study are as follows:

(1) An important consideration of CB service planning is addressed and integrated in our model, namely, stochastic bus arrival time. Furthermore, some practical issues, such as mixed fleet size and multi-terminal, are also considered. The proposed model can optimize bus assignment, routes and timetables, and mixed fleet size simultaneously.

(2) A hybrid heuristic algorithm that integrates the features of GA, SA, and ADAR is proposed to efficiently search for the approximate optimal solution.

(3) The relationship between model parameters and decision variables is explored, which may provide some decisions and strategy supports for operating CB service in the future.

The rest of this paper is structured as follows. Section 2 summarizes an overview of related existing literature. Section 3 describes the developed probabilistic model, including the objective function and associated constraints. The developed HGA-ADAR is discussed in Section 4, while the model and algorithm performance and efficiency are assessed in Section 5. Section 6 presents a case study, in which the applicability of the developed model is presented, and the optimized results are discussed. The relationship between decision variables and model parameters is explored in the results of sensitivity analysis. Finally, the research findings with a future study are presented in Section 7.

2. Literature Review

The discussion of the literature review covers previous studies in transit planning with deterministic bus arrival time, vehicle routing problem with stochastic travel time, bus arrival time distribution, and feasible solution algorithms.

2.1. Transit Planning with Deterministic Bus Arrival Time. Considering deterministic bus arrival time, Chien [6] optimized route choice, headway, and bus capacity for a feeder bus system, which minimized total cost. Ulusoy et al. [7] and Ulusoy and Chien [8] optimized the integration of all-stop, short-turn, and express service considering heterogeneous demand. Later, Qu et al. [9] enhanced the model by taking time-dependent demand into account. Chowdhury and Chien [10] optimized fare and headway considering demand elasticity that maximized total profit. Chen et al. [11] simultaneously optimized bus routing, operating frequency, wireless power device locations, and battery capacity for a multi-route bus network, which minimized total cost.

Several previous studies focusing on optimizing CB service considered deterministic bus arrival time. Tong et al. [12] optimized passenger assignment and routes considering spatiotemporal windows that maximized ridership. Guo et al. [13] proposed a model for optimizing passenger assignment and routing that minimized the total cost, and later they enhanced the model by considering time window [14] and time-dependent bus arrival time and path flexibility between stops [15]. Lyu et al. [16] introduced a planning framework to optimize stop location, routing, timetabling, and passengers' probability of choosing CB simultaneously, which maximized total profit. Huang et al. [17] modelled the decision-making process for CB service and optimized passenger assignment and bus routing that maximized total profit. Sun et al. [5] optimized trip assignment, routing, timetabling, and bus fleet size for CB service considering the mixed fleet size and multi-terminal, which minimized total cost.

In summary, to the best of our knowledge, the optimization of CB service considering stochastic bus arrival time has not been studied in the above studies. This variant constitutes an important extension to the classical CB service planning that has not been investigated in the literature for which we believe it is worth exploring.

2.2. Vehicle Routing Problem with Stochastic Travel Time. In the field of operational research, the optimization of CB service considering stochastic bus arrival time is an extension of stochastic vehicle routing problem (SVRP) [18, 19]. Gendreau et al. [20, 21] presented some comprehensive surveys on SVRP and made a classification of the articles relying on different stochastic parameters. The authors argued that the common stochastic parameters considered are stochastic demand, stochastic customer, and stochastic travel time. In this paper, we focus on the consideration of stochastic travel time. Thus, we reviewed the literature related to vehicle routing problem with stochastic travel time (VRPSTT) emphatically.

Kenyon and Morton [18] optimized VRPSTT considering two objectives: minimizing the expected completion time that all vehicles will return to the depot and maximizing the probability that the trip is complete on or before a prespecified deadline. Li et al. [19] optimized an VRPSTT, which minimized total transportation cost. It was found that evaluating the objective function of stochastic problems would consume lots of computation time. Adulyasak and Jaillet [22] studied an VRPSTT with soft time windows, where the service at each customer should be within a

predetermined time window or a violation occurs. The problem aimed to construct the routes to minimize the sum of probability of time window violations. Errico et al. [23] solved an VRPSTT with hard time windows, where a penalty is paid whenever the service at a customer is skipped. The objective function was to minimize the total expected travel cost and penalty cost.

The optimization of CB service is a more complex problem than the conventional VRP because it incorporates several specific model constraints, such as time windows, bus capacity, pairing, and precedence [24]. However, previous models discussed above focused on general VRPSTT, which are not suitable for CB service. Besides, the trade-off between operator cost and user cost is the main consideration in CB service, which is different from general VRPSTT.

2.3. Bus Arrival Time Distribution. Bus arrival time inherently fluctuates due to various interruptions caused by traffic signals, accidents, unplanned road works, and adverse weather [25–27]. The distributions of bus arrival time were investigated and assumed that they follow normal [18, 19, 22, 28–30], log-normal [31–33], and gamma [34, 35] distributions.

Assuming normally distributed bus arrival time, Chen et al. [36] optimized limited-stop bus service (i.e., stop-skipping), which minimized total cost. Results found that the randomness of bus arrival time should not be ignored for transit planning. Xiao et al. [37] optimized scheduled arrival time and slack time for a transfer hub of a bus transit considering the probabilistic bus travel time and walking speed, which minimized total system cost.

Considering log-normally distributed bus arrival time, Chien et al. [31] proposed a probabilistic model for optimizing disseminated bus arrival time for pretrip passengers, which minimized total wait time. Prakash [32] studied on determining the most reliable routes on the stochastic and time-dependent network, which maximized on-time arrival probability.

Taş et al. [34] developed a model for optimizing an SVRPSTT with soft time windows considering bus arrival time following a gamma distribution, which minimized the sum of transportation and service costs. Later, they enhanced that model, considering time-dependent stochastic bus arrival time [38].

2.4. Solution Algorithm. As discussed above, the studied CB optimization problem is an extension of VRP. Thus, it belongs to the NP-hard problem. An efficient algorithm is desired to search for the optimal or approximate optimal solution. Genetic algorithm (GA) is generally employed to deal with the NP-hard problem [5, 39]. Sun et al. [5] used a hybrid GA to optimize the routing and scheduling for a CB service. Chen et al. [11] developed a nested GA for planning an electric feeder system. GA has a great ability of global search in solution space, but it does not employ a local search and spends much time to converge to the optimal solution. So, it needs some modifications to improve the convergence and reduce the computation time.

The neighbourhood search method, such as ADAR, which is a local search-based strategy, has shown excellent performance in solving routing problems [40] and finds the optimal solution very quickly. Kitjacharoenchai et al. [41] developed an ADAR method that employs three destroy operators and three repair operators to solve a VRP. Dong et al. [42] optimized stop plans and timetables for commuter railways with an ADAR method by employing multiple destroy and repair operators. It was found that the local search-based strategy is easy to integrate with other algorithms to effectively find the optimal solution [43–45]. Soto et al. [43] proposed a local search hybridized with a tabu search (TS) to deal with a multi-depot VRP. Moshref-Javadi et al. [45] proposed a metaheuristic algorithm, in which several local search strategies were used to improve the solution obtained by SA, and it was proved to be able to solve VRP efficiently. Masmoudi et al. [46] introduced a hybrid GA to deal with a DARP by integrating local search. Belhaiza [44] designed a heuristic that integrated GA and ADAR to solve a DARP. Baniamerian et al. [47] proposed a hybrid heuristic algorithm based on variable neighbourhood search (VNS) and GA.

Based on the above literature review, we proposed a probabilistic optimization model to assist a transit agency in planning a CB service in the context of the stochastic environment. To the best of our knowledge, there are no studies on CB service optimization problems considering stochastic bus arrival times. Table 1 presents a comparison of our study with previous studies. It is found that the studied optimization problem is combinatorial with many decision variables including vehicle assignment, bus routes, timetables, and fleet size. Unlike previous studies, some practical issues, such as mixed fleet, multi-terminal, time window, and stochastic bus arrival time, are also considered. None of the previous research has covered all the above issues. Therefore, this study will fill the gap in comprehensive optimization for CB service and provide a significant basis for new areas of research to explore in the future.

3. Methodology

In this paper, a probabilistic model is formulated to optimize a CB service (i.e., vehicle assignment, routing, scheduling, and mixed fleet size) considering stochastic bus arrival time, which minimizes the total cost including operator cost, user cost, and early and late arrival penalties subject to some practical constraints. It is worth noting that the passengers need to be classified into groups before the CB service is optimized, and we adopt the passenger partition approach given in Sun et al. [5]. In the research of Sun et al., the authors partitioned passengers into groups based on the information submitted by passengers (i.e., OD information and desired arrival time) and a prespecified time interval.

3.1. Assumptions. To formulate the developed model, the following assumptions were made:

(1) The spatiotemporal passenger demand of the study area is known.

Table 1: Comparison of this study with previous studies.

Reference	Problem	Objective	Decision variables				Constraints						Solution algorithm
			R	*S*	*V*	*F*	TW	PP	MT	MB	VC	ST	
Li et al. [19]	VRP	Min. operator cost	√	√	√		√				√	√	Tabu search (TB)
Taş et al. [34]	VRP	Min. total cost	√	√	√		√				√	√	TB
Chen et al. [36]	PT	Min. total cost	√	√							√	√	Hybrid artificial bee colony (ABC)
Tong et al. [12]	CB	Max. ridership	√	√	√	√	√	√			√		Lagrangian decomposition
Han et al. [48]	CB	Min. total cost	√			√		√		√	√		Branch and bound (B&B)
Chen et al. [11]	PT	Min. total cost	√			√					√		Nested genetic algorithm (NGA)
Huang et al. [17]	CB	Max. profit	√	√	√	√	√	√			√		B&B
Han et al. [49]	CB	Min. travel distance	√	√	√	√	√	√	√		√		Branch-and-price
Dou et al. [50]	CB	Max. profit	√	√	√	√	√	√		√	√		Branch-and-price
Guo et al. [15]	CB	Min. total cost	√	√	√		√	√			√	√	Tabu search-variable neighbourhood search (TS-VNS)
Shen et al. [51]	CB	Min. total cost	√	√	√	√	√	√			√		Column generation algorithm
This study (2021)	CB	Min. total cost	√	√	√	√	√	√	√	√	√	√	HGA-ADAR

Note. *R*: route; *S*: schedule; *V*: vehicle assignment; *F*: fleet size; TW: time window; PP: pairing and precedence; MT: multi-terminal; MB: mixed bus fleet; VC: vehicle capacity; ST: stochastic bus arrival time.

(2) Bus sizes and associated information are given (i.e., capacity and operating cost).

(3) Buses are dispatched from a terminal and return to a different one after finishing a trip.

(4) In the advent of real-time information of bus arrivals, passengers' wait time at pick-up stops can be ignored.

The model is described on a complete graph G, which consists of a set of nodes denoted as N and a set of links denoted as A. N consists of a set of terminals denoted as W, a set of pick-up stops denoted as N^+, and a set of drop-off stops denoted as N^-. W, N^+, and N^- are mutually exclusive. Table 2 lists the parameters and variables used to formulate the proposed model.

3.2. Objective Function. The objective function minimizes total cost denoted as ZT which consists of operator cost denoted as ZO, user in-vehicle cost denoted as ZU, and early and late arrival penalties caused by stochastic bus arrival time at drop-off stops denoted as ZP. Thus,

$$\min \mathrm{ZT} = \mathrm{ZO} + \mathrm{ZU} + \mathrm{ZP}. \tag{1}$$

ZO consists of variable cost denoted as ZV and fixed cost denoted as ZF. ZV is the product of average variable cost per bus-hour v_b and travel time $\sum_{i\in N}\sum_{j\in N}\sum_{b\in B} t_{ij}x_{ijb}$, while ZF is the product of fixed cost per bus-hour f_b, number of buses used $\sum_{w\in W}\sum_{j\in N}\sum_{b\in B} x_{wjb}$, and service time horizon o.

$$\begin{aligned}\mathrm{ZO} &= \mathrm{ZV} + \mathrm{ZF} \\ &= \sum_{i\in N}\sum_{j\in N}\sum_{b\in B} v_b t_{ij} x_{ijb} + o\sum_{w\in W}\sum_{j\in N}\sum_{b\in B} f_b x_{wjb},\end{aligned} \tag{2}$$

where x_{ijb} is a binary variable, and $x_{ijb} = 1$ if a bus of type b serves link (i, j); otherwise, $x_{ijb} = 0$.

ZU is the product of value of time λ_3, in-vehicle time, and the number of passengers served q_{ij}. The in-vehicle time from pick-up stop i to drop-off stop j by a bus of type b is denoted as $(\widetilde{T}_{jb} - \widetilde{T}_{ib})$.

$$\mathrm{ZU} = \lambda_3 \sum_{i\in N^+}\sum_{j\in N^-}\sum_{b\in B}\left(\widetilde{T}_{jb} - \widetilde{T}_{ib}\right) q_{ij} y_{ib} y_{jb}, \tag{3}$$

where y_{ib} is a binary variable, and $y_{ib} = 1$ if stop i is served by a bus of type b; otherwise, $y_{ib} = 0$. $\widetilde{T}_{jb}$ is the arrival time of a bus of type b at stop j, which follows a probability distribution (i.e., $f(\widetilde{T}_{jb})$), as shown in Figure 1.

It is assumed that the expected bus arrival time at stop j is T'_j, and the acceptable arrival time window is tw. The earliest and latest arrival times within the time window are e_j and l_j, respectively (i.e., $e_j = T'_j - \mathrm{tw}/2$, $l_j = T'_j + \mathrm{tw}/2$). Figure 1 shows a probability distribution of arrival time for a bus of type b at stop j (i.e., $f(\widetilde{T}_{jb})$). Bus arrivals outside the time window are classified into (1) early arrival represented by area I where the bus arrives before the start of time window e_j (e.g., $\widetilde{T}_{jb} < e_j$) and (2) late arrival represented by area II where the bus arrives after the end of time window l_j (e.g., $\widetilde{T}_{jb} > l_j$).

The penalty incurred by early and late bus arrivals is denoted as ZP that can be determined based on the probability of bus arrival time at drop-off stops. For example, penalty caused by the early arrival of a bus of type b at stop j, denoted as PE_{jb}, is the product of unit penalty for early arrival λ_1, number of passengers alighting from stop j denoted as $|Q_{jb}|$, the probability of early arrival $\int_0^{e_j} f(\widetilde{T}_{jb})\mathrm{d}(\widetilde{T}_{jb})$, and early arrival time. The early arrival time is the elapsed time from the actual arrival time to the start of the time window denoted as $(e_j - \widetilde{T}_{jb})$. Thus,

$$\mathrm{PE}_{jb} = \lambda_1 \left|Q_{jb}\right| \int_0^{e_j} \left(e_j - \widetilde{T}_{jb}\right) f\left(\widetilde{T}_{jb}\right) \mathrm{d}\left(\widetilde{T}_{jb}\right). \tag{4}$$

Similarly, the penalty for late arrival of a bus of type b at stop j denoted as PL_{jb} is formulated as

$$\mathrm{PL}_{jb} = \lambda_2 \left|Q_{jb}\right| \int_{l_j}^{\infty} \left(\widetilde{T}_{jb} - l_j\right) f\left(\widetilde{T}_{jb}\right) \mathrm{d}\left(\widetilde{T}_{jb}\right), \tag{5}$$

where λ_2 is the unit penalty for late arrival.

Finally, ZP represents the total early and late arrival penalties. Thus,

$$\begin{aligned} \mathrm{ZP} &= \sum_{j\in N^-} \sum_{b\in B} \mathrm{PE}_{jb} + \sum_{j\in N^-} \sum_{b\in B} \mathrm{PL}_{jb} \\ &= \lambda_1 \sum_{j\in N^-} \sum_{b\in B} \left|Q_{jb}\right| \int_0^{e_j} \left(e_j - \widetilde{T}_{jb}\right) f\left(\widetilde{T}_{jb}\right) \mathrm{d}\left(\widetilde{T}_{jb}\right) \\ &\quad + \lambda_2 \sum_{j\in N^-} \sum_{b\in B} \left|Q_{jb}\right| \int_{l_j}^{\infty} \left(\widetilde{T}_{jb} - l_j\right) f\left(\widetilde{T}_{jb}\right) \mathrm{d}\left(\widetilde{T}_{jb}\right). \end{aligned} \tag{6}$$

To minimize ZT, several constraints (e.g., vehicle assignment, routing, scheduling, and bus capacity) are taken into account. Vehicle assignment constraints are formulated in equations (7)–(9). Equation (7) states that each stop might be served by at least one bus, while equations (8) and (9) ensure that the dispatched bus of type b must be linked to at most one start (and end) terminal.

$$\sum_{b\in B} y_{ib} \geq 1, \quad \forall i \in N^+ \cup N^-, \tag{7}$$

$$\sum_{w\in W} S_{wb} \leq 1, \quad \forall b \in B, \tag{8}$$

$$\sum_{w\in W} E_{wb} \leq 1, \quad \forall b \in B. \tag{9}$$

Bus routing constraints are stated in equations (10)–(12). Equations (10) and (11) make sure that each route begins from and ends at a terminal. Equation (12) represents a route continuity constraint.

$$\sum_{j\in N} x_{wjb} = S_{wb}, \quad \forall w \in W, b \in B, \tag{10}$$

$$\sum_{j\in N} x_{jwb} = E_{wb}, \quad \forall w \in W, b \in B, \tag{11}$$

$$\sum_{\substack{j\in N \\ j\neq i}} x_{ijb} = \sum_{\substack{j\in N \\ j\neq i}} x_{jib} = y_{ib}, \quad \forall i \in N^+ \cup N^-, b \in B. \tag{12}$$

Equation (13) is subtour elimination constraint. u_{ib} and u_{jb} are auxiliary variables, while n represents the number of all stops. If stop j is an immediate downstream stop of stop i served by a bus of type b, x_{ijb} is equal to 1, which implies that $u_{jb} \geq u_{ib} + 1 > u_{ib}$. A subtour $(i, j, \ldots, i)$ without a terminal will be prevented by $u_{ib} > \cdots > u_{jb} > u_{ib}$.

$$u_{ib} - u_{jb} + n x_{ijb} \leq n - 1, \quad \forall i, j \in N^+ \cup N^-, i \neq j, b \in B. \tag{13}$$

Bus capacity constraints are defined in equations (14)–(17). Equation (14) represents that number of in-vehicle passengers is zero as the bus departs from a terminal. Q_{jb} in equation (15) is positive if stop j is a pick-up stop; otherwise, negative if stop j is a drop-off stop. Equation (16) tracks the number of in-vehicle passengers according to the sequence of the visited stops and the number of passengers boarding/alighting at each stop. Equation (17) makes sure that the number of in-vehicle passengers shall not exceed bus capacity.

$$L_{wb} = 0, \quad \forall w \in W, b \in B, \tag{14}$$

$$Q_{jb} = \begin{cases} \displaystyle\sum_{i\in N^-} q_{ji} y_{ib} y_{jb}, & j \in N^+, b \in B, \\ \displaystyle(-1) \sum_{i\in N^+} q_{ij} y_{ib} y_{jb}, & j \in N^-, b \in B, \end{cases} \tag{15}$$

$$\begin{aligned} L_{jb} \geq \left(L_{ib} + Q_{jb}\right) - M\left(1 - x_{ijb}\right), \\ \forall i \in N, j \in N^+ \cup N^-, b \in B, \end{aligned} \tag{16}$$

$$0 \leq L_{ib} \leq C_b y_{ib}, \quad \forall i \in N, b \in B. \tag{17}$$

Equation (18) restricts that bus arrival time at node j must greater than that at the previous node i, while equation (19) ensures that each pick-up and drop-off stop pair is served by the same bus and the pick-up stop is visited before its drop-off stop.

$$\widetilde{T}_{jb} \geq \widetilde{T}_{ib} + t_{ib} + t_{ij} - M\left(1 - x_{ijb}\right), \quad \forall i \in N, j \in N, b \in B, \tag{18}$$

$$\left(\widetilde{T}_{jb} - \widetilde{T}_{ib}\right) q_{ij} y_{ib} y_{jb} \geq 0, \quad \forall i \in N^+, j \in N^-, b \in B. \tag{19}$$

Finally, decision variables on routing and scheduling are defined in equations (20)–(22).

$$x_{ijb} \in \{0, 1\}, \quad \forall i, j \in N, b \in B, \tag{20}$$

$$y_{ib} \in \{0, 1\}, \quad \forall i \in N, b \in B, \tag{21}$$

$$\widetilde{T}_{ib} \geq 0, \quad \forall i \in N, b \in B. \tag{22}$$

4. Solution Algorithm

The studied CB problem is an extension of VRP. Thus, it belongs to the NP-hard problem which is difficult to solve by exact algorithms, especially for large networks. An HGA-ADAR algorithm is developed, which starts with an initial solution generated by HGA followed by performing the developed ADAR to effectively find the solution.

Genetic algorithm (GA) has been widely and successfully applied to solve various vehicle routing and scheduling

Table 2: Notations.

Set	Definition	Unit
A	Set of links	—
B	Set of bus types	—
N^+	Set of pick-up stops	—
N^-	Set of drop-off stops	—
N	Set of all nodes	—
W	Set of terminals	—
	Parameter	
C_b	Capacity of a bus of type b	seat/veh
e_i	Start of time window at drop-off stop i	hh:mm
f_b	Fixed cost of a bus of type b per hour	\$/veh-hr
l_i	End of time window at drop-off stop i	hh:mm
M	A large positive number	—
n	No. of all stops	stops
o	Service time horizon	hr
q_{ij}	Passenger demand from stop i to j	Pass/day
t_{ib}	Dwell time of a bus of type b at stop i	hr
t_{ij}	Travel time from stop i to j	hr
T_i'	Expected bus arrival time at stop i	hh:mm
v_b	Average variable cost of a bus of type b per hour	\$/veh-hr
λ_1	Unit penalty for early arrival	\$/pass-hr
λ_2	Unit penalty for late arrival	\$/pass-hr
λ_3	Value of passengers' travel time	\$/pass-hr
	Variable	
E_{wb}	Binary var. $E_{wb}=1$, a bus of type b ends at terminal w; $E_{wb}=0$, otherwise	—
L_{ib}	No. of passengers in a bus of type b leaving from node i	Pass
Q_{ib}	No. of passengers boarding/alighting of a bus of type b at stop i	Pass
S_{wb}	Binary var. $S_{wb}=1$, a bus of type b begins from terminal w; $S_{wb}=0$, otherwise	—
u_{ib}	Auxiliary variable	—
x_{ijb}	Binary var. $x_{ijb}=1$, a bus of type b serves link (i, j); $x_{ijb}=0$, otherwise	—
y_{ib}	Binary var. $y_{ib}=1$, stop i is served by a bus of type b; $y_{ib}=0$, otherwise	—
T_{ib}	Arrival time of a bus of type b at node i	hh:mm

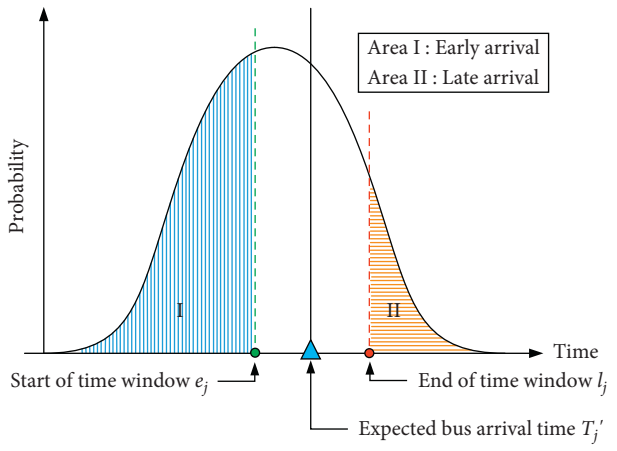


Figure 1: Probability distribution of arrival time for bus b at stop j.

problems, which is an important stochastic global search algorithm but suffers some issues with premature convergence and local optimum. Simulated annealing (SA) is a metaheuristic to approximate global optimization in a large discrete space (e.g., VRP), which can be integrated with GA, called HGA [5], to improve the performance for searching the optimal solution. However, the computation time consumed by the HGA significantly escalates in a transportation network considering stochastic travel time. The adaptive destroy-and-repair (ADAR) method is a local search-based method, employing destroy and repair operators that satisfy both diversify and intensify searching for the optimal solution, which can find the optimal solution very quickly. The advantage of hybridizing global search in form of HGA with local search in form of ADAR is the improvement in the convergence speed to the optimal or near-optimal solution [47].

4.1. Solution Evaluation. A fitness function $F(S)$ is considered to evaluate a solution S, which is formulated as the sum of the objective total cost as equation (1) and the penalty of bus capacity violation, denoted as ZT (S) and PC (S), respectively. Thus,

$$F(S) = \mathrm{ZT}(S) + \mathrm{PC}(S). \tag{23}$$

Note that if solution S meets the bus capacity constraint, PC (S) is equal to 0.

4.2. Chromosome Encoding. Chromosome is used to represent the solution, which is an integer string, consisting of a set of substrings. Each substring is a sequence of nodes

visited by a bus. For example, for a solution with 2 routes and 5 pick-up and drop-off stop pairs (PD pairs), the chromosome coding is as follows: $0\text{-}2^{+}\text{-}4^{+}\text{-}2^{-}\text{-}4^{-}\text{-}0\text{-}1^{+}\text{-}1^{-}\text{-}3^{+}\text{-}5^{+}\text{-}5^{-}\text{-}3^{-}\text{-}0$, where 1^{+}, 2^{+}, 3^{+}, 4^{+}, and 5^{+} represent pick-up stops, while 1^{-}, 2^{-}, 3^{-}, 4^{-}, and 5^{-} represent drop-off stops. 0 represents a terminal, which divides the chromosome into multiple routes. The start and end terminals are, respectively, contingent on the proximity of the nearest terminals to the first stop and the last stop. Note that the chromosome is randomly generated in this study.

4.3. Simulated Annealing (SA). SA is integrated to prevent HGA-ADAR from getting trapped into a local optimum, which conditionally accepts a slightly worse solution based on a probability, denoted as P, estimated by

$$P = \begin{cases} 1, & F(S) \geq F(S''), \\ \exp\left(-\dfrac{F(S'') - F(S)}{\gamma T}\right), & \text{otherwise}, \end{cases} \tag{24}$$

where $F(S)$ and $F(S'')$ are the fitness values of current solution S and new solution S'', respectively. Note that γT represents the current temperature, the product of temperature in the previous iteration (T) and temperature change rate (γ) (i.e., $0 < \gamma < 1$). Note that the probability of accepting a worse solution decreases as the temperature decreases at a rate of γ.

4.4. Destroy Operator. A set of three destroy operators, denoted as D, are applied in ADAR. Destroy operator I removes a number of requests (i.e., PD pairs) denoted as φ_1, which are randomly selected from the current solution S. The idea of randomly selecting requests helps to diversify the search space. Destroy operator II randomly removes one route, consisting of multi-PD pairs, from S. This idea helps to reduce the number of buses used to decrease the operator cost. Destroy operator III removes φ_2 requests resulting in greatest early and late arrival penalties. The idea is to prevent great deviations between excepted bus arrival time and actual arrival time. Note that the removed requests are placed in a destroy list, denoted as L.

The process of destroy operation is shown in Figure 2 using a simplified scenario. Two requests (i.e., the 2nd and 5th requests) are chosen and removed from the incumbent solution S. Then, the two requests are stored in the destroy list L and a temporary solution S' is produced.

4.5. Repair Operator. In this study, a set of two repair operators is employed by ADAR, which is denoted as R. Repair operator I (as shown in Figure 3) sequentially removes the requests one by one from the destroy list L (i.e., the first is the 2nd request and then the 5th request) and inserts them into the best positions of the existing routes. The best position is the one that incurs the least increase of fitness value (i.e., the positions marked in orange).

The illustration of repair operator II is shown in Figure 4. In each iteration, this operator repeatedly inserts every request in destroy list L (i.e., the 2nd and 5th requests) into the best position (i.e., the position marked in orange) of the existing routes. For each request, an increase in fitness value (calculated by equation (23)) is obtained. The request that causes the least increase in fitness value is chosen (i.e., the 5th request) and inserted into the best position of the existing routes. This process continues until the destroy list L is exhausted.

4.6. Operator Weight. Let ω_d ($d \in D$) and ω_r ($r \in R$) be weights of destroy and repair operators d and r, respectively, which are initially set equal to 1. In each iteration, operators d and r are randomly chosen according to the associated weights to produce new solution S'' based on current solution S. Note that ω_d and ω_r are justified iteratively with respect to reaction factor $\rho \in [0, 1]$ and scores of the operators denoted as ψ_d and ψ_r.

$$\begin{aligned} \omega_d &= \rho\omega_d + (1-\rho)\psi_d, \\ \omega_r &= \rho\omega_r + (1-\rho)\psi_r. \end{aligned} \tag{25}$$

ψ_d and ψ_r are determined by equations (26) and (27), which increase by σ_μ. Thus,

$$\psi_d = \psi_d + \sigma_\mu, \tag{26}$$

$$\psi_r = \psi_r + \sigma_\mu. \tag{27}$$

If $F(S'')$ is the least fitness value found in previous iterations, $\mu = 1$. If $F(S'') < F(S)$, $\mu = 2$. If $F(S'') > F(S)$ but S'' is accepted by SA, $\mu = 3$. Note that σ_μ varies within 0 and 1, and $\sigma_1 + \sigma_2 + \sigma_3 = 1$.

4.7. HGA-ADAR. Based on an initial population with GA, three operators (i.e., selection, crossover, and mutation) are applied to reproduce new solutions iteratively. SA is integrated with GA, called HGA, which permits conditionally accepting a worse solution. The HGA process terminates as the maximum number of generations is yielded, and the solution obtained is regarded as the initial solution of ADAR. A detailed procedure about the proposed solution algorithm HGA-ADAR is discussed and illustrated in Figure 5.

5. Performance Analysis

The effectiveness of HGA-ADAR is evaluated using different numbers of PD pairs in a CB network. The network scales range from 4 to 48 PD pairs. The minimized total cost yielded by the optimized solution was found by LINGO (Global Optimal Solver, version 18.0), HGA (MATLAB R2018b), and HGA-ADAR (MATLAB R2018b) on a personal laptop (Intel Core i5, 8G, 3.0 GHz).

5.1. Distribution of Bus Arrival Time. The proposed model and algorithm can virtually be applied to any distribution of bus arrival time. In this analysis, bus arrival time at stop j

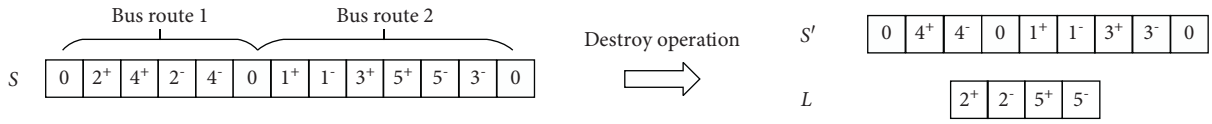


Figure 2: Process of destroy operation.

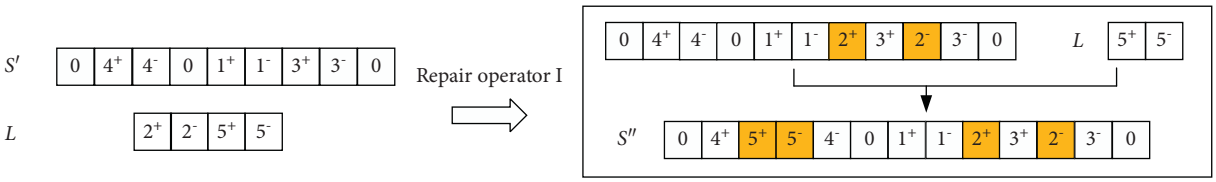


Figure 3: Illustration of repair operator I.

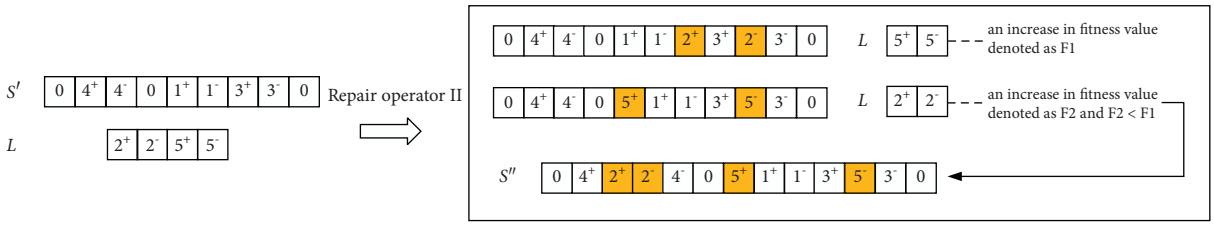


Figure 4: Illustration of repair operator II.

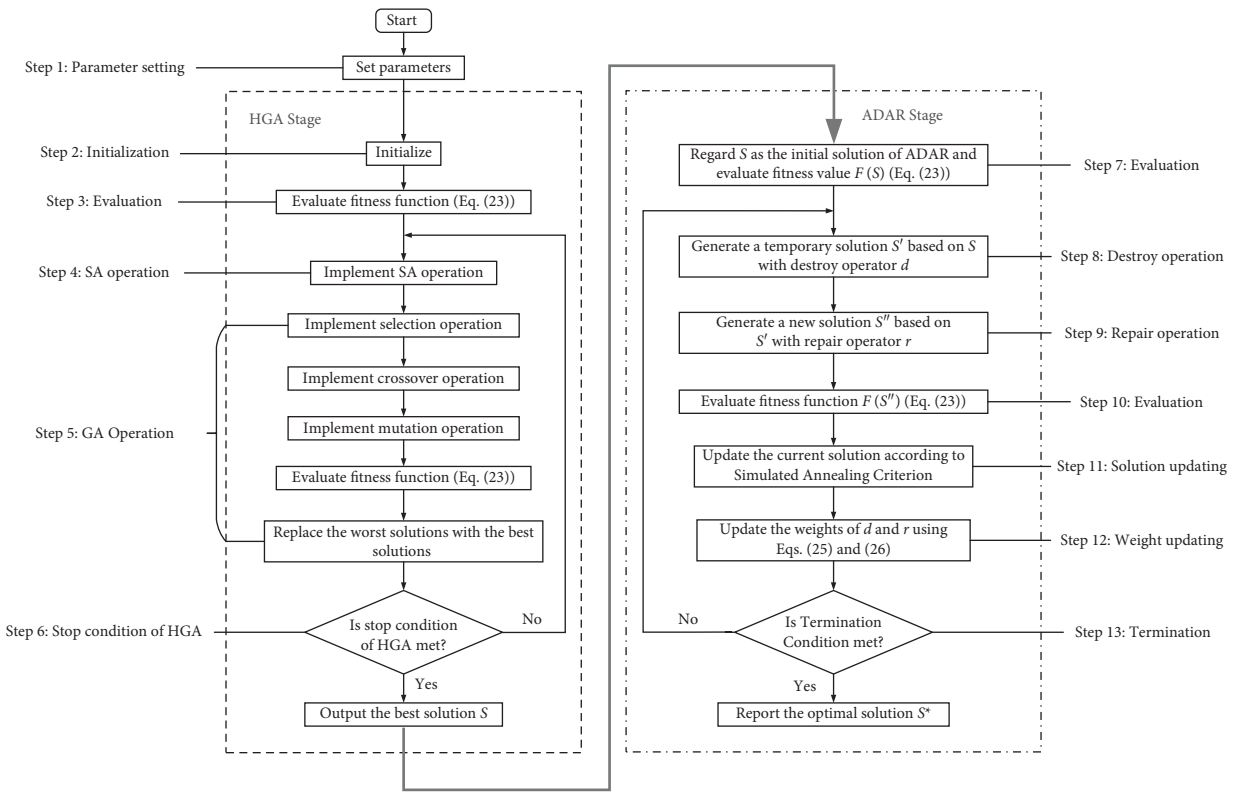


Figure 5: Proposed HGA-ADAR framework.

denoted as $\widetilde{T}_{jb}$ follows a gamma distribution with shape parameter α_{jb} and scale parameter θ_j [52–54]. The probability density function denoted as $f(\widetilde{T}_{jb})$ is given below:

$$f\left(\widetilde{T}_{jb}\right) = \frac{\left(\widetilde{T}_{jb}\right)^{\alpha_{jb}-1} e^{-\left(\widetilde{T}_{jb}/\theta_j\right)}}{\left(\theta_j\right)^{\alpha_{jb}} \Gamma\left(\alpha_{jb}\right)}, \tag{28}$$

where $\Gamma(\alpha_{jb}) = \int_0^{\infty} e^{-r} r^{\alpha_{jb}-1} dr$, $(\alpha_{jb} > 0)$. Based on the theory of gamma distribution, the mean and variance of $\widetilde{T}_{jb}$ denoted as $E(\widetilde{T}_{jb})$ and $\mathrm{Var}(\widetilde{T}_{jb})$ are formulated in equations (29) and (30), respectively.

$$E\left(\widetilde{T}_{jb}\right) = \alpha_{jb}\theta_j, \tag{29}$$

$$\mathrm{Var}\left(\widetilde{T}_{jb}\right) = \alpha_{jb}\theta_j^2. \tag{30}$$

As is well known, the sum of gamma-distributed variables that have the same scale parameter leads to a gamma-distributed variable. To make the problem tractable, we assume the same scale parameter (i.e., θ_j) for all stops. In addition, $E(\widetilde{T}_{jb})$ can be obtained from the mean of travel times on links and dwell times at stops covered by the bus until stop j. Thus, α_{jb} and $\mathrm{Var}(\widetilde{T}_{jb})$ can be determined by equations (29) and (30). Then, α_{jb} and θ_j are applied in equation (28) to calculate $f(\widetilde{T}_{jb})$.

5.2. Parameter Setting. As discussed earlier, φ_1 and φ_2 represent the number of requests to be destroyed, which are between 1 and ζH, where ζ is the destroy rate and H is the number of PD pairs. The scale parameter θ_j is set as 1. The value of in-vehicle time is 6 \$/pass-hr, while the unit penalty of early arrival and late arrival is 9 and 90 \$/pass-hr, respectively. The parameters of HGA and HGA-ADAR summarized in Table 3 were tested, which yielded acceptable solutions with the least computation time.

5.3. Solutions Obtained by LINGO and HGA-ADAR. LINGO solver is a well-known commercial software program, which is applied to optimize CB networks with deterministic bus arrival time (deterministic model). On the other hand, the proposed HGA-ADAR algorithm is applied to optimize the same CB networks considering both deterministic and stochastic bus arrival time (probabilistic model).

Ten simulation runs per experiment are conducted, and the results are listed in Table 4; the computation time is shown in Figure 6. Note that the computation time for LINGO is limited to 10,800 seconds. Considering deterministic bus arrival time, the difference between the minimized costs found by HGA-ADAR (e.g., $F(S_A)$) and LINGO (e.g., $F(S_L)$) is denoted as "ε_1," which is calculated by

$$\varepsilon_1 = \frac{F\left(S_A\right) - F\left(S_L\right)}{F\left(S_L\right)}. \tag{31}$$

The results indicate that LINGO can be applied to optimize small-scale networks (i.e., 4 and 6 PD pairs) within a given time limit. As the number of PD pairs increases, the computation time exponentially escalates and LINGO fails to find the solution. On the other hand, HGA-ADAR performs well, which efficiently finds the solutions for different network scales considering deterministic and stochastic bus arrival time. Note that the computation time consumed by optimizing stochastic cases is significantly longer.

5.4. Solutions Obtained by HGA and HGA-ADAR. To further assess the efficiency of HGA-ADAR, HGA and HGA-ADAR are both applied to optimize the network with stochastic bus arrival time (probabilistic model). The results are shown in Table 5. The difference between the minimized costs found by HGA-ADAR (e.g., $F(S_A)$) and HGA (e.g., $F(S_H)$) is denoted as "ε_2," which is calculated by equation (32). The improved computation time after applying ADAR is recorded in the "CPU^3" column, which is calculated by

$$\varepsilon_2 = \frac{F\left(S_A\right) - F\left(S_H\right)}{F\left(S_H\right)}, \tag{32}$$

$$\mathrm{CPU}^3 = \mathrm{CPU}^1 - \mathrm{CPU}^2, \tag{33}$$

where CPU^1 and CPU^2 are computation times of using HGA and HGA-ADAR, respectively. In addition, the total costs found by HGA and HGA-ADAR over iterations for 24 PD pairs are illustrated in Figure 7.

From Table 5, we can see that HGA-ADAR outperforms HGA in terms of shorter computation time as well as less minimized total cost. In addition, for the network with 24 PD pairs, the solution found by HGA shown in Figure 7 converges at the 620th generation. After applying the initial solution produced by HGA at the 100th generation, the solution found by HGA-ADAR converges at the 300th iteration. It indicates that HGA-ADAR increases convergence speed greatly when compared to HGA.

5.5. Impact of Stochastic Bus Arrival Time. To demonstrate the performance of the probabilistic model, routing and scheduling for various sizes of networks with different PD pairs are optimized considering deterministic (scenario I) and stochastic (scenario II) bus arrival time. The minimized costs summarized in Table 6 are determined based on ten simulation runs. The results show that minimized total cost reduced by 4.5–13.3% for networks with various PD pairs if stochastic bus arrival time is considered, in which early and late arrival penalties reduced by 21.5–63.7%, user cost decreased by 0–10.2%, and operator cost increased by 0–17% because of increased fleet size.

6. Case Study

A real-world CB network in Xi'an, China, is employed as a case study to evaluate the effectiveness and applicability of the proposed model. The study network includes 31 stops and 7 terminals as illustrated in Figure 8. An instance with 275 passenger requests is presented, which is generated based on the real-world passenger travel data in Xi'an. The service time horizon of CB service ranges between 6:00 am

Table 3: HGA and HGA-ADAR parameters.

Parameter	HGA-ADAR	HGA
Population size	100	100
Crossover probability	0.9	0.9
Mutation probability	0.1	0.1
Maximum HGA generations	50–200	300–1000
Maximum ADAR iterations	500–2000	—
Destroy rate ζ	0.15	—
σ_1	0.7	—
σ_2	0.2	—
σ_3	0.1	—
Reaction factor ρ	0.2	—
Initial annealing temperature	1000	1000
Temperature change rate γ	0.95	0.95

Table 4: LINGO and HGA-ADAR results (10 simulation runs).

No. of PD pairs	Deterministic model					Probabilistic model		
	LINGO	HGA-ADAR				HGA-ADAR		
	Minimized total cost ($)	Minimized total cost ($)	Average total cost ($)	Standard deviation ($)	ε_1 (%)	Minimized total cost ($)	Average total cost ($)	Standard deviation ($)
4	199.5	199.5	199.5	0	0	215.2	215.2	0
6	259.7	259.7	259.7	0	0	289.6	290.1	0.4
8	417.4	408.3	410.7	0.9	NA	445.8	448.7	1.8
10	NA	636.6	640	1.8	NA	680	684.1	2.5
16	NA	982.5	987.9	3.3	NA	1,030.5	1,036.3	3.7
24	NA	1,265.5	1,270.8	3.9	NA	1,356.5	1,363.2	4.4
32	NA	1,542.4	1,549.2	4.7	NA	1,639.6	1,647.8	5.6
40	NA	2,803.6	2,815.3	6.2	NA	3,002.3	3,016.5	7.8
48	NA	3,553.1	3,567.1	6.9	NA	3,813.5	3,829.8	8.1

Note. 417.4 $—a feasible solution found at 10,800 seconds; NA—no feasible solution.

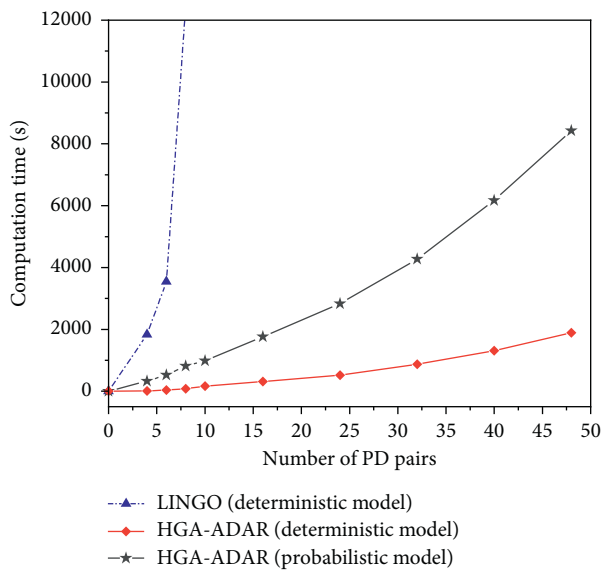


Figure 6: Computation time with LINGO and HGA-ADAR.

and 10:00 am. Based on the method described in the paper of Sun et al. [5], the passengers are partitioned into 35 groups which are shown in Table 7.

The fixed and variable costs as well as capacities of various bus types are shown in Table 8, which are provided by the transit agency, Xi'an Public Transport Corporation. Fixed cost per bus-hour f_b is determined by the expense on vehicle depreciation which is purchase cost divided by useful life. Average variable cost per bus-hour v_b is determined by expenses on fuel/energy, labour, maintenance, management, and social costs. The parameters of HGA and HGA-ADAR are summarized in Table 3.

6.1. Result Analysis. The results discussed in this section yield the minimum total costs in ten simulation runs. The optimized results under scenario I (deterministic bus arrival time) and scenario II (stochastic bus arrival time) are summarized in Table 9, while the associated CB routing and scheduling under scenario II are listed in Table 10.

In Table 9, the minimized total cost in scenario I is 3,622.7 $, consisting of user cost of 1,090.2 $, operator cost of 1,881.2 $, and early and late arrival penalties of 651.3 $. The average passenger travel time is 39.6 min, while the average travel time per bus is 92.3 min. 12 buses with different capacities are employed, in which the numbers of 10-, 15-, 20-, and 25-seat buses are 1, 1, 4, and 6, respectively. Under scenario II considering stochastic bus arrival time, the minimized total cost can be reduced by 8.4% (from 3,622.7 $ to 3,317.4 $). The user cost reduced by 10.8% (from 1,090.2 $ to 972.6 $), while early and late arrival penalties significantly reduced by 43.7% (from 651.3 $ to 366.9 $). The operator

Table 5: HGA and HGA-ADAR results (10 simulation runs; probabilistic model).

No. of PD pairs	HGA				HGA-ADAR					
	Minimized total cost ($)	Average total cost ($)	Standard deviation ($)	CPU[1] (s)	Minimized total cost ($)	Average total cost ($)	Standard deviation ($)	CPU[2] (s)	ε_2 (%)	CPU[3] (s)
4	215.2	215.2	0	617	215.2	215.2	0	331	0	286
6	289.6	291.2	0.7	925	289.6	290.1	0.4	523	0	402
8	446.3	450.2	2.9	1234	445.8	448.7	1.8	815	−0.1	419
10	696.1	700.4	3.3	2571	680	684.1	2.5	986	−2.3	1585
16	1,086.1	1,092.3	6.2	6582	1,030.5	1,036.3	3.7	1765	−5.1	4817
24	1,442.3	1,449.7	6.8	8963	1,356.5	1,363.2	4.4	2832	−5.9	6131
32	1,740.2	1,749.9	8.1	13985	1,639.6	1,647.8	5.6	4274	−5.7	9711
40	3,146.7	3,155.3	8.4	18172	3,002.3	3,016.5	7.8	6171	−4.6	12001
48	4,035.9	4,046.8	10.9	24745	3,813.5	3,829.8	8.1	8428	−5.5	16317

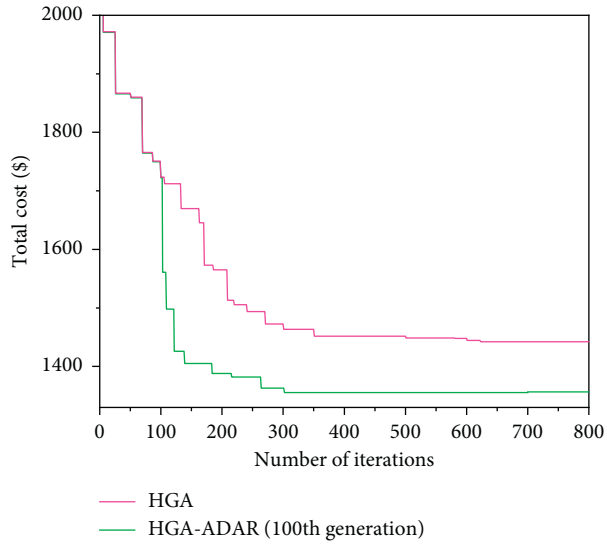


Figure 7: Total cost yielded by HGA and HGA-ADAR (24 PD pairs).

Table 6: Optimal results under various scenarios and number of PD pairs (HGA-ADAR).

No. of PD pairs	Scenario	ZT ($)	ZU ($)	ZO ($)	ZP ($)	Fleet size (veh)
4	I	225.4	89.8	105.3	30.3	1
	II	215.2	89.5	101.9	23.8	1
6	I	314.6	114.2	137.5	62.9	2
	II	289.6	116.4	137.7	35.5	2
8	I	503.8	190.9	185.8	127.1	2
	II	445.8	177.8	217.4	50.6	3
10	I	779.2	293.2	293.4	192.6	3
	II	680	266.2	337.9	75.9	5
16	I	1,188.5	480.2	460.4	247.9	5
	II	1,030.5	431.1	509.5	89.9	7
24	I	1,478.2	602.1	600.6	275.5	6
	II	1,356.5	551	693.3	112.2	9
32	I	1,793.4	720.8	765.2	307.4	7
	II	1,639.6	662.2	794.2	183.2	10
40	I	3,186.9	1,195.5	1,290.2	701.2	11
	II	3,002.3	1,101.3	1,432.4	468.6	15
48	I	4,050	1,388.5	1,827.7	833.8	17
	II	3,813.5	1,250	1,955.3	608.2	21

Note. ZT—total cost; ZU—user cost; ZO—operator cost; ZP—early and late arrival penalties.

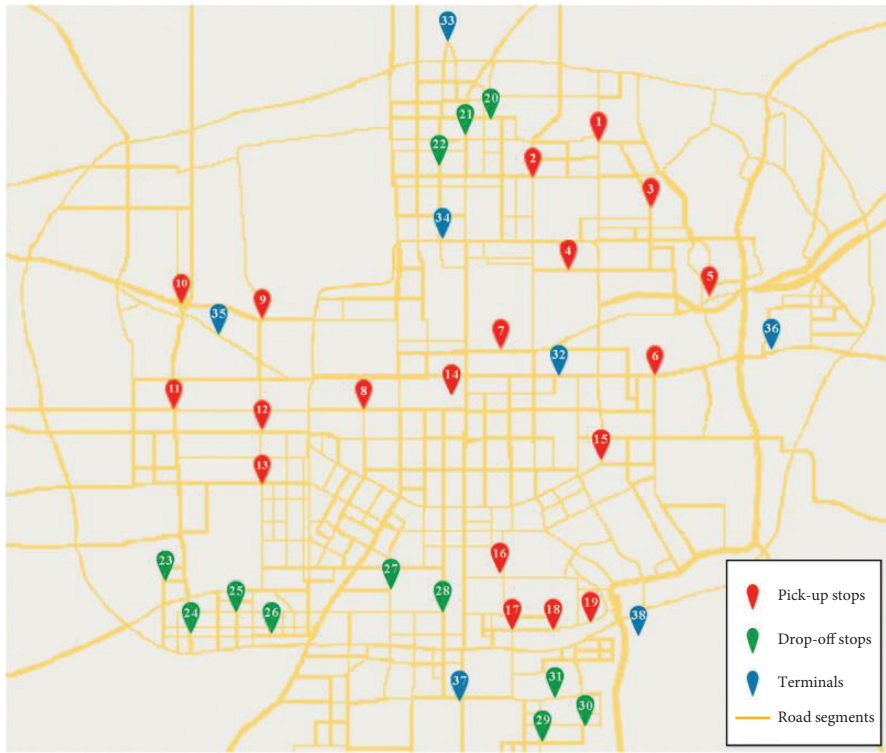


Figure 8: Stop and terminal locations of CB network in Xi'an.

Table 7: Spatiotemporal passenger PD demand.

No.	PD pair	Passengers (pass)	Expected arrival time window
1	[1, 22]	9	[7:55, 8:05]
2	[2, 21]	9	[8:10, 8:20]
3	[3, 22]	7	[8:25, 8:35]
4	[4, 20]	6	[8:25, 8:35]
5	[5, 21]	9	[8:10, 8:20]
6	[5, 22]	7	[8:25, 8:35]
7	[5, 26]	7	[8:40, 8:50]
8	[6, 20]	13	[8:10, 8:20]
9	[6, 26]	6	[8:40, 8:50]
10	[7, 28]	6	[8:10, 8:20]
11	[7, 21]	8	[8:40, 8:50]
12	[7, 30]	8	[7:25, 7:35]
13	[7, 30]	8	[7:55, 8:05]
14	[8, 28]	6	[8:10, 8:20]
15	[8, 29]	6	[8:25, 8:35]
16	[9, 22]	8	[8:25, 8:35]
17	[9, 27]	8	[7:55, 8:05]
18	[9, 29]	7	[8:55, 9:05]
19	[10, 22]	9	[8:25, 8:35]
20	[10, 29]	14	[8:40, 8:50]
21	[11, 31]	13	[8:25, 8:35]
22	[12, 30]	5	[8:40, 8:50]
23	[13, 30]	6	[8:40, 8:50]
24	[14, 22]	9	[8:40, 8:50]
25	[14, 27]	9	[7:55, 8:05]
26	[14, 30]	8	[8:10, 8:20]
27	[15, 23]	7	[8:10, 8:20]
28	[15, 29]	6	[8:10, 8:20]

Table 7: Continued.

No.	PD pair	Passengers (pass)	Expected arrival time window
29	[16, 22]	5	[8:25, 8:35]
30	[16, 24]	9	[8:10, 8:20]
31	[16, 30]	9	[8:10, 8:20]
32	[17, 22]	7	[8:25, 8:35]
33	[17, 25]	7	[8:10, 8:20]
34	[18, 26]	8	[8:40, 8:50]
35	[19, 23]	6	[8:55, 9:05]

Note. Stops 1 through 19 represent pick-up stops. Stops 20 through 31 represent drop-off stops.

Table 8: Bus type and associated capacity and cost.

Bus type		*A*	*B*	*C*	*D*
C_b (seat/veh)		10	15	20	25
f_b (\$/veh-hr)	Vehicle depreciation cost	5	6	7	8
v_b (\$/veh-hr)	Fuel/energy/labour/maintenance/management/social cost	54	72	90	108

Note. C_b—bus capacity; f_b—fixed cost; v_b—variable cost.

Table 9: Optimized results.

	Scenario I	Scenario II
Minimized total cost (\$)	3,622.7	3,317.4
User cost (\$)	1,090.2	972.6
Operator cost (\$)	1,881.2	1,977.9
Early and late arrival penalties (\$)	651.3	366.9
Avg. passenger travel time (min/pass)	39.6	35.4
Avg. bus travel time (min/veh)	92.3	77.6
Fleet size (veh)		
10-seat	1	2
15-seat	1	3
20-seat	4	4
25-seat	6	5
Total (fleet size)	12	14

Table 10: Optimal routing/scheduling, vehicle assignment, and early and late arrival penalties (scenario II).

Route ID	Bus capacity	Bus route and schedule	Number of passengers (pass)	Avg. passenger travel time (min)	Early arrival penalty (\$)	Late arrival penalty (\$)
1	10	36 ⟶ 5 ⟶ 21 ⟶ 34 7:41–7:47–8:10–8:16	9	23	2.97	1.71
2	10	32 ⟶ 14 ⟶ 27 ⟶ 14 ⟶ 22 ⟶ 34 7:15–7:23–7:44–8:05–8:40–8:46	18	28	20.1	9.97
3	15	32 ⟶ 7 ⟶ 30 ⟶ 17 ⟶ 16 ⟶ 22 ⟶ 34 6:21–6:32–7:16–7:32–7:43–8:25–8:31	20	46.7	26.13	13.58
4	15	36 ⟶ 5 ⟶ 3 ⟶ 22 ⟶ 34 7:48–7:55–8:02–8:25–8:32	14	26.5	4.99	3.62
5	15	32 ⟶ 7 ⟶ 30 ⟶ 28 ⟶ 37 6:54–7:05–7:49–8:10–8:21	14	53	11.88	6.41
6	20	34 ⟶ 2 ⟶ 1 ⟶ 21 ⟶ 22 ⟶ 7 ⟶ 21 ⟶ 34 7:25–7:35–7:42–7:52–7:59–8:15–8:40–8:46	26	19.5	33.61	9.34
7	20	35 ⟶ 9 ⟶ 10 ⟶ 22 ⟶ 34 7:36–7:44–7:52–8:25–8:33	17	36.7	8.5	5.98
8	20	36 ⟶ 6 ⟶ 4 ⟶ 20 ⟶ 33 7:20–7:30–7:48–8:10–8:17	19	34.3	19.05	6.24
9	20	35 ⟶ 9 ⟶ 8 ⟶ 27 ⟶ 28 ⟶ 29 ⟶ 38 7:07–7:15–7:32–7:52–8:04–8:25–8:32	20	40.3	14.89	8.03

Table 10: Continued.

Route ID	Bus capacity	Bus route and schedule	Number of passengers (pass)	Avg. passenger travel time (min)	Early arrival penalty ($)	Late arrival penalty ($)
10	25	35 ⟶ 9 ⟶ 10 ⟶ 29 ⟶ 38 7:50–7:59–8:06–8:41–8:48	21	37.3	19.75	9.01
11	25	36 ⟶ 5 ⟶ 6 ⟶ 18 ⟶ 26 ⟶ 37 7:36–7:45–7:55–8:20–8:40–8:54	21	38.8	13.31	8.39
12	25	35 ⟶ 13 ⟶ 12 ⟶ 11 ⟶ 31 ⟶ 30 ⟶ 38 7:25–7:38–7:48–7:55–8:25–8:30–8:39	24	38	23.19	9.08
13	25	32 ⟶ 15 ⟶ 16 ⟶ 17 ⟶ 25 ⟶ 24 ⟶ 23 ⟶ 35 7:07–7:17–7:32–7:53–7:59–8:05–8:12–8:27	23	31.5	22.69	11.57
14	25	32 ⟶ 14 ⟶ 15 ⟶ 16 ⟶ 30 ⟶ 29 ⟶ 19 ⟶ 23 ⟶ 35 7:00–7:09–7:27–7:40–8:03–8:10–8:25–8:55–9:09	29	37.1	27.98	14.93

Note. Stops 1 through 19 are pick-up stops. Stops 20 through 31 are drop-off stops. Stops 32 through 38 are terminals.

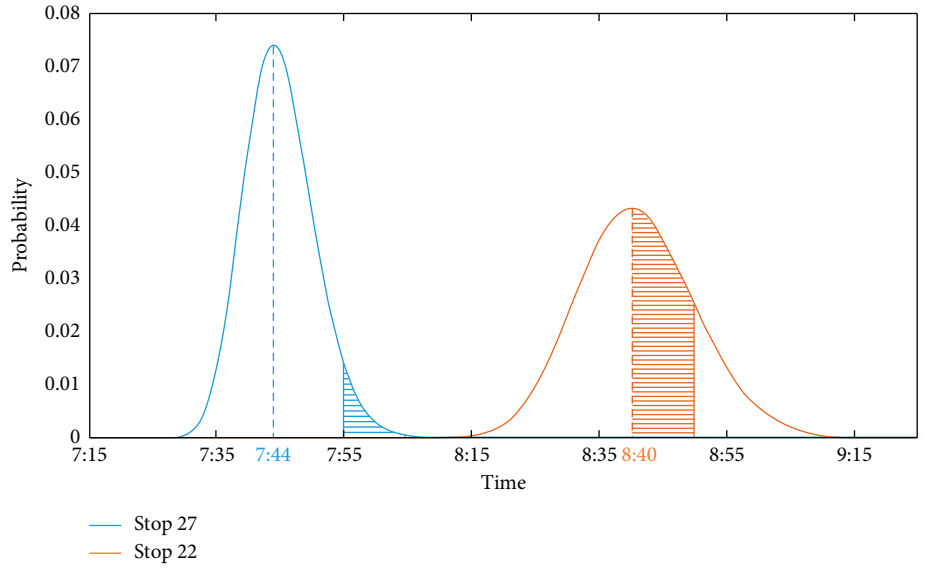


Figure 9: Probability distributions of bus arrival time (route 2; time window = 10 min; $\theta_j = 1$).

Table 11: Optimal results under different distributions of bus arrival time.

	Gamma	Normal	Log-normal
Minimized total cost ($)	3,317.4	3,282.7	3,320.6
User cost ($)	972.6	948.3	927.8
Operator cost ($)	1,977.9	1,946.7	1,940.2
Early and late arrival penalties ($)	366.9	387.7	452.6
Fleet size (veh)			
10-seat	2	3	3
15-seat	3	3	4
20-seat	4	4	3
25-seat	5	4	4
Total (fleet size)	14	14	14

cost slightly increased mainly due to the increase of fleet size (from 12 to 14 vehicles). The results indicate that considering stochastic bus arrival time shall yield a better solution.

Table 10 illustrates the optimal CB routing and scheduling, vehicle assignment, average passenger travel time, and early and late arrival penalties. The results suggest 14 routes, in which route 1 is the shortest one consisting of 2 stops (excluding terminals), route 14 is the longest one consisting of 7 stops, and the average number of stops per route is 4. The suggested CB service begins at 6:20 am and ends at 9:10

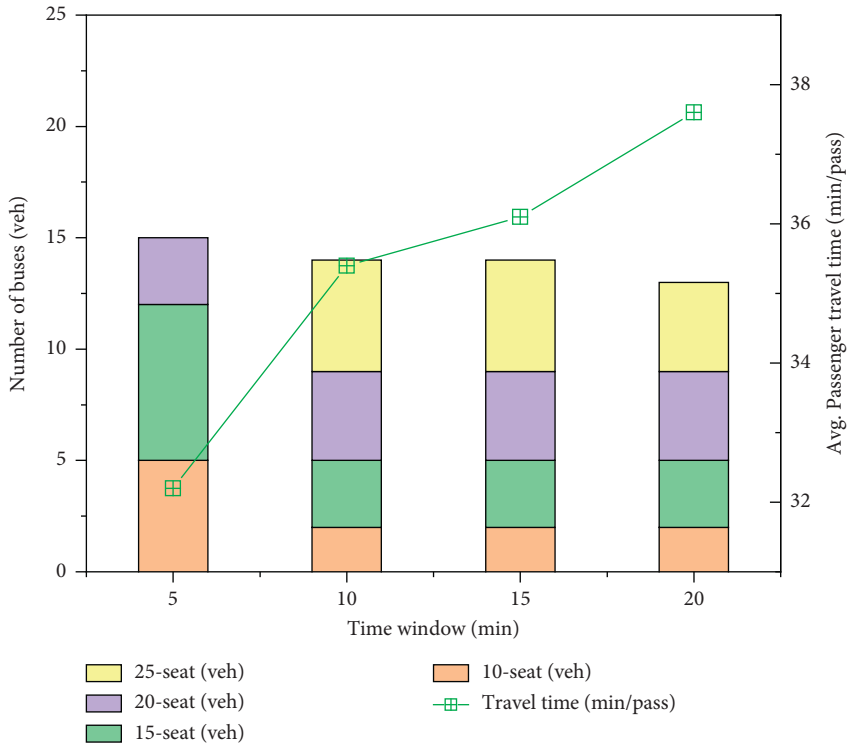


FIGURE 10: Fleet size and passenger travel time vs. time window ($\theta_j = 1$).

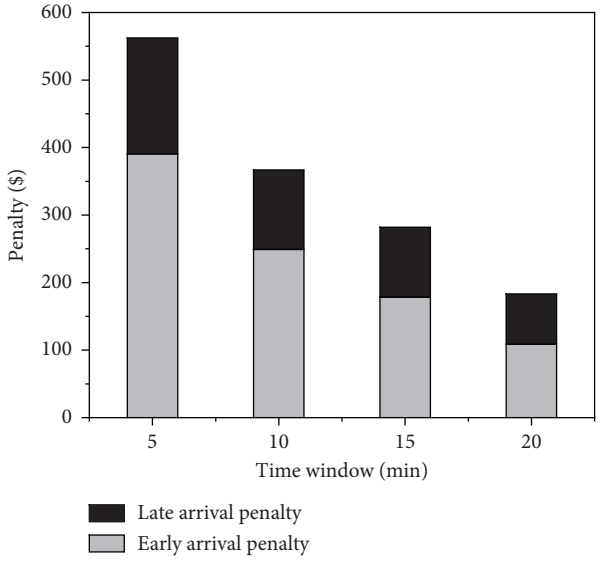


FIGURE 11: Early and late arrival penalties vs. time window ($\theta_j = 1$).

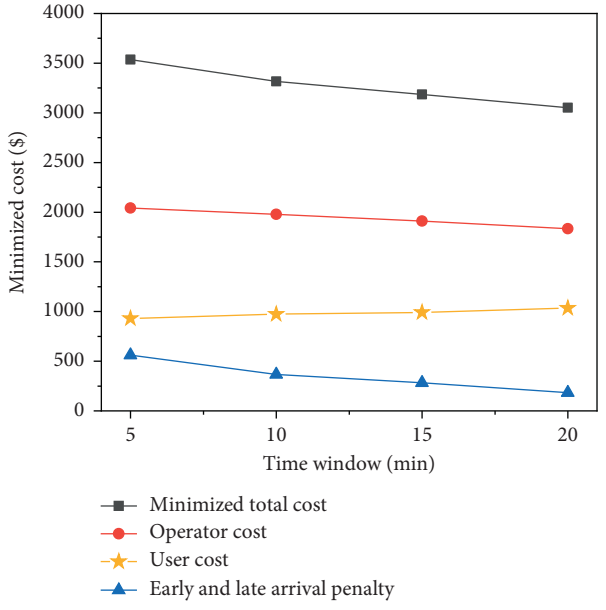


FIGURE 12: Minimized cost vs. time window ($\theta_j = 1$).

am. It is found that the early arrival penalty is always greater than the late arrival penalty for all bus routes, indicating that the probability of passengers arrive destinations early is much greater than late arrivals because of the huge late arrival penalty. The probability distributions of bus arrival time at drop-off stops for route 2 are illustrated in Figure 9. Route 2 is 32-14-27-14-22-34. The bus starts from terminal 32, which takes 9 passengers at stop 14, drops them off at stop 27, then returns back to stop 14, takes 9 passengers, drops them off at stop 22, and then parks at terminal 34.

Two curves in Figure 9 represent the probability distributions of bus arrival time at stops 27 and 22. The scheduled bus arrival times are 7:44 am and 8:40 am with variance of 5.5 min and 9.2 min, respectively. Note that the shaded areas represent the probability of passengers will

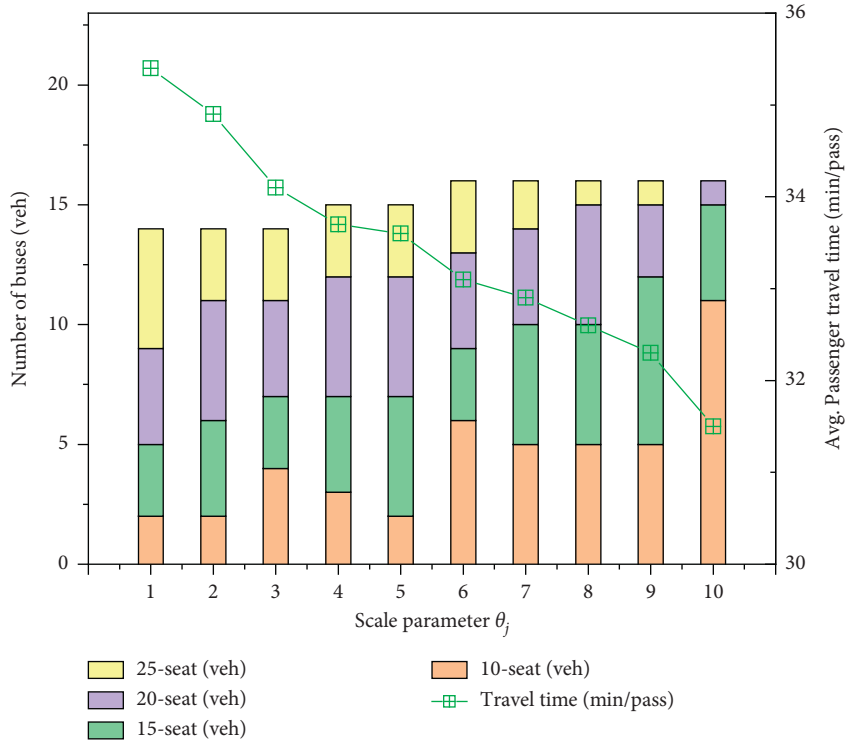


FIGURE 13: Fleet size and passenger travel time vs. scale parameter θ_j (time window = 10 min).

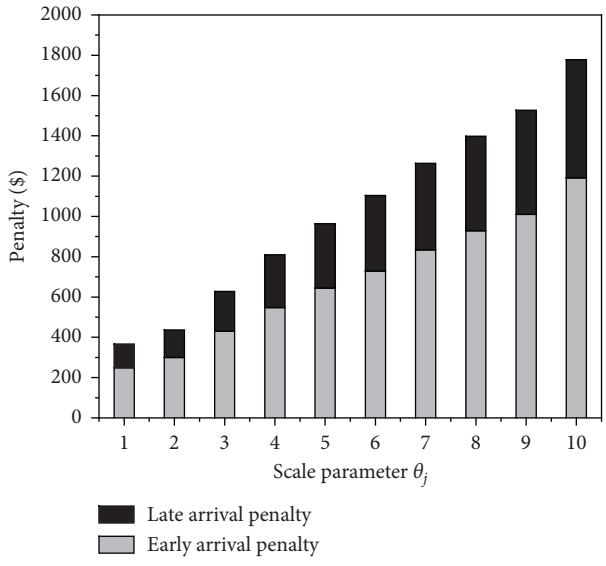


FIGURE 14: Early and late arrival penalties vs. scale parameter θ_j (time window = 10 min).

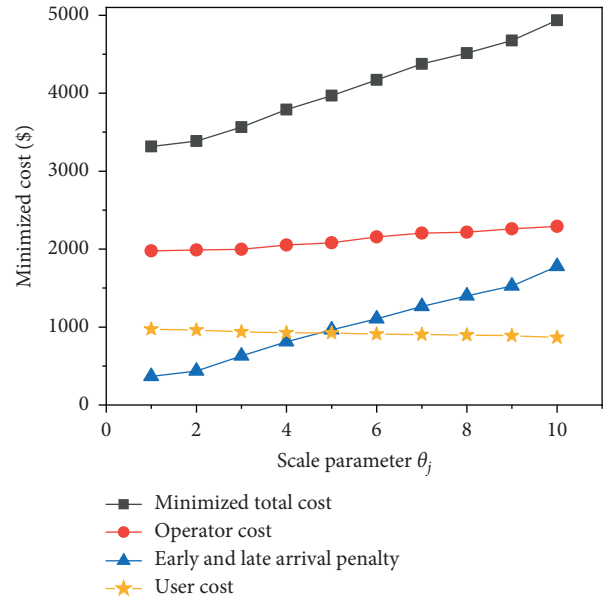


FIGURE 15: Minimized cost vs. scale parameter θ_j (time window = 10 min).

arrive at the stops within the desired time windows. The left and right parts of the shaded area represent the probabilities of early and late bus arrivals, respectively. The probabilities of late arrivals are 0.0005 for stop 27 and 0.1408 for stop 22, and those for early arrivals are 0.9706 and 0.5144, respectively. The early (and late) penalty is the product of early (and late) arrival probability and the number of passengers alighting from the drop-off stop multiplied by the unit penalty, which is 15.09 $ (and 0.01 $) at stop 27 and 4.98 $ (and 9.96 $) at stop 22.

6.2. Impact of Distribution of Bus Arrival Time. As discussed above, the most commonly used distributions of bus arrival time are normal, log-normal, and gamma distributions. To

verify the application and rationality of the proposed model and algorithm, the three distributions of bus arrival time are applied in the case study. Ten simulation runs per experiment are conducted, and the optimal results under the different distributions are shown in Table 11. To make the problem comparable and tractable, we assume that the mean and variance of bus arrival time with different distributions are all generated by the same standard. Thus, the parameters (δ and ξ) of the log-normal distribution are calculated as

$$\begin{aligned} e^{\delta+(\xi^2/2)} &= E\left(\widetilde{T}_{jb}\right), \\ \left(e^{\xi^2}-1\right)e^{2\delta+\xi^2} &= \mathrm{Var}\left(\widetilde{T}_{jb}\right). \end{aligned} \tag{34}$$

The probability density function denoted as $f_{\mathrm{log}n}(\widetilde{T}_{jb})$ is given below:

$$f_{\mathrm{log}n}\left(\widetilde{T}_{jb}\right) = \frac{1}{\sqrt{2\pi}\widetilde{T}_{jb}\xi}e^{-\left(\left(\ln\widetilde{T}_{jb}-\delta\right)^2/2\xi^2\right)}. \tag{35}$$

The parameters (ϕ and κ) of the normal distribution are calculated as

$$\begin{aligned} \phi &= E\left(\widetilde{T}_{jb}\right), \\ \kappa^2 &= \mathrm{Var}\left(\widetilde{T}_{jb}\right). \end{aligned} \tag{36}$$

The probability density function denoted as $f_{\mathrm{norm}}(\widetilde{T}_{jb})$ is given below:

$$f_{\mathrm{norm}}\left(\widetilde{T}_{jb}\right) = \frac{1}{\sqrt{2\pi}\kappa}e^{-\left(\left(\widetilde{T}_{jb}-\phi\right)^2/2\kappa^2\right)}. \tag{37}$$

From Table 11, we can see that the proposed model and algorithm can be applied to the three distributions of bus arrival time, and the difference under different distributions only lies in the optimal results. It is worth noting that the optimal results under the three distributions of bus arrival time are close, indicating that it is reasonable to use gamma distribution in this paper.

6.3. Sensitivity Analysis. A sensitivity analysis is performed to explore the relationship between model parameters (i.e., time window tw and scale parameter θ_j) and optimal results (i.e., fleet size, travel time, and minimized cost).

Figure 10 suggests that as the time window increases from 5 to 20 minutes, smaller buses (10- and 15-seat) decrease from 12 to 5, while larger buses (20- and 25-seat) increase from 3 to 8. It indicates that a loose time window could make it more flexible for the operator to arrange the CB service, and larger buses are preferred to reduce operator cost. However, the decreased fleet size would increase the number of stops per route as well as passenger travel time.

Figure 11 suggests that an increased time window may reduce early and late arrival penalties and increase the probability of on-time passenger arrival. It is worth noting that the early arrival penalty is always greater than the late arrival penalty, indicating that there are more passengers who arrive earlier than those who arrive later from the desired arrival time, due to a greater unit penalty for late arrival.

Figure 12 shows that as the time window increases, minimized total cost consisting of user cost, operator cost, and early and late arrival penalties decreases. A decrease in operator cost due to the decreased fleet size would increase average passenger in-vehicle time and user cost.

Figure 13 indicates that as θ_j increases (i.e., a greater variance of bus arrival time), fleet size increases, and smaller buses are preferable. It is worth noting that the number of buses is equal to the number of routes. Therefore, the increased fleet size would increase the number of routes and decrease the number of stops per route. This would result in fewer passengers per route with less travel time, and smaller buses with lower operation cost are preferable.

As θ_j increases as shown in Figure 14, the early and late arrival penalties increase because of the increasing variance of bus arrivals at stops and the increasing probability of early and late arrival.

Figure 15 suggests that as θ_j increases, minimized total cost increases. Operator cost is expected to increase because of increased fleet size. A minor decrease in user cost results from reduced passenger travel time. The increased probability of early and late bus arrivals could increase the early and late arrival penalties.

7. Conclusions

In this paper, we propose a probabilistic model for optimizing CB service considering stochastic bus arrival time and some practical conditions, including time window and capacity in mixed bus fleet. An HGA-ADAR algorithm is applied to efficiently search for the solution. The developed model and algorithm can deal with the impact of probabilistic bus arrival time on the optimization of CB service planning. The objective total cost, consisting of user cost, operator cost, and early and late arrival penalties, is minimized subject to several practical constraints. The decision variables include vehicle assignment, bus routing, timetabling, and fleet size.

According to the performance analysis, the developed model and HGA-ADAR have demonstrated themselves effective in minimizing the total cost with the least computation time. The reduced total cost is mainly from saving early and late penalties by elevating on-time passenger arrivals by routing and scheduling a mixed bus fleet. The solution suggested by the proposed model with stochastic bus arrival time is superior to that of the deterministic model. The minimized total cost can be reduced by 13.3%, and early and late penalties can be reduced by 63.7% in networks with various PD pairs.

The proposed model is applied to optimize the service for a CB network in Xi'an, China. Three distributions of bus arrival time (i.e., gamma, normal, and log-normal distributions) are used to verify the application and rationality of the proposed model and algorithm. The results of sensitivity analysis suggest that as the time window increases, minimized total cost decreases because of reduced fleet size and operator cost. At this moment, larger buses shall be employed to meet the demand, albeit user cost slightly increases due to the increase in

passenger travel time. Results also suggest that as scale parameter θ_j increases, minimized total cost increases because of increased fleet size. Smaller buses are advisable for operator. In addition, early and late arrival penalties tend to increase as the variance of bus arrival time increases, and they decrease as time window increases.

The immediate extension of this study will focus on enhancing the model to deal with time-dependent traffic conditions which would permit the probabilistic model to be applied for cost-effective CB operation (e.g., dynamic bus dispatching and holding control). In the future, it is desirable to model the problem as a multi-objective optimization one and compute the Pareto front. The behaviour of HGA-ADAR shall be further explored, especially on the sensitivity of the initial solution and parameters (i.e., numbers and weights of destroy and repair operators, destroy rate, etc.), to expedite the converging speed and ensure the quality of the solution. In addition, the computation for solving the stochastic problem is time consuming, and HGA-ADAR still has much room for improvement in the computing speed of solving such stochastic problems.

Data Availability

The data used in the case study to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was in part sponsored by the Young Scientists Fund of the Natural Science Foundation of Shaanxi Province, China (grant no. 2020JQ-399), the Special Scientific Research Project of Shaanxi Provincial Department of Education (grant no. 20JK0356), and the Fundamental Research Funds for the Central Universities (300102220101).

References

[1] X. Lu, J. Yu, X. Yang, S. Pan, and N. Zou, "Flexible feeder transit route design to enhance service accessibility in urban area," *Journal of Advanced Transportation*, vol. 50, no. 4, pp. 507–521, 2016.

[2] P. Franco, R. Johnston, and E. McCormick, "Demand responsive Transport: generation of activity patterns from mobile phone network data to support the operation of new mobility services," *Transportation Research Part A: Policy and Practice*, vol. 131, pp. 244–266, 2020.

[3] T. Y. Ma, J. Y. J. Chow, S. Klein, and Z. Ma, "A user-operator assignment game with heterogeneous user groups for empirical evaluation of a microtransit service in Luxembourg," *Transportmetrica: Transportation Science*, vol. 17, pp. 1–28, 2020.

[4] T. Liu and A. Ceder, "Analysis of a new public-transport-service concept: customized bus in China," *Transport Policy*, vol. 39, pp. 63–76, 2015.

[5] Q. Sun, S. Chien, D. Hu, G. Chen, and R. S. Jiang, "Optimizing multi-terminal customized bus service with mixed fleet," *IEEE Access*, vol. 8, Article ID 156456, 2020.

[6] S. I. J. Chien, "Optimization of headway, vehicle size and route choice for minimum cost feeder service," *Transportation Planning and Technology*, vol. 28, no. 5, pp. 359–380, 2005.

[7] Y. Y. Ulusoy, S. I. Chien, and C. H. Wei, "Optimal all-stop, short-turn, and express transit services under heterogeneous demand," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2197, no. 1, pp. 8–18, 2010.

[8] Y. Y. Ulusoy and S. I. Chien, "Optimal bus service patterns and frequencies considering transfer demand elasticity with genetic algorithm," *Transportation Planning and Technology*, vol. 38, no. 4, pp. 409–424, 2015.

[9] H. Z Qu, S. I. J. Chien, X. B Liu, P. T Zhang, and A. Bladikas, "Optimizing bus services with variable directional and temporal demand using genetic algorithm," *Journal of Central South University*, vol. 23, no. 7, pp. 1786–1798, 2016.

[10] M. S. Chowdhury and S. I. Chien, "Optimizing fare and headway to facilitate timed transfer considering demand elasticity," *Transportation Planning and Technology*, vol. 42, no. 1, pp. 56–69, 2018.

[11] G. Chen, D. Hu, and S. Chien, "Optimizing battery-electric-feeder service and wireless charging locations with nested genetic algorithm," *IEEE Access*, vol. 8, Article ID 67166, 2020.

[12] L. Tong, L. Zhou, J. Liu, and X. Zhou, "Customized bus service design for jointly optimizing passenger-to-vehicle assignment and vehicle routing," *Transportation Research Part C: Emerging Technologies*, vol. 85, pp. 451–475, 2017.

[13] R. Guo, W. Guan, and W. Zhang, "Route design problem of customized buses: mixed integer programming model and case study," *Journal of Transportation Engineering, Part A: Systems*, vol. 144, 2018.

[14] R. Guo, W. Guan, W. Zhang, F. Meng, and Z. Zhang, "Customized bus routing problem with time window restrictions: model and case study," *Transportmetrica: Transportation Science*, vol. 15, no. 2, pp. 1804–1824, 2019.

[15] R. Guo, W. Zhang, W. Guan, and B. Ran, "Time-dependent urban customized bus routing with path flexibility," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, pp. 1–10, 2020.

[16] Y. Lyu, C. Y. Chow, V. C. S. Lee, J. K. Y. Ng, Y. Li, and J. Zeng, "CB-Planner: a bus line planning framework for customized bus systems," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 233–253, 2019.

[17] D. Huang, Y. Gu, S. Wang, Z. Liu, and W. Zhang, "A two-phase optimization model for the demand-responsive customized bus network design," *Transportation Research Part C: Emerging Technologies*, vol. 111, pp. 1–21, 2020.

[18] A. S. Kenyon and D. P. Morton, "Stochastic vehicle routing with random travel times," *Transportation Science*, vol. 37, no. 1, pp. 69–82, 2003.

[19] X. Li, P. Tian, and S. C. H. Leung, "Vehicle routing problems with time windows and stochastic travel and service times: models and algorithm," *International Journal of Production Economics*, vol. 125, no. 1, pp. 137–145, 2010.

[20] M. Gendreau, G. Laporte, and R. Séguin, "Stochastic vehicle routing," *European Journal of Operational Research*, vol. 88, no. 1, pp. 3–12, 1996.

[21] M. Gendreau, O. Jabali, and W. Rei, "50th anniversary invited article-future research directions in stochastic vehicle routing," *Transportation Science*, vol. 50, no. 4, pp. 1163–1173, 2016.

[22] Y. Adulyasak and P. Jaillet, "Models and algorithms for stochastic and robust vehicle routing with deadlines," *Transportation Science*, vol. 50, no. 2, pp. 608–626, 2016.

[23] F. Errico, G. Desaulniers, M. Gendreau, W. Rei, and L. Rousseau, "A priori optimization with recourse for the vehicle routing problem with hard time windows and stochastic service times," *European Journal of Operational Research*, vol. 249, no. 1, pp. 55–66, 2016.

[24] J. F. Cordeau, "A branch-and-cut algorithm for the dial-a-ride problem," *Operations Research*, vol. 54, no. 3, pp. 573–586, 2006.

[25] L. Zhao and S. I. Chien, "Investigating the impact of stochastic vehicle arrivals to optimal stop spacing and headway for a feeder bus route," *Journal of Advanced Transportation*, vol. 49, no. 3, pp. 341–357, 2015.

[26] L. Zhao, S. I. Chien, L. N. Spasovic, and X. Liu, "Modeling and optimizing urban bus transit considering headway variation for cost and service reliability analysis," *Transportation Planning and Technology*, vol. 41, no. 7, pp. 706–723, 2018.

[27] B. Y. Chen, Y. Wang, D. Wang, and W. H. K. Lam, "Understanding travel time uncertainty impacts on the equity of individual accessibility," *Transportation Research Part D: Transport and Environment*, vol. 75, pp. 156–169, 2019.

[28] L. Fu, "Scheduling dial-a-ride paratransit under time-varying, stochastic congestion," *Transportation Research Part B: Methodological*, vol. 36, no. 6, pp. 485–506, 2002.

[29] T. S. Chang, Y. W. Wan, and W. T. Ooi, "A stochastic dynamic traveling salesman problem with hard time windows," *European Journal of Operational Research*, vol. 198, no. 3, pp. 748–759, 2009.

[30] Y. Shi, T. Boudouh, O. Grunder, and D. Wang, "Modeling and solving simultaneous delivery and pick-up problem with stochastic travel and service times in home health care," *Expert Systems with Applications*, vol. 102, pp. 218–233, 2018.

[31] S. Chien, S. K. Daripally, and K. Kim, "Development of a probabilistic model to optimize disseminated real-time bus arrival information for pre-trip passengers," *Journal of Advanced Transportation*, vol. 41, pp. 195–215, 2010.

[32] A. Prakash, "Algorithms for most reliable routes on stochastic and time-dependent networks," *Transportation Research Part B: Methodological*, vol. 138, pp. 202–220, 2020.

[33] K. Chen, L. Yu, and J. Guo, "Characteristics analysis of road network reliability in beijing based on the data logs from Taxis," in *Proceedings of the Transportation Research Board Meeting*, Washington DC, USA, January 2007.

[34] D. Taş, N. Dellaert, T. V. Woensel, and T. D. Kok, "Vehicle routing problem with stochastic travel times including soft time windows and service costs," *Computers & Operations Research*, vol. 40, pp. 214–224, 2013.

[35] R. A. Russell and T. L. Urban, "Vehicle routing with soft time windows and erlang travel times," *Journal of the Operational Research Society*, vol. 59, no. 9, pp. 1220–1228, 2008.

[36] J. Chen, Z. Liu, S. Zhu, and W. Wang, "Design of limited-stop bus service with capacity constraint and stochastic travel time," *Transportation Research Part E: Logistics and Transportation Review*, vol. 83, pp. 1–15, 2015.

[37] M. Xiao, S. Chien, and D. Hu, "Optimizing coordinated transfer with probabilistic vehicle arrivals and passengers' walking time," *Journal of Advanced Transportation*, vol. 50, no. 8, pp. 2306–2322, 2016.

[38] D. Taş, N. Dellaert, T. V. Woensel, and T. D. Kok, "The time-dependent vehicle routing problem with soft time windows and stochastic travel times," *Transportation Research Part C: Emerging Technologies*, vol. 48, pp. 66–83, 2014.

[39] S. Chien, Z. Yang, and E. Hou, "Genetic algorithm approach for transit route planning and design," *Journal of Transportation Engineering*, vol. 127, no. 3, pp. 200–207, 2001.

[40] B. Li, D. Krushinsky, and H. A. Reijers, "The share-a-ride problem with stochastic travel times and stochastic delivery locations," *Transportation Research Part C: Emerging Technologies*, vol. 67, pp. 95–108, 2016.

[41] P. Kitjacharoenchai, B. C. Min, and S. Lee, "Two echelon vehicle routing problem with drones in last mile delivery," *International Journal of Production Economics*, vol. 225, 2020.

[42] X. Dong, D. Li, Y. Yin, S. Ding, and Z. Cao, "Integrated Optimization of Train Stop Planning and Timetabling for Commuter Railways with an Extended Adaptive Large Neighborhood Search Metaheuristic Approach," *Transportation Research Part C: Emerging Technologies*, vol. 117, 2020.

[43] M. Soto, M. Sevaux, A. Rossi, and A. Reinholz, "Multiple neighborhood search, tabu search and ejection chains for the multi-depot open vehicle routing problem," *Computers & Industrial Engineering*, vol. 107, pp. 211–222, 2017.

[44] S. Belhaiza, "A hybrid adaptive large neighborhood heuristic for a real-life dial-a-ride problem," *Algorithms*, vol. 12, 2019.

[45] M. Moshref-Javadi, A. Hemmati, and M. Winkenbach, "A truck and drones model for last-mile delivery: a mathematical model and heuristic approach," *Applied Mathematical Modelling*, vol. 80, pp. 290–318, 2020.

[46] M. A. Masmoudi, K. Braekers, M. Masmoudi, and A. Dammak, "A hybrid genetic algorithm for the heterogeneous dial-a-ride problem," *Computers & Operations Research*, vol. 81, pp. 1–13, 2017.

[47] A. Baniamerian, M. Bashiri, and R. M. Tavakkoli, "Modified variable neighborhood search and genetic algorithm for profitable heterogeneous vehicle routing problem with cross-docking," *Applied Soft Computing*, vol. 75, pp. 441–460, 2019.

[48] Z. Han, Y. Chen, H. Li, K. Zhuang, and J. Sun, "Customized bus network design based on individual reservation demands," *Sustainability*, vol. 11, 2019.

[49] S. Han, H. Fu, J. Zhao, J. Lin, and W. Zeng, "Modelling and simulation of hierarchical scheduling of real-time responsive customised bus," *IET Intelligent Transport Systems*, vol. 14, no. 12, pp. 1615–1625, 2020.

[50] X. Dou, Q. Meng, and K. Liu, "Customized bus service design for uncertain commuting travel demand," *Transportmetrica: Transportation Science*, vol. 17, pp. 1–26, 2020.

[51] C. Shen, Y. Sun, Z. Bai, and H. Cui, "Real-time customized bus routes design with optimal passenger and vehicle matching based on column generation algorithm," *Physica A: Statistical Mechanics and Its Applications*, vol. 571, 2021.

[52] J. G. Wardrop, "Road paper. Some theoretical aspects of road traffic research," *Proceedings - Institution of Civil Engineers*, vol. 1, no. 3, pp. 325–362, 1952.

[53] A. Polus, "A study of travel time and reliability on arterial routes," *Transportation*, vol. 8, no. 2, pp. 141–151, 1979.

[54] G. C. Dandy and E. A. McBean, "Variability of individual travel time components," *Journal of Transportation Engineering*, vol. 110, no. 3, pp. 340–356, 1984.

Hindawi
Journal of Advanced Transportation
Volume 2021, Article ID 6619539, 28 pages
https://doi.org/10.1155/2021/6619539

Research Article

Two-Echelon Multidepot Logistics Network Design with Resource Sharing

Siyu Luo,[1] **Yong Wang**,[1] **Jinjun Tang**,[2] **Xiangyang Guan**,[3] **and Maozeng Xu**[1]

[1]*School of Economics and Management, Chongqing Jiaotong University, Chongqing 400074, China*
[2]*Smart Transport Key Laboratory of Hunan Province, School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China*
[3]*Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195, USA*

Correspondence should be addressed to Yong Wang; yongwx@cqjtu.edu.cn and Jinjun Tang; jinjuntang@csu.edu.cn

Received 17 August 2021; Revised 9 September 2021; Accepted 18 September 2021; Published 11 October 2021

Academic Editor: Wenxiang Li

Resource sharing within a logistics network offers an effective way to solve problems resulting from inefficient and costly operations of individual logistics facilities. However, the existing analysis of resource sharing and profit allocation is still limited. Therefore, this study aims to model resource sharing in two-echelon delivery and pickup logistics networks to improve the overall efficiency and decrease the total network operating cost. A bi-objective integer programming model is first proposed for two-echelon collaborative multidepot pickup and delivery problems with time windows (2E-CMDPDTW) to seek the minimization of operating costs and number of vehicles. Integrating a customer clustering algorithm, a greedy algorithm, and an improved nondominated sorting genetic algorithm-II (Im-NSGA-II), a hybrid method is then designed to handle the 2E-CMDPDTW model. The customer clustering and the greedy algorithms are employed to generate locally optimized initial solutions to accelerate the calculating velocity and guarantee the diversity of feasible solutions. The Im-NSGA-II combines the order crossover operation and the polynomial mutation process to find the optimal solution of the 2E-CMDPDTW. The comparative results show that the proposed hybrid method outperforms the NSGA-II and the multiobjective genetic algorithm. Furthermore, a Shapley value method is used for allocating total profits of established alliances and finding an optimal coalition sequence of the logistics facilities joining alliances based on the strictly monotonic path strategy. Finally, a case study of 2E-CMDPDTW in Chongqing China is conducted to validate the feasibility. Results indicate that this study contributes to long-term partnerships between logistics facilities within multi-echelon logistics networks in practice and contributes to the long-term sustainability of urban logistics pickup and delivery networks' development.

1. Introduction

The multidepot pickup and delivery problem (MDPDP) is a generalization of the multidepot vehicle routing problem (MDVRP) aiming to satisfy customer demands through optimizing vehicle routes among logistics facilities, delivery customers, and pickup customers in a multidepot logistics network. To achieve efficiency improvement under practical constraints, the collaborative multidepot pickup and delivery problem with time windows (CMDPDTW) has been proposed [1, 2]. Demands for more efficient logistics, for example, the urban express delivery, are expanding as the logistics sector rapidly develops and the complexity of logistics networks increases. In 2019, the total expenditure of logistics in China reached 298 trillion RMB, and the express volume of commodities rose by 24% to 63 billion RMB from the year earlier [3]. It is important to improve the performance of the logistics networks simultaneously considering pickup and delivery demands via transportation resource optimization and cost reduction. Logistics pickup and delivery networks (LDPN) are facing a shift from simple structures (e.g., one echelon) to complex ones (e.g., multiple echelons) because of extensive logistics demands and diverse customer requirements. In addition, the optimization of two-echelon logistics pickup and delivery networks (2E-LPDN) and the design of collaborative mechanisms are

conducive to build intelligent transportation systems and design sustainable logistics networks.

In traditional two-echelon multidepot logistics pickup and delivery networks (2E-MDLPDN), each logistics facility operates independently. In other words, each logistics facility always serves its customers regardless of where a customer is located. The independent operation mode causes operating problems, such as long-haul and cross transportations, long delivery periods, and transportation resource underutilization in the traditional logistics networks [4]. In addition to operating problems within the network, service functionality and capability of logistics facilities are also critical. For a logistics facility, cross-regional demands of customers may cause service unavailable or high service costs, and insufficient or redundant service capabilities are common and result in customer loss or resource underutilization in practice. Thus, reasonable resource configuration and vehicle scheduling among multiple logistics facilities are crucial for designing an optimized logistics network structure.

For the underused resources and service capabilities, the mode of a sharing economy provides individuals and groups potential ways to save costs and obtain profits, which is promising and validated in transportation-related issues such as car-sharing services such as Uber and Zipcar [5, 6]. When designing a logistics transportation network, designers usually face the problem of how to simultaneously maximize the profits and minimize the operating cost of the whole network when each logistics facility has idle vehicles and excessive service requirements owing to independent operation. The idea of sharing economy is referred to as a considerable candidate to solve the mentioned problem and further optimize the designed logistics network. Therefore, both the transportation routing optimization and the global resource integration in the 2E-MDLPDN are required and necessary.

In the 2E-MDLPDN, products are gathered and then transported among logistics facilities, such as pickup centers and distribution centers, in the first echelon, and pickup and delivery activities between the appropriate logistics facilities and their customers happen in the second echelon. This configuration of the network provides a collaborative basis for logistics facilities in the network, and it is critical and effective to establish collaborative alliances for facilitating resource sharing [7]. In terms of sharing customer information and transportation resources, the scattered customers originally belonging to different logistics facilities can be reallocated to appropriate ones to reduce costs and meet the constraints (e.g., time windows for service). Despite the advantages of collaboration, a critical issue on allocating the cost savings (i.e., profits produced via alliances) is its contributions to the stability of an alliance [8, 9]. With cost savings and reasonable profit allocations, alliances can be more stable and attractive to potential logistics facilities trapped in a traditional network. Therefore, this study aims to provide an effective approach to promote sharing of logistics resources and profit allocation in the two-echelon collaborative multidepot pickup and delivery problem with time windows (2E-CMDPDTW) to optimize the logistics network and minimize logistics costs.

In this study, a 2E-CMDPDTW is proposed and solved considering the collaboration among multiple pickup centers and distribution centers within the logistics network. The transportation resources and customer information are shared in the 2E-CMDPDTW based on establishing collaborative alliances, and then, the services of different logistics facilities can be integrated to reduce operating costs and increase revenues of logistics facilities. A bi-objective integer programming model is first proposed to minimize the total operating cost of the whole logistics network and the number of vehicles used by all logistics facilities. Then, a hybrid method is developed to solve the 2E-CMDPDTW optimization, which includes a customer clustering algorithm and a greedy algorithm to generate the initial solution and the improved nondominated sorting genetic algorithm-II (Im-NSGA-II) to find the optimal solution. Finally, a profit allocation method, namely, the Shapley value model, is employed to achieve an efficient allocation of the total profit to the logistics facilities participating in the collaboration, and the strictly monotonic path (SMP) strategy is used for determining the optimal joining sequence that guarantees the alliance stability. Both the developed bi-objective optimization model and the designed intelligent algorithm provide insights into solving two-echelon multidepot logistics network optimization problems with resource sharing in the urban logistics transportation. Besides, sharing resources within the urban logistics transportation system is demonstrated and validated using real-world case data. Consequently, this study is conducive to improving the operation and management of intelligent transportation systems as well as the development of smart cities.

The remainder of this study is organized as follows. The related studies are reviewed and summarized in Section 2. The problem statement and mathematical model of the 2E-CMDPDTW optimization are presented in Section 3. A three-stage method consisting of customer reassignment and initial solution generation, two-echelon route optimization, and profit allocation and alliance stability analysis is discussed in Section 4. A case study of a 2E-LPDN in Chongqing is conducted to illustrate and demonstrate the proposed model and the hybrid method in Section 5, followed by conclusions and future research directions.

2. Literature Review

Traditional multidepot vehicle routing problems with time windows (MDVRPTW) attempt to seek the optimal routes to minimize traveled distances of the vehicles and the total cost and satisfy customers' demands. Due to the complexity of the network structure and the diversity of customers' demands, 2E-CMDPDTW referring to an extension of the MDVRPTW concentrates on the collaboration between logistics facilities [7]. Therefore, the literature pertaining to the MDVRPTW, collaboration and resource sharing in logistics networks, solution methods for the 2E-CMDPDTW, and profit allocation methods in collaborative networks are reviewed.

2.1. MDVRPTW. In contrast to the traditional VRP, the MDVRP enables vehicles to start from multiple depots and return to the original depot or other depots after providing corresponding pickup or delivery services. In practice, the time windows of pickup and delivery processes can greatly affect customers' satisfaction degrees with logistics service quality. Thus, related constraints for time windows are commonly introduced in the MDVRP, in which a new problem is produced and named MDVRPTW [10]. The MDVRPTW and its derivatives have received much attention both in the theoretical and practical fields. For example, Rabbouch et al. [11] reviewed the relevant literature concerning the MDVRPTW and proved that the satisfaction of time windows is important and necessary.

The MDVRPTW is the same as the MDVRP referring to an NP-hard optimization problem and mainly solved with heuristic algorithms [12]. Many researchers, such as Surekha and Sumathi [13] and the review of Karakatič and Podgorelec [14], employed genetic algorithms or related genetic approaches to solve the MDVRPTW. In addition, Cordeau et al. [15] and Schneider [16] studied a tabu algorithm and an improved tabu algorithm. An ant colony optimization (ACO) is also a prominent method to solving these problems [17–19]. To reduce transportation costs and environmental impacts, a collaborative scheme is integrated into MDVRPTW optimizations [20]. For the collaboration in the multicenter vehicle routing problems (MCVRP), Xu et al. [21] established a mixed-integer linear programming model to study a collaborative multidepot petrol station replenishment problem with multicompartments and the assignment of the customers' time windows. Considering the horizontal collaboration between distribution centers, Muñoz-Villamizar et al. [22] optimized the urban logistics network to minimize the overall cost and reduce environmental impacts. The existing research of the MDVRPTW provides a research framework for the proposed 2E-CMDPDTW in this study. Moreover, research on collaboration provides a research basis for resource sharing in the MDPDP.

2.2. Collaboration and Logistics Resource Sharing. The collaboration and resource sharing of logistics facilities in a logistics network attracts researchers and practitioners as they contribute to cost savings and efficiency improvement [23, 24]. In recent years, many scholars have studied the collaboration and resource sharing of logistics facilities in the MDVRP with simultaneous pickup and delivery problems (PDP) [25, 26]. Quintero-Araujo et al. [27] discussed the horizontal collaboration of logistics facilities in logistics networks and proved that the efficiency of fully collaborative logistics networks is higher than that of semicollaborative or noncollaborative logistics networks. Besides the economic benefits, such as transportation cost reduction, both the society and the environment are beneficial through alleviating traffic congestions and reducing greenhouse gas emissions. To ensure the stability of collaborative logistics networks, the negative impacts of the uncertainties on resource matching are required to be reduced [28]. With the increasing complexity of logistics networks, resource sharing has been modeled into the optimization model of the multi-echelon logistics delivery and pickup network design and can further reduce operating costs of logistics networks [1, 27, 28]. Considering the resource sharing of multiple distribution centers in a two-echelon distribution network, Wang et al. [23] proposed a bi-objective mixed-integer programming model to minimize the operation cost and carbon dioxide emissions of the whole logistics network. Lin [29] studied the coordination and sharing of transportation resources in a PDP with time windows, in which the transportation resources are divided into van trains and foot couriers and proposed a coordination strategy to optimize the service time and cost. For the last-mile PDP, Bhasker et al. [30] presented an optimization model to reduce the delivery and pickup cost of the collaborative logistics network and analyzed the sharing of the extra profits generated by the collaboration of logistics enterprises. Therefore, the resource sharing through the collaboration of multiple pickup centers and distribution centers is integrated to optimize the 2E-MDLPDN.

2.3. Relevant Solution Methods for the 2E-CMDPDTW. The previous efforts on solving VRP and their variants provide a solid foundation for the 2E-CMDPDTW, and they can be categorized according to the logistics tasks, configurations of logistics networks, and operating modes. For example, when a task is conducted by a single DC, vehicle routing problems with simultaneous pickups and deliveries (VRPSPD) are usually developed without time window violations [31, 32]; otherwise, simultaneous delivery and pickup problems with time windows (SDPPTW) are formulated [33, 34]. In addition, a multiproduct vehicle routing problem with time windows (MPVRPTW) [35] is formulated when delivery tasks with several products are conducted by a single DC. A basic two-echelon multidepot vehicle routing problem (2E-MDVRP) [36] can be described as an extension of a two-echelon vehicle routing problem with time windows (2E-VRPTW) [37]. Both of them focus on optimizing the vehicle routes in the first- and second-echelon logistics networks to achieve certain objectives (e.g., minimizing logistics operating costs and maximizing customer satisfaction).

Due to the diversities of fleets, pickup, and delivery demands, number of depots, and region division, various problems are further developed, such as multivehicle multidepot pickup and delivery problems with time windows (M-MDPDPTW) [38], multidepot heterogeneous-fleet pickup, delivery problems with soft time windows (MDHPDSTW) [39], multidepot vehicle routing problems with multiple pickup and delivery demands (MDVRPMPDD) [40], and two-region multidepot pickup and delivery problems (2R-MDPDP) [41]. Moreover, two-echelon vehicle routing problems with pickup and delivery (2E-VRPPD) [42] and two-echelon vehicle routing problems with simultaneous pickup and delivery (2E-VRPSPD) [43] are formulated based on the number of depots (e.g., single or multiple) and the existence of time window violations. For

operating modes, cooperation usually occurs among DCs or multidepots, and thus, horizontal collaboration in multidepot vehicle routing problems (HC-MDVRP) [44] and collaborative multicenter vehicle routing problems (CMCVRP) [45] are further investigated. Similarly, two-echelon logistics pickup and delivery problems (2E-LPDP) [46] are formulated and contribute to performance improvement via cooperation and the satisfaction of customers' demands. The comparative results of related efforts and this work are presented in Table 1.

Research on PDP, MDVRPTW, and their variants established an optimization model with a single objective to optimize the logistics network, especially to minimize the cost, waiting time of customers, and number of transportation tools of the network [31, 37, 45]. However, compared with the optimization model with a single objective, the optimization model with multiobjective can evaluate the efficiency of the logistics network effectively to better optimize the logistics network [47]. For example, Govindan et al. [35] established a bi-objective mixed-integer programming model, and the network cost was considered as the first objective function, and the environmental and social impacts were integrated into the second objective function to optimize the supplier network of multiple commodities. Pérez-Bernabeu et al. [44] developed an iterated local search algorithm to solve the problem of horizontal cooperation between road transportation enterprises and minimize transportation costs and exhaust emissions of vehicles. Besides, the number of vehicles was chosen as one of the objectives. To optimize the MDVRPMPDD, Kachitvichyanukul et al. [40] established a multiobjective metaheuristic model that the objective functions consisted of the total logistics cost, number of vehicles, and number of assigned demands. The PDP is an NP-hard optimization problem, and the 2E-CMDPDTW expands single-echelon logistics networks of the PDP to two-echelon logistics networks. Hence, the 2E-CMDPDTW is also an NP-hard optimization problem and searching for the optimal solution with exact algorithms in large-scale PDPs is cost expensive. Therefore, an efficient heuristic algorithm is required. Goksal et al. [32], Wang et al. [45], and Belgin et al. [43] proposed heuristic solution approaches to study the VRPSPD, CMCVRP, and 2E-VRPSPD, respectively.

Despite the powerful computation capacity of heuristic algorithms, more efforts are also made to simplify the design and optimization problems in the two-echelon logistics networks. Clustering algorithms have been combined with heuristic algorithms to solve MDVRP, and the integrated algorithms are capable of effectively reducing computational time [25]. To improve the convergence speed, Crainic et al. [48] developed clustering-based heuristics to solve the 2E-VRPTW. Meihua et al. [49] improved the ACO with multiple neighborhood descent in terms of distance-based clusters. Liu and Liao [50] integrated the k-means clustering with an adaptive large neighborhood search algorithm to simplify the process of solving a collaborative 2E-VRPTW. Moreover, clustering algorithms also play significant roles in improving the performance of heuristic algorithms and simplifying the computing process for large-scale optimization problems. Therefore, in this study, the solution efficiency of the proposed 2E-CMDPDTW can be improved by adopting the clustering algorithm.

In recent years, multiobjective heuristic algorithms have become widespread since multiple optimization objectives are inherently conflicting in logistics networks. Paul and Shill [51] proposed an NSGA-II to optimize the two validity indices of the clustering processes, and the results reflected that the proposed NSGA-II has good performance in both gene representation and nongene representation datasets. Vachhani et al. [52] proposed a new agglomerative hierarchical clustering process to accelerate the convergence process of the NSGA-II, and the experiments on the standard test of the multiobjective evolutionary algorithm showed that the improved NSGA-II is better than NSGA-II in terms of the existing diversity method. Due to the high efficiency in solving multiobjective optimization problems, the NSGA-II has been widely used and improved to solve the model for the design and optimization of the logistics network. Bandyopadhyay and Bhattacharya [53] proposed an improved NSGA-II to solve the triobjective problem including minimizing the total delivery cost, the variance of order quantity, and the total inventory cost of the two-echelon logistics network. Li et al. [54] improved a hybrid evolutionary algorithm integrating a local search and the NSGA-II to optimize a reverse logistics network with the minimum total cost and maximum benefits. Wang et al. [46] presented an improved NSGA-II based on the clustering algorithm to solve the MDVRP in a 2E-LPDN.

2.4. Profit Allocation. Both the operating efficiency and cost of logistics systems can be improved via resource sharing in collaborative networks. Reasonable profit allocation methods can further facilitate the stable development of collaborative alliances. Considering the fairness preferences of different supply chain members, Wen and Li [55] developed two profit allocation models to analyze the impact of profit allocation on the stability of the supply chain. Focusing on asymmetric information, Liu et al. [56] also presented a profit allocation model for the alliance consisting of different logistics enterprises. Moreover, the game theory is also prominent to solve profit allocation problems. Wang et al. [57] used the game quadratic programming to allocate profits in a cooperative distribution network with time window assignments. In reverse logistics, Liu and Liao [50] proposed an improved Shapley value model to promote the distribution of total profit, and the calculation of transportation costs and carbon emissions is the basis of the model.

Despite the abovementioned contributions to 2E-CMDPDTW optimization, its following aspects still need attention. (1) The collaboration among multiple distribution centers or pickup centers in a delivery or pickup network has been discussed, but the resource sharing between two-echelon logistics facilities, especially in the 2E-MDLPDN, is still limited. (2) Multiobjective optimization problems are overlooked. The economic factors such as network operating

Table 1: Comparison between the abovementioned research studies and this work.

Reference	Multiple depots	Multiple echelons	Pickup and delivery	Time windows	Logistics collaboration	Resources sharing	Multiple objectives	Customer clustering	Heuristic algorithm	Hybrid algorithm	Problem studied
Subramanian et al. [31]			✓						✓	✓	VRPSPD
Goksal et al. [32]			✓						✓	✓	VRPSPD
Wang and Chen [33]			✓	✓					✓		SDPPTW
Mahmoudi and Zhou [34]			✓	✓							SDPPTW
Govindan et al. [35]	✓	✓		✓			✓		✓	✓	MPVRPTW
Zhou et al. [36]	✓	✓	✓						✓	✓	2E-MDVRP
Li et al. [37]		✓		✓				✓	✓		2E-VRPTM
Dridi et al. [38]	✓		✓	✓					✓		M-MDPDPTW
Bettinelli et al. [39]	✓		✓	✓							MDHPDSTW
Kachitvichyanukul et al. [40]	✓		✓	✓			✓		✓		MDVRPMPDD
Soriano et al. [41]	✓	✓	✓	✓					✓		2R-MDPDP
Martins et al. [42]		✓	✓	✓					✓	✓	2E-VRPPD
Belgin et al. [43]	✓	✓	✓						✓	✓	2E-VRPSPD
Pérez-Bernabeu et al. [44]	✓				✓	✓	✓	✓	✓	✓	HC-MDVRP
Wang et al. [45]	✓				✓			✓	✓	✓	CMCVRP
Wang et al. [46]	✓	✓	✓		✓	✓			✓	✓	2E-LPDP
Adelzadeh et al. [47]	✓			✓			✓	✓	✓	✓	MDVRPTW
This work	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	2E-CMDPDTW

Note: ✓ denotes that the content is considered.

cost are usually modeled in the 2E-MDLPDN optimization problems. Transportation resources, for instance, the number of vehicles, should be taken into account for the 2E-MDLPDN optimization. (3) Difficulties (e.g., time-consuming convergence and low-quality solution) often occur when using the conventional heuristic algorithms to solve a large-scale 2E-MDLPDN. On one hand, existing customer clustering algorithms developed to improve computational efficiency are not appropriate for the 2E-CMDPDTW proposed in this study. On the other hand, convergence speed acceleration via the integration of heuristic algorithms to improve the quality of the initial feasible solutions needs to be further explored. (4) More efforts should be made on a collaborative 2E-MDLPDN instead of the current focus on horizontal collaboration and profit allocation in a single-echelon logistics network. The profits of a collaborative alliance should be reasonably distributed in a multi-echelon logistics network to stabilize long-term partnerships. In addition, the evaluation of the stability of established alliances and the analysis of the optimal order of each member joining the alliance are necessary.

In response to the research needs, the major contributions of this study can be summarized in the following aspects. (1) Constructing a collaborative 2E-MDLPDN, and discussing the influence of resource sharing on the network optimization. (2) Establishing a bi-objective integer programming model for minimizing the total operating cost and the number of vehicles for the 2E-CMDPDTW. (3) Designing a hybrid method integrating the customer clustering process, the greedy algorithm, and the Im-NSGA-II to solve the optimization model efficiently. (4) Proposing an optimal alliance sequence selection method on the basis of the Shapley value method and the SMP method to solve the problems of profit allocation and the order of joining an alliance.

3. Problem Statement and Mathematical Model for 2E-CMDPDTW

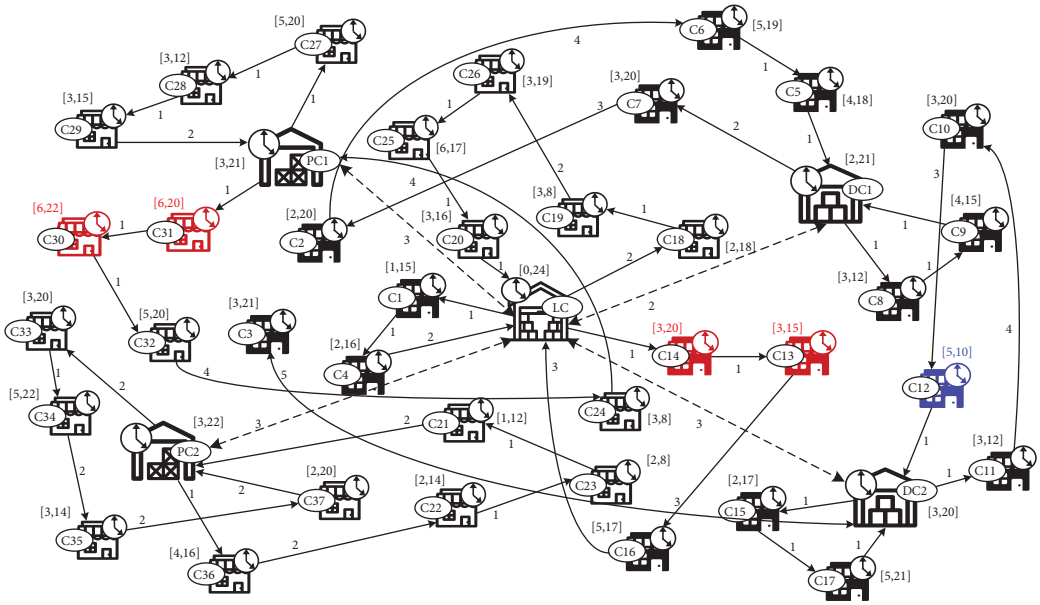
3.1. Problem Statement. Collaboration between logistics facilities in the network contributes to resource integration and optimization so as to minimize the overall operating cost and maximize the revenue of individual logistics facilities. In this study, the 2E-CMDPDTW consists of different logistics facilities containing a logistics center (LC), multiple pickup centers (PCs), and multiple distribution centers (DCs) as the first echelon and a certain number of delivery and pickup customers (Cs) as the second echelon in the 2E-LPDN. Of these facilities, the LC is responsible for collecting and distributing products; and, PCs and DCs provide pickup and delivery service for their customers, respectively. As for the transportation resources, transportation tasks between different logistics facilities in the first-echelon are carried out by semitrailer trucks. For the second-echelon transportation service between a logistics facility and its customers, vehicles are utilized to provide door-to-door services. Fixed time windows are requested by all facilities and customers in the network. If customers were not served within their time windows, the network operating cost would increase along with the decrease in service efficiency.

The structures of the 2E-CMDPDTW optimization before the collaboration and after the collaboration are illustrated in Figure 1. Without collaboration, logistics facilities operate independently. In comparison, collaborating logistics facilities in the alliance can share customers and transportation resources. In Figure 1, the numbers attached to the links represent the travel time of semitrailer trucks or vehicles on the routes.

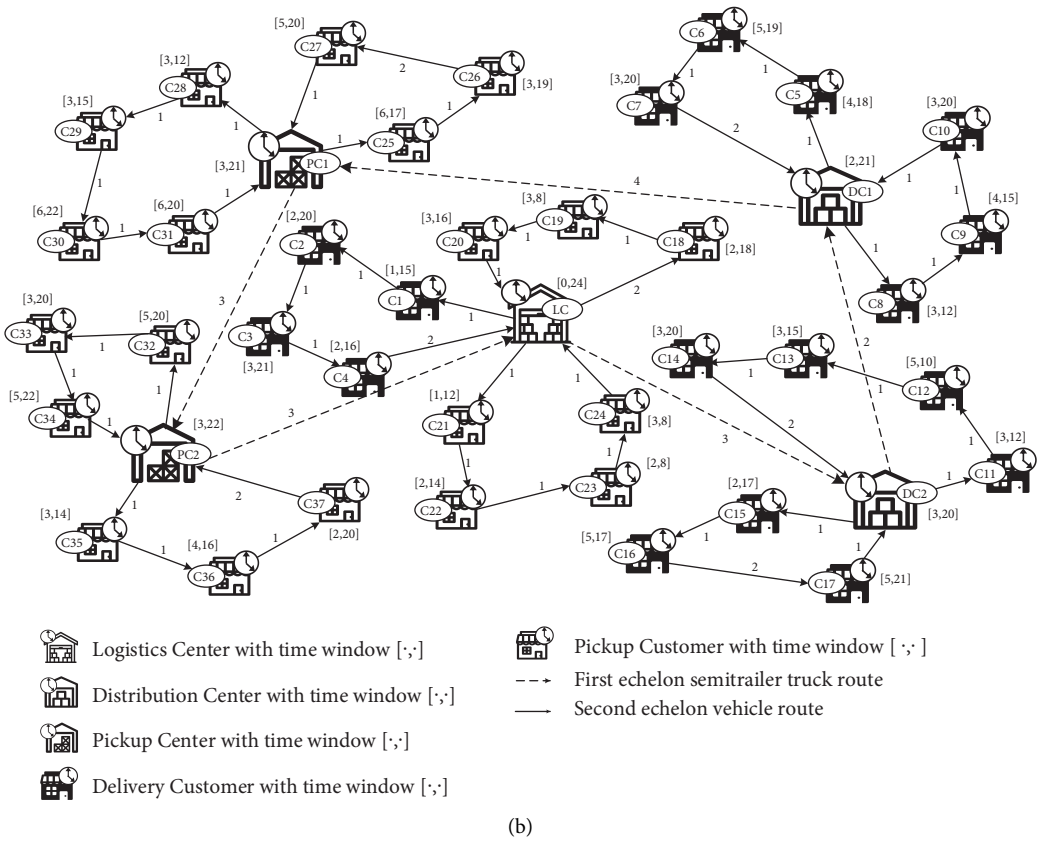
In the noncollaborative 2E-LPDN before optimization (Figure 1(a)), the LC serves DCs, PCs, and its customers with various demands simultaneously; however, DCs and PCs only serve their customers with delivery and pickup demands, respectively. Three problems arise consequently. (1) There are unreasonable transportation routes (e.g., long-haul and cross transportation). For example, $C2$ is closer to the LC but served by DC1. This situation can be observed in practice in terms of customer loyalty and long-term contractual partnership. (2) Customers' time windows are frequently violated. The vehicle, for instance, starting from the LC must wait for open time windows when arriving at $C13$ and $C14$. The demand of $C12$ cannot be satisfied on time since the arrival of the vehicle from DC2 is beyond the given time window. (3) Transportation resources are underutilized. For example, the demand for DC1 served by the LC can be satisfied with a vehicle rather than a semitrailer truck. In other words, other facilities can also be served with the truck. In response to the aforementioned problems, collaboration with resource sharing is effective for efficiency improvement and reducing resource underutilization.

Figure 1(b) presents the collaborative 2E-LPDN after optimization. The LC plays an important role in organizing delivery tasks and pickup tasks and coordinating logistics facilities within the network. The advantages of collaboration are concluded as follows. First of all, with customer information sharing, the customers in the alliance can be reassigned to new service providers closer to them. In addition, the transportation routes both in the first and second echelon can be optimized to guarantee the given service time windows. Finally, sharing transportation resources, such as vehicles and semitrailer trucks, can effectively avoid idle transportation resources in the collaborative logistics network. For example, the customers re-arranged in the route LC $\longrightarrow C21 \longrightarrow C22 \longrightarrow C23 \longrightarrow C24 \longrightarrow$ LC, and the other route PC2 $\longrightarrow$ C32 $\longrightarrow$ C33 $\longrightarrow$ PC2 can be served orderly by the two vehicles on a single transportation task. As a consequence, the aim of reducing the total cost and the number of vehicles is achieved in the optimized 2E-CMDPDTW.

Table 2 shows the comparative results of the costs, number of trucks, and number of vehicles before and after the 2E-CMDPDTW optimization. Assume that the transportation cost per unit time of the semitrailer trucks equals \$30, and all of the delivery cost and pickup cost of vehicles equal \$15. The unit time penalty cost of time windows violation is \$10, and each of DCs and PCs pays



(a)



(b)

Figure 1: The 2E-LPDN before and after collaboration. (a) Two-echelon noncollaborative pickup and delivery network. (b) Two-echelon collaborative pickup and delivery network.

Table 2: Comparison before and after collaboration in the two-echelon pickup and delivery network.

Case	Transportation cost ($)	Pickup and delivery cost ($)	Penalty cost ($)	Collaborative cost ($)	Total operating cost ($)	Time windows violation	Number of trucks	Number of vehicles
Before collaboration	660	1320	70	0	2050	5	4	12
After collaboration	450	825	0	500	1775	0	1	11

for a fixed collaborative cost of $100 when deciding to collaborate.

In Table 2, the total operating cost of the optimized 2E-CMDPDTW is decreased by $275, and 3 semitrailer trucks and 1 vehicle are saved. In the network before collaboration, the number of violating time windows equals 5, while all the customers' time windows are satisfied in the collaborative one. The results reflect that the total operating cost, the number of trucks, the number of vehicles, and the number of violating time windows in the 2E-CMDPDTW are reduced. Therefore, resource sharing can reduce operating costs and save transportation resources in the entire network.

3.2. Definitions. In this section, the definition, notations, and decision variables used in this study for the 2E-CMDPDTW optimization are introduced. The definitions and the decision variables are listed in Tables 3 and 4, respectively.

3.3. Mathematical Model. In this study, a bi-objective mathematical model is developed to optimize the total operating cost of the network and the number of vehicles in the 2E-CMDPDTW. The objective functions in equations (1) and (2) are as follows:

$$\text{Min } F_1 = \text{TC}_1 + \text{TC}_2 + \text{PC}_1 + \text{PC}_2 + \text{TC}_3, \tag{1}$$

$$\text{Min } F_2 = \sum_{m\in M}\sum_{v\in V}\sum_{u\in U_d\cup U_p} x_{muv}. \tag{2}$$

TC_1 expresses the total operating cost of semitrailer trucks, and it includes the total transportation cost and the total maintenance cost in alliances:

$$\text{TC}_1 = \sum_{m,m'\in M, m\neq m'}\sum_{s\in S} x_{mm's}\times L_{mm's}\times f_s\times p_s + \max_{s\in S}\left\{\sum_{m\in L}\sum_{i\in I} x_{mis}, \sum_{m\in L}\sum_{j\in J} x_{mjs}\right\}\times\frac{M_s}{B}, \tag{3}$$

where $\sum_{m,m'\in M,m\neq m'}\sum_{s\in S} x_{mm's}\times L_{mm's}\times f_s\times p_s$ represents the transportation cost of semitrailer trucks and $\max_{s\in S}\left\{\sum_{m\in L}\sum_{i\in I} x_{mis}, \sum_{m\in L}\sum_{j\in J} x_{mjs}\right\}\times(M_s/B)$ depicts the maintenance cost of semitrailer trucks.

TC_2 expresses the total delivery cost, pickup cost, and maintenance cost of vehicles in the second echelon in a working period:

$$\text{TC}_2 = \left(\sum_{\substack{d,u_d^2\in D\\ d\neq u_d^2}}\sum_{v\in V} x_{du_d^2v}\times L_{du_d^2v}\times f_v\times p_v + \sum_{m\in I}\sum_{u\in U_d^2}\sum_{v\in V} x_{muv}\times\frac{M_v}{B}\right) + \left(\sum_{\substack{p,u_p^2\in P\\ p\neq u_p^2}}\sum_{v\in V} x_{pu_p^2v}\times L_{pu_p^2v}\times f_v\times p + \sum_{m\in J}\sum_{u\in U_p^2}\sum_{v\in V} x_{muv}\times\frac{M_v}{B}\right)$$
$$+ \left(\sum_{\substack{l_d,u_d^1\in L_d\\ l_d\neq u_d^1}}\sum_{v\in V} x_{l_du_d^1v}\times L_{l_du_d^1v}\times f_v\times p_v + \sum_{m\in L}\sum_{u\in U_d^1}\sum_{v\in V} x_{muv}\times\frac{M_v}{B}\right) + \left(\sum_{\substack{l_p,u_p^1\in L_p\\ l_p\neq u_p^1}}\sum_{v\in V} x_{l_pu_p^1v}\times L_{l_pu_p^1v}\times f_v\times p_v + \sum_{m\in L}\sum_{u\in U_p^1}\sum_{v\in V} x_{muv}\times\frac{M_v}{B}\right), \tag{4}$$

Table 3: Definitions and notations related to the mathematical model.

Sets	Definitions
L	Set of LC, $L = \{L_0\}$, $l \in L$
I	Set of DCs, $I = \{i \mid i = 1, 2, 3, \ldots, h\}$
J	Set of PCs, $J = \{j \mid j = 1, 2, 3, \ldots, h'\}$
U_d	Set of all delivery customers in the collaborative 2E-LPDN, $U_d = \{u_d \mid u_d = 1, 2, 3, \ldots, u\}$
U_p	Set of all pickup customers in the collaborative 2E-LPDN, $U_p = \{u_p \mid u_p = 1, 2, 3, \ldots, u'\}$
M	Set of all logistics facilities, containing the LC, DCs, and PCs in the collaborative 2E-LPDN, $M = \{m \mid m = 1, 2, 3, \ldots, h, h+1, \ldots, h+h', h+h'+1\}$
U_d^1	Set of delivery customers served by the LC, $U_d^1 = \{u_d^1 \mid u_d^1 = 1, 2, 3, \ldots, u^1\}$
U_d^2	Set of delivery customers served by the DCs, $U_d^2 = \{u_d^2 \mid u_d^2 = 1, 2, 3, \ldots, u^2\}$
U_p^1	Set of pickup customers served by the LC, $U_p^1 = \{u_p^1 \mid u_p^1 = 1, 2, 3, \ldots, u^{1\prime}\}$
U_p^2	Set of pickup customers served by the PCs, $U_p^2 = \{u_p^2 \mid u_p^2 = 1, 2, 3, \ldots, u^{2\prime}\}$
D	Set of the DCs and their delivery customers, $D = I \cup U_d^2 = \{d \mid d = 1, 2, 3, \ldots, h, h+1, \ldots, h+u^2\}$
P	Set of the PCs and their pickup customers, $P = J \cup U_p^2 = \{p \mid p = 1, 2, 3, \ldots, h', h'+1, \ldots, h'+u^{2\prime}\}$
L_d	Set of the LC and its delivery customers, $L_d = L \cup U_d^1 = \{l_d \mid l_d = 1, 2, 3, \ldots, u^1+1\}$
L_p	Set of the LC and its pickup customers, $L_p = L \cup U_p^1 = \{l_p \mid l_p = 1, 2, 3, \ldots, u^2+1\}$
S	Set of semitrailer trucks used for product transportation among LC, DCs, and PCs, $S = \{s \mid tsn = q1, 2, 3h, \ldots x, 7s'\}$
V	Set of vehicles used for product transportation between logistics facilities and customers, $V = \{v \mid v = 1, 2, 3, \ldots, v'\}$
O_s	Set of the DCs and the PCs served by semitrailer truck s, $s \in S$
S_{iv}^d	Set of delivery customers whose delivery service are provided by vehicle v from DC i, $S_{iv}^d \subseteq U_d^2$, $i \in I$ and $v \in V$
S_{jv}^p	Set of pickup customers whose pickup service are provided by vehicle v from PC j, $S_{jv}^p \subseteq U_p^2$, $j \in J$ and $v \in V$
S_{lv}^d	Set of delivery customers whose delivery service are provided by vehicle v from LC l, $S_{lv}^d \subseteq U_d^1$, $l \in L$ and $v \in V$
S_{lv}^p	Set of pickup customers whose pickup service are provided by vehicle v from LC l, $S_{lv}^p \subseteq U_p^1$, $l \in L$ and $v \in V$
Parameters	
$L_{mm's}$	Distance of semitrailer truck s starts from logistics facility m to m', $s \in S$, m and $m' \in M$, and $m \neq m'$
$L_{du_d^2 v}$	Distance of vehicle v starts from the DC or delivery customer d to u_d^2, $v \in V$, $d \in D$, and $u_d^2 \in U_d^2$
$L_{pu_p^2 v}$	Distance of vehicle v starts from the PC or pickup customer p to u_p^2, $v \in V$, $p \in P$, and $u_p^2 \in U_p^2$
$L_{l_d u_d^1 v}$	Distance of vehicle v starts from the LC or delivery customer l_d to u_d^1, $v \in V$, $l_d \in L_d$, and $u_d^1 \in U_d^1$
$L_{l_p u_p^1 v}$	Distance of vehicle v starts from the LC or pickup customer l_p to u_p^1, $v \in V$, $l_p \in L_p$, and $u_p^1 \in U_p^1$
f_s	Diesel fuel consumption rate per mile for semitrailer truck s (gallon/mile), $s \in S$
f_v	Petrol fuel consumption rate per mile for vehicle v (gallon/mile), $v \in V$
p_s	Unit price of diesel used in semitrailer trucks (dollar/gallon)
p_v	Unit price of petrol used in vehicles (dollar/gallon)
M_s	Average annual cost for the maintenance of semitrailer trucks
M_v	Average annual cost for the maintenance of vehicles
B	Number of working periods per year
Q_i	Demand quantity of DC i for delivery, $i \in I$
Q_j	Demand quantity of PC j for pickup, $j \in I$
$R_{u_d^1}$	Demand quantity of delivery customer u_d^1 served by the LC, $u_d^1 \in U_d^1$
$R_{u_p^1}$	Demand quantity of pickup customer u_p^1 served by the LC, $u_p^1 \in U_p^1$
$R_{u_d^2}$	Demand quantity of delivery customer u_d^2 served by the DCs, $u_d^2 \in U_d^2$
$R_{u_p^2}$	Demand quantity of pickup customer u_p^2 served by the PCs, $u_p^2 \in U_p^2$
K_s	Approved load capacity of semitrailer truck s, $s \in S$
K_v	Approved load capacity of vehicle v, $v \in V$
F_m	Fixed cost of logistics facility m, $m \in M$
θ	Variable cost coefficient for products' transportation
c_m	Discount on the costs of logistics facility m, if collaboration occurs (dollar), $m \in M$
$\lvert O_s \rvert$	Number of DCs and PCs that provides transportation services by semitrailer truck s, $s \in S$
$\lvert S_{iv}^d \rvert$	Number of delivery customers whose delivery service are provided by vehicle v from DC i, $i \in I$ and $v \in V$
$\lvert S_{jv}^p \rvert$	Number of pickup customers whose pickup service are provided by vehicle v from PC j, $j \in J$ and $v \in V$
$\lvert S_{lv}^d \rvert$	Number of delivery customers whose delivery service are provided by vehicle v from LC l, $l \in L$ and $v \in V$
$\lvert S_{lv}^p \rvert$	Number of pickup customers whose pickup service are provided by vehicle v from LC l, $l \in L$ and $v \in V$
$[\alpha_i, \beta_i]$	Time window of the DC i, $i \in I$
$[\alpha_j, \beta_j]$	Time window of the PC j, $j \in J$
$[\alpha_{u_d}, \beta_{u_d}]$	Time window of delivery customer u_d, $u_d \in U_d$

Table 3: Continued.

Sets	Definitions
$[\alpha_{u_p}, \beta_{u_p}]$	Time window of pick customer u_p, $u_p \in U_p$
T_s	Maximum travel time allowed for semitrailer truck s on every route, $s \in S$
T_v	Maximum travel time allowed for vehicle v on every route, $v \in V$
w_e	The unit penalty cost for semitrailer trucks and vehicles arriving early
w_d	The unit penalty cost for semitrailer trucks and vehicles arriving delay
at_{is}	Time of semitrailer truck s arrival at the DC i, $s \in S$ and $i \in I$
at_{js}	Time of semitrailer truck s arrival at the PC j, $s \in S$ and $j \in J$
at_{iv}	Time of vehicle v arrival at DC i, $v \in V$ and $i \in I$
at_{jv}	Time of vehicle v arrival at PC j, $v \in V$ and $j \in J$
$\mathrm{at}_{u_d v}$	Time of vehicle v arrival at delivery customer u_d, $v \in V$ and $u_d \in U_d$
$\mathrm{at}_{u_p v}$	Time of vehicle v arrival at pickup customer u_p, $v \in V$ and $u_p \in U_p$
$\mathrm{at}_{u_d^2 v}$	Time of vehicle v arrival at delivery customer u_d^2, $v \in V$ and $u_d^2 \in U_d^2$
$\mathrm{at}_{u_p^2 v}$	Time of vehicle v arrival at pickup customer u_p^2, $v \in V$ and $u_p^2 \in U_p^2$
$\mathrm{at}_{u_d^1 v}$	Time of vehicle v arrival at delivery customer u_d^1, $v \in V$ and $u_d^1 \in U_d^1$
$\mathrm{at}_{u_p^1 v}$	Time of vehicle v arrival at pickup customer u_p^1, $v \in V$ and $u_p^1 \in U_p^1$
dt_{ms}	Time of semitrailer truck s depart from logistics facility m, $m \in M$ and $s \in S$
dt_{dv}	Time of vehicle v depart from DC i or delivery customer u_d^2, $v \in V$, $i \in I$, and $u_d^2 \in U_d^2$
dt_{pv}	Time of vehicle v depart from PC j or pickup customer u_p^2, $v \in V$, $j \in J$, and $u_p^2 \in U_p^2$
$\mathrm{dt}_{l_d v}$	Time of vehicle v depart from LC l or delivery customer u_d^1, $v \in V$, $l_d \in L_d$, and $u_d^1 \in U_d^1$
$\mathrm{dt}_{l_p v}$	Time of vehicle v depart from LC l or pickup customer u_p^1, $v \in V$, $l_p \in L_p$, and $u_p^1 \in U_p^1$
$\mathrm{dt}_{u_d^2 v}$	Time of vehicle v depart from delivery customer u_d^2, $v \in V$ and $u_d^2 \in U_d^2$
$\mathrm{dt}_{u_p^2 v}$	Time of vehicle v depart from pickup customer u_p^2, $v \in V$ and $u_p^2 \in U_p^2$
dr_{ls}	Return time of semitrailer truck s to LC l, $l \in L$ and $s \in S$
dr_{lv}	Return time of vehicle v to LC l, $l \in L$ and $v \in V$
$t_{mm's}$	Time of semitrailer truck s travel from logistics facility m to m', $m \in M$ and $s \in S$
t_{mis}	Time of semitrailer truck s travel from logistics facility m to DC i, $m \in M$, $i \in I$, and $s \in S$
t_{mjs}	Time of semitrailer truck s travel from logistics facility m to PC j, $m \in M$, $i \in I$, and $s \in S$
$t_{du_d^2 v}$	Time of vehicle v travel from DC i or u_d^2 to u_d^2, $v \in V$, $d \in D$, $i \in I$, and $u_d^2 \in U_d^2$
$t_{pu_p^2 v}$	Time of vehicle v travel from PC j or u_p^2 to u_p^2, $v \in V$, $p \in P$, $j \in J$, and $u_p^2 \in U_p^2$
$t_{l_d u_d^1 v}$	Time of vehicle v travel from LC l or u_d^1 to u_d^1, $v \in V$, $l_d \in L_d$, and $u_d^1 \in U_d^1$
$t_{l_p u_p^1 v}$	Time of vehicle v travel from LC l or u_p^1 to u_p^1, $v \in V$, $l_p \in L_p$, and $u_p^1 \in U_p^1$
$t_{u_d^2 iv}$	Time of vehicle v travel from u_d^2 to DC i, $v \in V$, $i \in I$, and $u_d^2 \in U_d^2$
$t_{u_p^2 jv}$	Time of vehicle v travel from u_p^2 to PC j, $v \in V$, $j \in J$, and $u_p^2 \in U_p^2$

Table 4: Decision variables related to the mathematical model.

Variables	Definitions
$x_{mm's}$	$x_{mm's} = 1$, if semitrailer truck s starts from logistics facility m to m'; otherwise, $x_{mm's} = 0$, m and $m' \in M$ and $s \in S$
x_{mis}	$x_{mis} = 1$, if semitrailer truck s starts from logistics facility m to DC i; otherwise, $x_{mis} = 0$, $m \in M$ and $s \in S$
x_{mjs}	$x_{mjs} = 1$, if semitrailer truck s starts from logistics facility m to PC j; otherwise, $x_{mjs} = 0$, $m \in M$ and $s \in S$
$x_{du_d^2 v}$	$x_{du_d^2 v} = 1$, if vehicle v starts from d to u_d^2; otherwise, $x_{du_d^2 v} = 0$, $d \in D$, $v \in V$, and $u_d^2 \in U_d^2$
$x_{pu_p^2 v}$	$x_{pu_p^2 v} = 1$, if vehicle v starts from p to u_p^2; otherwise, $x_{pu_p^2 v} = 0$, $p \in P$, $v \in V$, and $u_p^2 \in U_p^2$
$x_{l_d u_d^1 v}$	$x_{l_d u_d^1 v} = 1$, if vehicle v starts from l_d to u_d^1; otherwise, $x_{l_d u_d^1 v} = 0$, $l_d \in L_d$, $u_d^1 \in U_d^1$, and $v \in V$
$x_{l_p u_p^1 v}$	$x_{l_p u_p^1 v} = 1$, if vehicle v starts from l_p to u_p^1; otherwise, $x_{l_p u_p^1 v} = 0$, $l_p \in L_p$, $u_p^1 \in U_p^1$, and $v \in V$
x_{muv}	$x_{muv} = 1$, if vehicle v starts from logistics facility m to customer u; otherwise, $x_{muv} = 0$, $m \in M$, and $u \in U_d \cup U_p$
Z_m	$Z_m = 1$, if logistics facility m agrees to join the alliance to cooperate; otherwise, $Z_m = 0$, $m \in M$
q_{mis}	Quantity of products transported by semitrailer truck s from logistics facility m to DC i, $m \in M$, $s \in S$, and $i \in I$
q_{mjs}	Quantity of products transported by semitrailer truck s from logistics facility m to PC j, $m \in M$, $s \in S$, and $j \in J$
$q_{mu_d v}$	Quantity of products delivered by vehicle v from logistics facility m to u_d, $v \in V$, $m \in M$, and $u_d \in U_d$
$q_{mu_p v}$	Quantity of products collected by vehicle v from logistics facility m to u_p, $v \in V$, $m \in M$, and $u_p \in U_p$
$q_{iu_d^2 v}$	Quantity of products delivered by vehicle v from DC i to u_d^2, $v \in V$, $i \in I$, and $u_d^2 \in U_d^2$
$q_{ju_p^2 v}$	Quantity of products collected by vehicle v from PC j to u_p^2, $v \in V$, $j \in J$, and $u_p^2 \in U_p^2$
$q_{du_d^2 v}$	Quantity of products delivered by vehicle v from DC i or u_d^2 to u_d^2, $v \in V$, $d \in D$, $i \in I$, and $u_d^2 \in U_d^2$
$q_{pu_p^2 v}$	Quantity of products collected by vehicle v from PC j or u_p^2 to u_p^2, $v \in V$, $p \in P$, $j \in J$, and $u_p^2 \in U_p^2$
$q_{l_d u_d^1 v}$	Quantity of products delivered by vehicle v from LC l or u_d^1 to u_d^1, $v \in V$, $l_d \in L_d$, and $u_d^1 \in U_d^1$
$q_{l_p u_p^1 v}$	Quantity of products collected by vehicle v from LC l or u_p^1 to u_p^1, $v \in V$, $l_p \in L_p$, and $u_p^1 \in U_p^1$

where $\left(\sum_{\substack{d,u_d^2\in D\\ d\neq u_d^2}}\sum_{v\in V}x_{du_d^2v}\times L_{du_d^2v}\times f_v\times p_v+\sum_{m\in I}\sum_{u\in U_d^2}\sum_{v\in V}x_{muv}\times(M_v/B)\right)$ is the delivery cost and maintenance cost of delivery vehicles starting from DCs, $\left(\sum_{\substack{p,u_p^2\in P\\ p\neq u_p^2}}\sum_{v\in V}x_{pu_p^2v}\times L_{pu_p^2v}\times f_v\times p+\sum_{m\in J}\sum_{u\in U_p^2}\sum_{v\in V}x_{muv}\times(M_v/B)\right)$ represents the pickup cost and maintenance cost of pickup vehicles starting from PCs, $\left(\sum_{\substack{l_d,u_d^1\in L_d\\ l_d\neq u_d^1}}\sum_{v\in V}x_{l_du_d^1v}\times L_{l_du_d^1v}\times f_v\times p_v+\sum_{m\in L}\sum_{u\in U_d^1}\sum_{v\in V}x_{muv}\times(M_v/B)\right)$ depicts the delivery cost and maintenance cost of delivery vehicles from the LC, and $\left(\sum_{\substack{l_p,u_p^1\in L_p\\ l_p\neq u_p^1}}\sum_{v\in V}x_{l_pu_p^1v}\times L_{l_pu_p^1v}\times f_v\times p_v+\sum_{m\in L}\sum_{u\in U_p^1}\sum_{v\in V}x_{muv}\times(M_v/B)\right)$ denotes the pickup cost and maintenance cost of pickup vehicles from the LC.

PC_1 expresses the penalty cost of the semitrailer trucks for early and delayed arrival at the DCs and PCs:

$$\begin{aligned}PC_1=&\sum_{m\in M}\sum_{i\in I}\sum_{s\in S}q_{mis}\times w_e\times[\max\{\alpha_i-at_{is},0\}]+\sum_{m\in M}\sum_{i\in I}\sum_{s\in S}q_{mis}\times w_d\times[\max\{at_{is}-\beta_i,0\}]\\&+\sum_{m\in M}\sum_{j\in J}\sum_{s\in S}q_{mjs}\times w_e\times\left[\max\left\{\alpha_j-at_{js},0\right\}\right]+\sum_{m\in M}\sum_{j\in J}\sum_{s\in S}q_{mjs}\times w_d\times\left[\max\left\{at_{js}-\beta_j,0\right\}\right],\end{aligned}\tag{5}$$

where $\sum_{m\in M}\sum_{i\in I}\sum_{s\in S}q_{mis}\times w_e\times[\max\{\alpha_i-at_{is},0\}]$ and $\sum_{m\in M}\sum_{i\in I}\sum_{s\in S}q_{mis}\times w_d\times[\max\{at_{is}-\beta_i,0\}]$ indicate the penalty cost of earliness and delay of the semitrailer trucks for DCs and $\sum_{m\in M}\sum_{j\in J}\sum_{s\in S}q_{mjs}\times w_e\times[\max\{\alpha_j-at_{js},0\}]$ and $\sum_{m\in M}\sum_{j\in J}\sum_{s\in S}q_{mjs}\times w_d\times[\max\{at_{js}-\beta_j,0\}]$ denote the penalty cost of earliness and delay of the semitrailer trucks for PCs.

PC_2 represents the penalty cost of the vehicles for early and delayed arrival at the delivery customers and pickup customers:

$$\begin{aligned}PC_2=&\sum_{m\in M\cup U_d}\sum_{u_d\in U_d}\sum_{v\in V}q_{mu_dv}\times w_e\times\left[\max\left\{a_{u_d}-at_{u_dv},0\right\}\right]+\sum_{m\in M\cup U_d}\sum_{u_d\in U_d}\sum_{v\in V}q_{mu_dv}\times w_d\times\left[\max\left\{at_{u_dv}-\beta_{u_d},0\right\}\right]\\&+\sum_{m\in M\cup U_p}\sum_{u_p\in U_p}\sum_{v\in V}q_{mu_pv}\times w_e\times\left[\max\left\{a_{u_p}-at_{u_pv},0\right\}\right]+\sum_{m\in M\cup U_p}\sum_{u_p\in U_p}\sum_{v\in V}q_{mu_pv}\times w_d\times\left[\max\left\{at_{u_pv}-\beta_{u_p},0\right\}\right],\end{aligned}\tag{6}$$

where $\sum_{m\in M\cup U_d}\sum_{u_d\in U_d}\sum_{v\in V}q_{mu_dv}\times w_e\times[\max\{a_{u_d}-at_{u_dv},0\}]$ and $\sum_{m\in M\cup U_d}\sum_{u_d\in U_d}\sum_{v\in V}q_{mu_dv}\times w_d\times[\max\{at_{u_dv}-\beta_{u_d},0\}]$ depict the penalty cost of earliness and delay of the vehicles for delivery customers and $\sum_{m\in M\cup U_p}\sum_{u_p\in U_p}\sum_{v\in V}q_{mu_pv}\times w_e\times[\max\{a_{u_p}-at_{u_pv},0\}]$ and $\sum_{m\in M\cup U_p}\sum_{u_p\in U_p}\sum_{v\in V}q_{mu_pv}\times w_d\times[\max\{at_{u_pv}-\beta_{u_p},0\}]$ represent the penalty cost of earliness and delay of the vehicles for pickup customers.

TC_3 expresses the total cost of logistics facilities including the fixed and variable costs and the collaborative discount for establishing alliances between logistics facilities:

$$TC_3=\sum_{m\in M}F_m+\sum_{i\in I}\sum_{m\in M}\theta q_{mis}+\sum_{j\in J}\sum_{m\in M}\theta q_{mjs}-\sum_{m\in M}Z_mc_m,\tag{7}$$

where $\sum_{m\in M}F_m$ represents the total fixed cost of the LC, DCs, and PCs, $\sum_{i\in I}\sum_{m\in M}\theta q_{mis}$ and $\sum_{j\in J}\sum_{m\in M}\theta q_{mjs}$ depict the total variable cost of DCs and PCs, respectively, and $\sum_{m\in M}Z_mc_m$ denotes the collaboration discount for establishing alliances in terms of government subsidies.

The constraints of the model are divided into three parts such as constraints for the first echelon, constraints for the second echelon, and constraints for the first and the second echelons.

3.3.1. Constraints for the First Echelon. Constraint (8) ensures that, if $m'=L_0$, a semitrailer truck starts from the LC and returns to the LC after the task is completed. Otherwise, if m' is a DC or PC, this constraint imposes the balance of semitrailer trucks entering and leaving the logistics facility. Constraints (9) and (10) guarantee that each DC and PC only can be served by one semitrailer truck, respectively. Constraint (11) eliminates subtours of every route in the first echelon. Constraints (12) and (13) express that the demands of each DC and PC can be satisfied, respectively. Constraints (14) and (15) ensure that the maximum load of semitrailer trucks is able to satisfy the total demand of the logistics facilities which are served by the semitrailer trucks. Constraints (16) and (17) imply that semitrailer trucks serve DCs and PCs within their time windows, respectively. Constraint (18) represents that the semitrailer truck arrives at the DC. Constraint (19) defines that the semitrailer truck arrives at

the PC. Constraints (20) and (21) denote that the time of each semitrailer truck departure from and return to the logistics facilities meet the time window of the LC. Constraint (22) limits that the total time a semitrailer truck travels on each route must be less than or equal to the allowed time of each route:

$$\sum_{m\in M, m\neq m'} x_{mm's} - \sum_{m\in M, m\neq m'} x_{m'ms} = 0, \quad \forall m' \in M \text{ and } \forall s \in S, \tag{8}$$

$$\sum_{m\in M, m\neq i} \sum_{s\in S} x_{mis} \leq 1, \quad \forall i \in I, \tag{9}$$

$$\sum_{m\in M, m\neq j} \sum_{s\in S} x_{mjs} \leq 1, \quad \forall j \in J, \tag{10}$$

$$\sum_{m\in O_s} \sum_{m'\in O_s, m'\neq m} x_{mm's} = |O_s| - 1, \quad \forall s \in S \text{ and } |O_s| \geq 2, \tag{11}$$

$$\sum_{s\in S} \sum_{m\in M} q_{mis} \times x_{mis} \geq Q_i, \quad \forall i \in I \text{ and } i \neq m, \tag{12}$$

$$\sum_{s\in S} \sum_{m\in M} q_{mjs} \times x_{mjs} \geq Q_i, \quad \forall j \in J \text{ and } j \neq m, \tag{13}$$

$$\sum_{i\in I} \sum_{m\in M} q_{mis} \times x_{mis} \leq K_s, \quad \forall s \in S, \tag{14}$$

$$\sum_{j\in J} \sum_{m\in M} q_{mjs} \times x_{mjs} \leq K_s, \quad \forall s \in S, \tag{15}$$

$$\alpha_i \leq \mathrm{at}_{is} \leq \beta_i, \quad \forall i \in I \text{ and } \forall s \in S, \tag{16}$$

$$\alpha_j \leq \mathrm{at}_{js} \leq \beta_j, \quad \forall j \in J \text{ and } \forall s \in S, \tag{17}$$

$$\mathrm{dt}_{ms} + t_{mis} \leq \mathrm{at}_{is}, \quad \forall m \in M, \forall i \in I, \text{ and } \forall s \in S, \tag{18}$$

$$\mathrm{dt}_{ms} + t_{mjs} \leq \mathrm{at}_{js}, \quad \forall m \in M, \forall j \in J, \text{ and } \forall s \in S, \tag{19}$$

$$\alpha_l \leq \mathrm{dt}_{ms} \leq \beta_l, \quad \forall l \in L, \forall s \in S, \text{ and } m = l_0, \tag{20}$$

$$\alpha_l \leq \mathrm{dr}_{ls} \leq \beta_l, \quad \forall l \in L, \forall s \in S, \text{ and } m = l_0, \tag{21}$$

$$\sum_{m,m'\in M, m\neq m'} t_{mm's} \times x_{mm's} \leq T_s, \quad s \in S. \tag{22}$$

3.3.2. Constraints for the Second Echelon. Constraints (23) and (24) ensure that a delivery vehicle serving DCs' customers originates from a DC to serve customers and returns to the DC when it completed a task. Constraint (25) imposes the balance of delivery vehicles entering and leaving the delivery customers served by DCs. Constraints (26) and (27) guarantee that a pickup vehicle serving PCs' customers originates from a PC and returns to the PC after completing a task. Constraint (28) imposes the balance of pickup vehicles entering and leaving the pickup customers served by PCs. Constraints (29) and (30) express that a delivery vehicle serving the LC originates from the LC and returns to the LC. Constraint (31) imposes the balance of delivery vehicles entering and leaving the delivery customers served by the LC. Constraints (32) and (33) imply that a pickup vehicle serving the LC originates from the LC and returns to the LC. Constraint (34) imposes the balance of pickup vehicles entering and leaving the pickup customers served by the LC. Constraints (35)–(38) are used to eliminate subtours of every route in the second echelon. Constraints (39)–(42) are used to guarantee the demands of each delivery or pickup customer can be satisfied. Constraints (43)–(46) stipulate that the maximum load of each vehicle can satisfy the total demand of customers served by this vehicle. Constraints (47) and (48) represent that all vehicles serve delivery and pickup customers within their time windows. Constraints (49)–(52) define that the vehicle arrives at the delivery and pickup customers. Constraints (53)–(56) limit that the total travel time for a vehicle must be less than or equal to the allowed time of each route:

$$\sum_{u_d^2\in U_d^2} x_{du_d^2v} = 1, \quad \forall d \in I \text{ and } \forall v \in V, \tag{23}$$

$$\sum_{u_d^2\in U_d^2} x_{u_d^2dv} = 1, \quad \forall d \in I \text{ and } \forall v \in V, \tag{24}$$

$$\sum_{u_d^2\in D} x_{du_d^2v} - \sum_{u_d^2\in D} x_{u_d^2dv} = 0, \quad \forall d \in D, d \neq u_d^2, \text{ and } \forall v \in V, \tag{25}$$

$$\sum_{u_p^2\in U_p^2} x_{pu_p^2v} = 1, \quad \forall p \in J \text{ and } \forall v \in V, \tag{26}$$

$$\sum_{u_p^2\in U_p^2} x_{u_p^2pv} = 1, \quad \forall p \in J \text{ and } \forall v \in V, \tag{27}$$

$$\sum_{u_p^2\in P} x_{pu_p^2v} - \sum_{u_p^2\in P} x_{u_p^2pv} = 0, \quad \forall p \in P, p \neq u_p^2, \text{ and } \forall v \in V, \tag{28}$$

$$\sum_{u_d^1\in U_d^1} x_{l_du_d^1v} = 1, \quad \forall l_d \in L_d \text{ and } \forall v \in V, \tag{29}$$

$$\sum_{u_d^1\in U_d^1} x_{u_d^1l_dv} = 1, \quad \forall l_d \in L_d \text{ and } \forall v \in V, \tag{30}$$

$$\sum_{u_d^1\in L_d} x_{l_du_d^1v} - \sum_{u_d^1\in L_d} x_{u_d^1l_dv} = 0, \quad \forall l_d \in L_d, l_d \neq u_d^1, \text{ and } \forall v \in V, \tag{31}$$

$$\sum_{u_p^1\in U_p^1} x_{l_pu_p^1v} = 1, \quad \forall l_p \in L_p \text{ and } \forall v \in V, \tag{32}$$

$$\sum_{u_p^1 \in U_p^1} x_{u_p^1 l_p v} = 1, \quad \forall l_p \in L_p \text{ and } \forall v \in V, \tag{33}$$

$$\sum_{u_p^1 \in L_p} x_{l_p u_p^1 v} - \sum_{u_p^1 \in L_p} x_{u_p^1 l_p v} = 0, \quad \forall l_p \in L_p, l_p \neq u_p^1, \text{ and } \forall v \in V, \tag{34}$$

$$\sum_{d \in S_{iv}^d} \sum_{u_d^2 \in S_{iv}^d} x_{d u_d^2 v} = \left|S_{iv}^d\right| - 1, \quad \forall i \in I, d \neq u_d^2, \forall v \in V, S_{iv}^d \subseteq U_d^2, \text{ and } \left|S_{iv}^d\right| \geq 2, \tag{35}$$

$$\sum_{p \in S_{jv}^p} \sum_{u_p^2 \in S_{jv}^p} x_{p u_p^2 v} = \left|S_{jv}^p\right| - 1, \quad \forall j \in J, p \neq u_p^2, \forall v \in V, S_{jv}^p \subseteq U_p^2, \text{ and } \left|S_{jv}^p\right| \geq 2, \tag{36}$$

$$\sum_{l_d \in S_{lv}^d} \sum_{u_d^1 \in S_{lv}^d} x_{l_d u_d^1 v} = \left|S_{lv}^d\right| - 1, \quad \forall l \in L, l_d \neq u_d^1, \forall v \in V, S_{lv}^d \subseteq U_d^1, \text{ and } \left|S_{lv}^d\right| \geq 2, \tag{37}$$

$$\sum_{l_p \in S_{lv}^p} \sum_{u_p^1 \in S_{lv}^p} x_{l_p u_p^1 v} = \left|S_{lv}^p\right| - 1, \quad \forall l \in L, l_p \neq u_p^1, \forall v \in V, S_{lv}^p \subseteq U_p^1, \text{ and } \left|S_{lv}^p\right| \geq 2, \tag{38}$$

$$\sum_{d \in D} \sum_{v \in V} q_{d u_d^2 v} \times x_{d u_d^2 v} \geq R_{u_d^2}, \quad \forall u_d^2 \in U_d^2 \text{ and } d \neq u_d^2, \tag{39}$$

$$\sum_{p \in P} \sum_{v \in V} q_{p u_p^2 v} \times x_{p u_p^2 v} \geq R_{u_p^2}, \quad \forall u_p^2 \in U_p^2 \text{ and } p \neq u_p^2, \tag{40}$$

$$\sum_{l_d \in L_d} \sum_{v \in V} q_{l_d u_d^1 v} \times x_{l_d u_d^1 v} \geq R_{u_d^1}, \quad \forall u_d^1 \in U_d^1 \text{ and } l_d \neq u_d^1, \tag{41}$$

$$\sum_{l_p \in L_p} \sum_{v \in V} q_{l_p u_p^1 v} \times x_{l_p u_p^1 v} \geq R_{u_p^1}, \quad \forall u_p^1 \in U_p^1 \text{ and } l_p \neq u_p^1, \tag{42}$$

$$\sum_{d \in D} \sum_{u_d^2 \in U_d^2, d \neq u_d^2} q_{d u_d^2 v} \times x_{d u_d^2 v} \leq K_v, \quad \forall v \in V, \tag{43}$$

$$\sum_{p \in P} \sum_{u_p^2 \in U_p^2, p \neq u_p^2} q_{p u_p^2 v} \times x_{p u_p^2 v} \leq K_v, \quad \forall v \in V, \tag{44}$$

$$\sum_{u_d^1 \in U_d^1} \sum_{l_d \in L_d, l_d \neq u_d^1} q_{l_d u_d^1 v} \times x_{l_d u_d^1 v} \leq K_v, \quad \forall v \in V, \tag{45}$$

$$\sum_{u_p^1 \in U_p^1} \sum_{l_p \in L_p, l_p \neq u_p^1} q_{l_p u_p^1 v} \times x_{l_p u_p^1 v} \leq K_v, \quad \forall v \in V, \tag{46}$$

$$\alpha_{u_d} \leq \text{at}_{u_d v} \leq \beta_{u_d}, \quad \forall u_d \in U_d \text{ and } \forall v \in V, \tag{47}$$

$$\alpha_{u_p} \leq \text{at}_{u_p v} \leq \beta_{u_p}, \quad \forall u_p \in U_p \text{ and } \forall v \in V, \tag{48}$$

$$\text{dt}_{dv} + t_{d u_d^2 v} \leq \text{at}_{u_d^2 v}, \quad \forall u_d^2 \in U_d^2, u_d^2 \neq d, \forall d \in D, \text{ and } \forall v \in V, \tag{49}$$

$$\text{dt}_{pv} + t_{p u_p^2 v} \leq \text{at}_{u_p^2 v}, \quad \forall u_p^2 \in U_p^2, u_p^2 \neq p, \forall p \in P, \text{ and } \forall v \in V, \tag{50}$$

$$\text{dt}_{l_d v} + t_{l_d u_d^1 v} \leq \text{at}_{u_d^1 v}, \quad \forall u_d^1 \in U_d^1, u_d^1 \neq l_d, \forall l_d \in L_d, \text{ and } \forall v \in V, \tag{51}$$

$$\text{dt}_{l_p v} + t_{l_p u_p^1 v} \leq \text{at}_{u_p^1 v}, \quad \forall u_p^1 \in U_p^1, u_p^1 \neq l_p, \forall l_p \in L_p, \text{ and } \forall v \in V, \tag{52}$$

$$\sum_{d, u_d^2 \in D, d \neq u_d^2} t_{d u_d^2 v} \times x_{d u_d^2 v} \leq T_v, \quad \forall v \in V, \tag{53}$$

$$\sum_{p, u_p^2 \in P, p \neq u_p^2} t_{p u_p^2 v} \times x_{p u_p^2 v} \leq T_v, \quad \forall v \in V, \tag{54}$$

$$\sum_{l_d, u_d^1 \in L_d, l_d \neq u_d^1} t_{l_d u_d^1 v} \times x_{l_d u_d^1 v} \leq T_v, \quad \forall v \in V, \tag{55}$$

$$\sum_{l_p, u_p^1 \in L_p, l_p \neq u_p^1} t_{l_p u_p^1 v} \times x_{l_p u_p^1 v} \leq T_v, \quad \forall v \in V. \tag{56}$$

3.3.3. Constraints for the First and the Second Echelon. Constraints (57) and (58) are used to connect the transportation of products in the first echelon and the product delivery and pickup process in the second echelon. Constraint (57) specifies the product quantity transported from facility m to DC i, and the delivery quantity is equal to the total demands of the customers served by DC i. Constraint (58) defines the product quantity transported from logistics facility m to PC j, and the pickup quantity is equal to the total demands of the customers served by PC j. Constraints (59) and (60) define that all vehicles must serve DCs and PCs within their time windows, respectively. Constraints (61) and (62) represent that the vehicle arrives at the DCs and PCs, respectively. Constraints (63) and (64) denote that each vehicle departure from and return to the LC meets the time windows of the LC:

$$\sum_{m \in M} \sum_{s \in S} q_{mis} = \sum_{v \in V} \sum_{u_d^2 \in U_d^2} q_{i u_d^2 v} \times x_{i u_d^2 v}, \quad \forall i \in I \text{ and } i \neq m, \tag{57}$$

$$\sum_{m \in M} \sum_{s \in S} q_{mjs} = \sum_{v \in V} \sum_{u_p^2 \in U_p^2} q_{j u_p^2 v} \times x_{j u_p^2 v}, \quad \forall j \in J \text{ and } j \neq m, \tag{58}$$

$$\alpha_i \leq \text{at}_{iv} \leq \beta_i, \quad \forall v \in V, \tag{59}$$

$$\alpha_j \leq \text{at}_{jv} \leq \beta_j, \quad \forall v \in V, \tag{60}$$

$$\mathrm{dt}_{u_d^2 v} + t_{u_d^2 iv} \le \mathrm{at}_{iv}, \quad \forall u_d^2 \in U_d^2, \forall i \in I, \text{ and } \forall v \in V, \tag{61}$$

$$\mathrm{dt}_{u_p^2 v} + t_{u_p^2 jv} \le \mathrm{at}_{jv}, \quad \forall u_p^2 \in U_p^2, \forall j \in J, \text{ and } \forall v \in V, \tag{62}$$

$$\alpha_l \le \mathrm{dt}_{l_d v} \le \beta_l, \quad \forall l \in L, l_d = l_0, \text{ and } \forall v \in V, \tag{63}$$

$$\alpha_l \le \mathrm{dr}_{lv} \le \beta_l, \quad \forall l \in L \text{ and } \forall v \in V. \tag{64}$$

4. Hybrid Solution Approach

To solve the developed mathematical model, a hybrid method including the customers clustering, the greedy algorithm, and the Im-NSGA-II is proposed, and a three-stage method is designed for the 2E-CMDPDTW optimization. In the first stage, customer clustering is used to reassign the customers in the 2E-LPDN. Then, based on the clustering results and the greedy algorithm, initial feasible solutions can be generated. In the second stage, the NSGA-II is employed to search for the optimal route of the two-echelon network. Especially, the second-echelon vehicle routes are first optimized, and then, the first-echelon semitrailer truck routes are generated based on the optimal route of vehicles. The first stage and the second stage are conducted using the proposed hybrid heuristic algorithm. In the third stage, the Shapley value method is used for distributing profits to alliance members. Finally, according to the principle of the SMP, the sequence of the logistics facilities joining the alliance is determined, and the optimal scheme for alliance participants to join the alliance is obtained.

4.1. Hybrid Heuristics for the 2E-CMDPDTW Optimization. The NSGA-II proposed by Deb et al. [58] has been prominent in academic and practice domains in recent years. It is widely accepted that the NSGA-II can solve problems with multiple objectives effectively and efficiently. In the 2E-CMDPDTW, logistics facilities with different functionalities and amounts of customers with delivery and pickup demands are considered, which increases the computational complexity of the network. Therefore, the customer clustering algorithm and the greedy algorithm are introduced into the hybrid heuristic approach. The two-echelon logistics network is partitioned off as small regions using customer clustering, and the greedy algorithm aims to preliminarily generate the local optimal initial solution of each region [59]. Finally, the improved genetic operators are employed in the NSGA-II to find the global solution. The hybrid method is designed and implemented, as shown in Figure 2.

The parameters in Figure 2 can be explained as follows: $k = \{k_1, k_2\}$ represents the number of clusters, in which k_1 and k_2 are the numbers of cluster centers, *pop* is the initial population, t and $t$max represent the current and the maximum generations, respectively, and *run* and $r$max denote the current and the maximum optimization runs, respectively.

4.1.1. Customer Clustering Algorithm. Customer clustering has been proved to be an effective method to simplify the large-scale and multidepot logistics network [7, 60]. Its basic process refers to reassign customers to logistics facilities closer to them. On one hand, to improve convergence speed and reduce computing efforts, the customer clustering algorithm is employed to generate initial feasible solutions. On the other hand, the customer clustering algorithm plays an important role in avoiding long-haul and cross transportation and reducing transportation costs. Algorithm 1 shows the procedure of the designed clustering algorithm.

4.1.2. Greedy Algorithm. The greedy algorithm referring to an approximate algorithm for optimization problems is usually employed to generate initial solutions [61]. The basic principle of the greedy algorithm is to guarantee the best choice in each search. With the greedy algorithm, this study can generate the local optimal solutions instead of the random ones, thereby improving the convergence speed. The major steps for implementing the greedy algorithm are introduced as follows.

Step 1: initialize inputs according to the outputs of customer clustering. C_l^1, C_{di}^2, and C_{pj}^2 represent the sets of customers served by LC l, DC i, and PC j, respectively. Both customers with delivery and pickup demands can be served by the LC; thus, R_d^1 and R_p^1 represent the delivery and pickup route sets of vehicles from LC l. Similarly, R_{di}^2 and R_{pj}^2 represent the route sets of vehicles from DC i and PC j.

Step 2: define conditions for adding subroutes. A new subroute can be added into the vehicle route if the two customers in the subroute appeared less than twice in the vehicle route, either of the two customers does not appear in the vehicle route, and their time windows will not be violated.

Step 3: generate the distance matrix of the customers and sort alternative subroutes. Calculate the distances between the customers with the same kind of requirements, and generate the distance matrix. Then, the lengths of alternative subroutes in the matrix are sorted from the shortest to the longest.

Step 4: add the alternative subroutes according to the sorted results. If the alternative subroute fits the conditions above, it will be added to one of the route sets (e.g., R_d^1, R_p^1, R_{di}^2, or R_{pj}^2) according to the customer sets (i.e., C_l^1, C_{di}^2, and C_{pj}^2). In addition, if the customers in the subroutes have appeared in the vehicle route, the subroute will be directly connected to the existing route with the same customer. Repeat this step until the transportation quantity exceeds the maximum capacity of a vehicle.

Step 5: generate closed vehicle routes. According to the clustering results, the customers in a certain vehicle route are served by the same facility so that a vehicle route is closed by adding the facility its customers belong to.

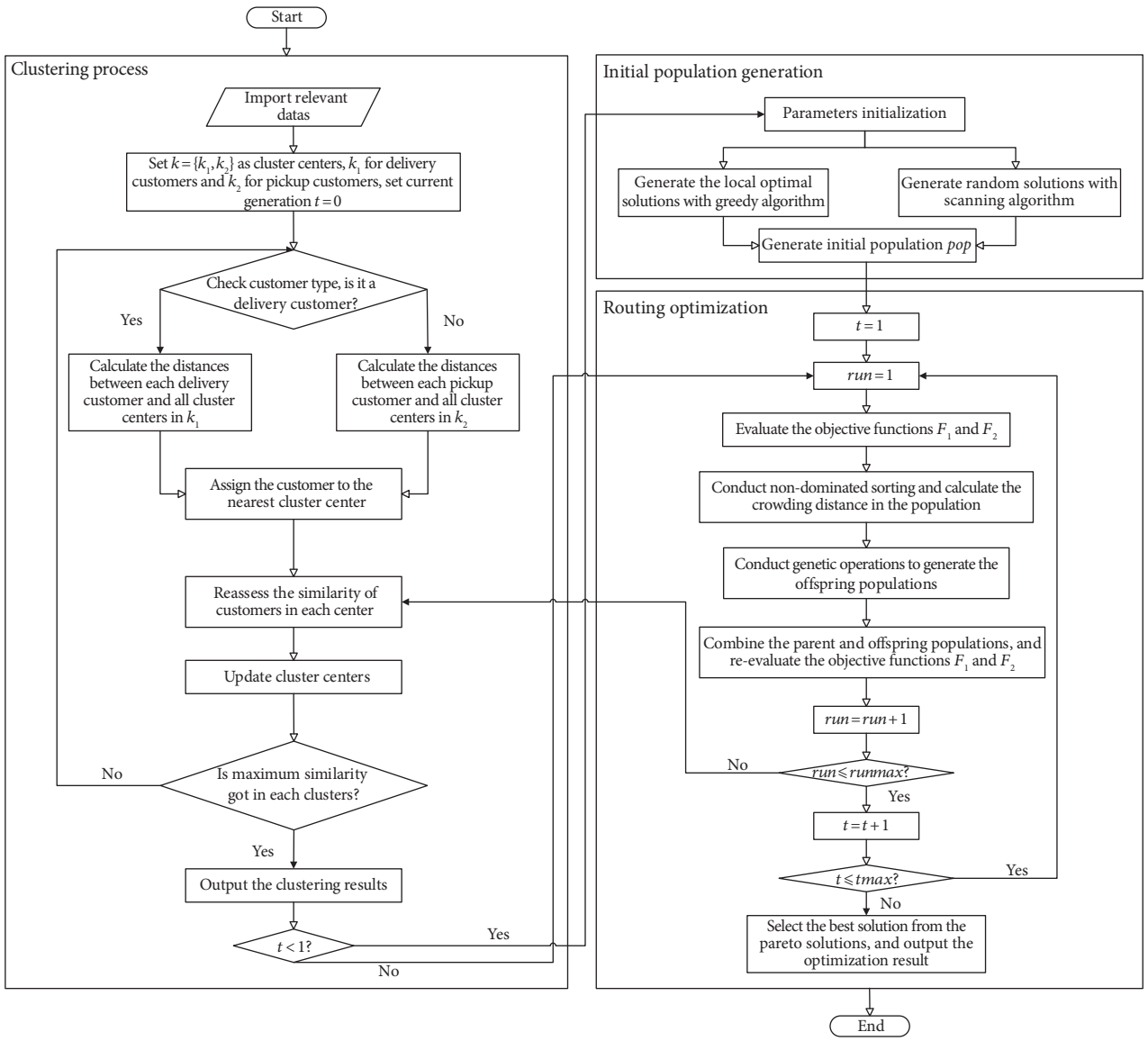


Figure 2: Flowchart of the hybrid heuristic algorithm.

Customer clustering algorithm
Inputs: Geographical coordinates of logistics facilities and geographical coordinates of customers
Outputs: Clustering results

(1) **Step 1:** Set the number of LC and DCs k_1 occupied in delivery tasks, and k_2 indicates the number of LC and PCs handling product pickup demands.
(2) **Step 2:** Select elements in k_1 and k_2 as centers of corresponding clusters.
(3) **For customers with delivery demands**
(4) **Step 3:** Evaluate the distance between each delivery customer and each cluster center belonging to k_1.
(5) **For customers with pickup demands**
(6) **Step 4:** Evaluate the distance between each pickup customer and each cluster center belonging to k_2.
(7) **If** the same distances between a certain customer and several cluster centers exist
Step 5: Assign the customer to one of the cluster centers randomly.
Else
Step 6: Assign the customer to the closest cluster center.
End
(8) **Step 7:** Reassess the similarity of customers in each center, and update the clusters.
(9) **Step 8:** Repeat **Steps 3-6** until every customer is assigned to the closest cluster center.
(10) Print the clustering results

Algorithm 1: Procedure of the designed customer clustering algorithm.

4.1.3. The Improved Nondominated Sorting Genetic Algorithm-II. The NSGA-II, a variant of the NSGA, contributes to balancing multiple objectives and generating the optimal decision on optimization problems [62, 63]. Considering the proposed model, this paper improves the method of generating the initial population and using the more appropriate genetic operators based on the original NSGA-II operating process. To improve the performance of global search and maintain the diversity of a population, the genetic operators including order crossover (OX) [45] and polynomial mutation [57] are exploited. The major procedures of implementing the Im-NSGA-II are presented in Figure 3.

In Figure 3, the individuals in the initial population are first generated using the greedy algorithm and the scanning algorithm enabling the diversity of the population. Second, according to the nondomination sorting method [64], the individuals in population P_t are sorted. Each individual in P_t is assigned with a rank value representing the nondominated level. Selection, cross-over, and mutation operators are then employed to generate offspring in Q_t. A new population R_t can be generated by combining the individuals in P_t and Q_t. Third, a new population P_{t+1} is generated by performing nondominated sorting and crowding calculation in R_t. The above process is repeated until the optimal Pareto solution is selected from all the Pareto solutions.

Generally, it is hard to pursue all optimal objectives simultaneously in a bi-objective problem since they can be conflicting in nature [65]. According to these objectives, the solutions referring to Pareto solutions cannot be compared simply. Figure 4 shows the optimal Pareto solution of the bi-objective problem. Both individuals A and B are in the optimal Pareto front, and the others dominated by the solutions in the optimal Pareto front are in the feasible region. For the bi-objective problem, when objective 1 (e.g., number of vehicles) is deemed as the selection criterion, individual B is selected as the optimal Pareto solution because of fewer vehicles.

The Im-NSGA-II is proposed to seek bi-objective optimal Pareto solutions in the 2E-CMDPDTW, as shown in Algorithm 2. The processes of nondominated sorting, crowding distance calculating, and genetic operations are presented as follows.

(1) *Nondominated Sorting and Crowding Distance Calculating Processes.* The bi-objective values of each individual in the population can be first calculated. Then, the Pareto fronts are generated in accordance to the dominant relationships between the individuals [66]. Finally, the crowding distance is calculated for the individuals, which are selected to generate a new population for iteration based on the rank of Pareto fronts and crowding distances. The calculation process of the crowding distances of individuals is depicted in Figure 5.

In Figure 5, p_r and p_l are two adjacent individuals of individual p with two objective values of O_1 and O_2. The crowding distance of individual p is expressed by D_p, where $D_p = (p_r \times O_1 - p_l \times O_1) + (p_r \times O_2 - p_l \times O_2)$. When selecting individuals to generate a new population, the individuals with the smaller rank in the nondominated solution set will be preferred. If the total number of the selected individuals is bigger than the size of the population, the individuals with the longer crowding distances in the last selected rank will be selected and added to the population [66]. After the maximum number of iterative operations, the solution in the optimal Pareto front (i.e., Rank 1) is the optimal Pareto solution of the bi-objective optimization problem.

(2) *Genetic Operation.* Genetic operation mainly includes crossover operation and mutation operation. The application of the crossover operator can improve the search capability of the Im-NSGA-II. The convergence speed, stability, and global search capability can be improved by employing proper mutation operators [67]. The main processes of crossover operation and mutation operation are as follows.

(1) The order crossover operator is employed in the designed Im-NSGA-II as it outperforms many other operators [68]. First, two chromosomes such as Parent 1 and Parent 2 in the parent generation are selected to eliminate the values representing the logistics facility, and the customers encoding in the two chromosomes are both served by the same facility. Then, randomly select the starting and ending positions of the two preprocessed parent chromosomes. A section of the chromosome is retrieved in terms of two random and disparate points to generate proto-children. Finally, the rest positions in the proto-children will be filled with the genes not retrieved in Parent 1.

(2) The polynomial mutation plays an important role in multiobjective optimization problems [69]. Considering its advantages, the polynomial mutation is used in our proposed Im-NSGA-II. The offspring $v_i^{(1,t+1)}$ are obtained from the parent individuals $v_i^{(1,t)}$ in equation (65), and the relative parameters are calculated in equation (66):

$$v_i^{(1,t+1)} = v_i^{(1,t)} + \beta_i, \tag{65}$$

$$\beta_i = \begin{cases} (2u_i)^{1/(\eta_m+1)} - 1, & u_i < 0.5, \\ 1 - [2(1-u_i)]^{(1/(1+\eta_m))}, & u_i \geq 0.5, \end{cases} \tag{66}$$

where η_m represents the distribution index and u_i is a random value, $u_i \in [0, 1)$.

4.2. Profit Allocation and Alliance Stability

4.2.1. Shapley Value Method. Alliances consisting of logistics facilities contribute to cost savings and efficiency improvement. The cost savings referring to the profit of the alliance should be distributed to alliance members to motivate them to take profitable actions such as resource sharing [70]. Accordingly, the methods for profit allocation are important and necessary. The Shapley value method is utilized in this study. Here, A and A_s are defined as the grand alliance consisting of all logistics facilities in the 2E-CMDPDTW and a suballiance including some of them,

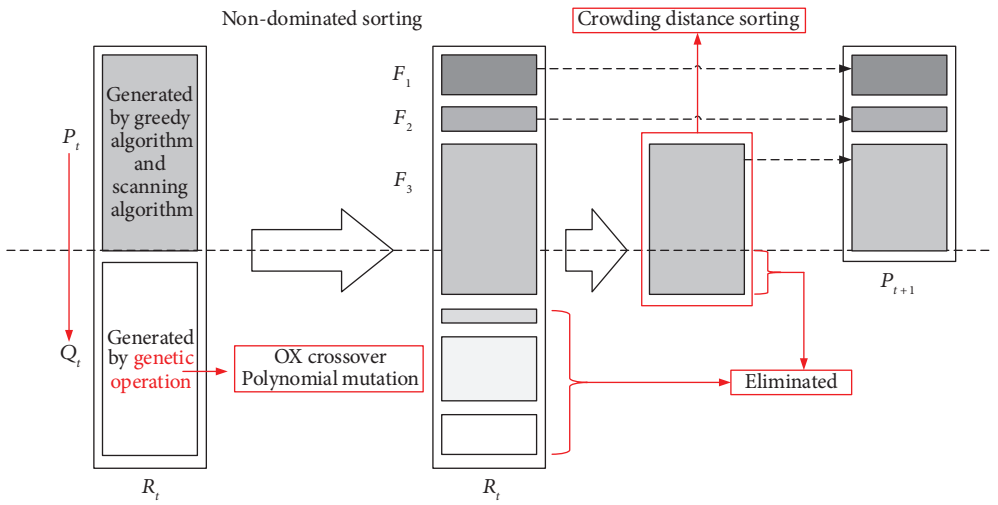


Figure 3: Major procedures of the Im-NSGA-II.

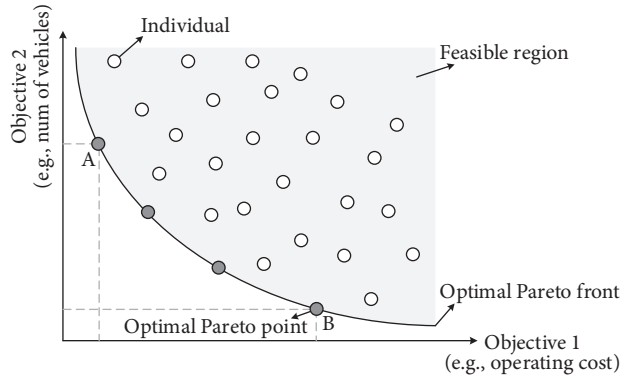


Figure 4: Pareto solutions of the multiobjective optimization problem.

Im-NSGA-II algorithm
Inputs: Population size pop_{size}, maximum number of generations *tmax*, maximum number of optimization runs *rmax*, crossover probability p_{cro}, mutation probability p_{mut}
Outputs: Optimal Pareto solutions
(1) Initialize parameters.
(2) Set system parameters: pop_{size}, *tmax*, *rmax*
(3) Set genetic parameters: p_{cro}, p_{mut}
Implement
(4) **for** run = 1: *rmax*
(5) Generate the initial population by greedy algorithm and scanning algorithm
(6) **for** $i = 1$: *tmax*
(7) **for** $i = 1$: pop_{size}
(8) Evaluate the objective function
(9) **End**
(10) Carry out *order crossover* and *polynomial mutation* for the population
(11) Generate the offspring population Q_i
(12) Combine the parent population P_i and the offspring population Q_i to generate a new population R_i
(13) Perform the *non-dominated sorting* and *crowding distance calculating* for the R_i
(14) Select individuals from the R_i to generate a new population P_{i+1}
(15) **end**
(16) **end**
(17) Obtain the optimal solutions

Algorithm 2: Procedure of the proposed Im-NSGA-II.

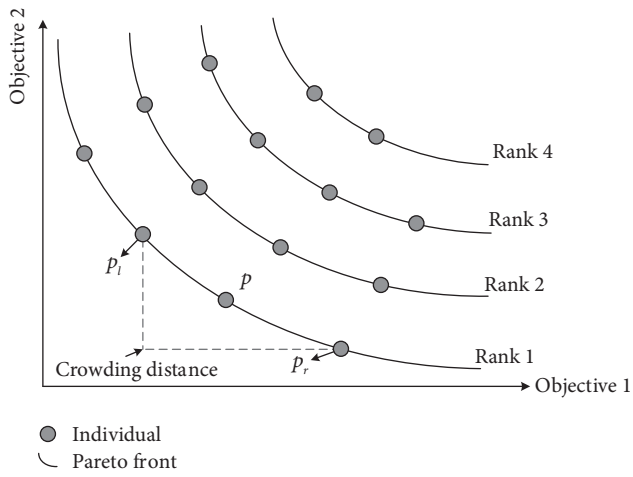


Figure 5: Calculation of crowding distances of individuals.

respectively. The profit of each logistic facility m in the grand alliance A is represented by $\varphi_m(A, v)$ and calculated in equation (67). In general, an alliance organizer responsible to send messages and coordinate members is deserved to gain extra profits. Thus, a given proportion ($\sigma \in (0, 1)$) of the total profit of the alliance is rewarded to the organizer. As for each suballiance, the total profit can be calculated in equation (68):

$$\varphi_m(A, v) = \sum_{m \in A_s, A_s \subseteq A} \frac{\left(|A_s| - 1\right)!\left(|A| - |A_s|\right)!}{|A|!} \times \left[v(A_s) - v(A_s - \{m\})\right], \quad (67)$$

$$v(A_s) = (1 - \sigma) \times \max\left\{ \sum_{m \in A_s} c(\{m\}) - c(A_s), 0 \right\}, \quad (68)$$

where $c(\{m\})$ represents the operating cost of a logistics facility m before taking part in the alliance A_s and $c(A_s)$ depicts the total operating cost of suballiance A_s.

4.2.2. Strictly Monotonic Path Method. The total profit of the alliance is significantly affected by the adding sequence of logistics facilities since the process of establishing the alliance is dynamic. The logistics facility with higher cost savings can be prioritized to join the alliance. The SMP is used to evaluate and select schemes of cooperative organization establishment [69]. In this study, the optimal sequence for logistics facilities to join the alliance is generated based on the SMP method. Let π represent a set of logistics facilities adding in the alliance with sequences and $\eta(m, \pi, k)$denote the decreased percentage of the operating cost of the logistics facility m when joining in the alliance. The decreased percentage can be calculated in equation (69):

$$\eta(m, \pi, k) = \frac{\varphi_m\left(\cup_{\pi(\mu) \le k, \mu \in A} \mu, v\right)}{c(\{m\})}, \quad \pi(m) \le k. \quad (69)$$

As a consequence, the optimal sequence of adding in the alliance can be generated by conducting the SMP method. The selection processes are introduced as follows. According to the diagonal principle [57], the minimum decreased percentages on the diagonal values of the cost decreased percentage matrix are selected. Among these selected decreased percentages, the sequence with the highest decreased percentage is selected as the optimal sequence. If the lowest decreased percentages on the diagonal values are equal, comparisons of the second-lowest decreased percentages of different alliance sequences are then conducted. Among them, the optimal sequence is also determined by the highest decreased percentage. Finally, repeat the above processes until the optimal alliance sequence is obtained.

4.2.3. Alliance Stability. Alliances are the basis for resource sharing and profit allocation in the 2E-CMDPDTW. The stability of the established alliance is the major focus. Despite the implemented profit allocation, the alliance stability is required to be evaluated. Based on the proximity between the schemes of different profit allocation methods and core center values, alliance stability can be guaranteed [68]. The core center value for each alliance is calculated in equation (70):

$$\frac{v(A) - v(A - \{m\})}{v(A)} \times \alpha + \sum_{c \in A, c \ne m} y_c = v(A - \{m\}), \quad (70)$$

where $v(A)$ represents the total profit of the grand alliance A, α is a variable that controls the profit changes before and after the logistics facility m joining the alliance A, and y_c is the core center value of the logistics facility m joining the grand alliance A.

5. Empirical Analyses

5.1. Algorithm Comparison and Analysis. The comparison between the Im-NSGA-II and the other two famous multiobjective optimization algorithms is conducted to evaluate the performance of the designed algorithm. In this section, the NSGA-II and the multiobjective genetic algorithm (MOGA) are employed since they are of interest to researchers in multiobjective optimization fields [71, 72]. The logistics networks in the modified "Solomon datasets" [73] and the "MDVRPTW instances[2]" [69] are selected as benchmarks. The parameters in the datasets are shown in Table 5.

The parameters for optimization algorithm implementation are listed as follows, $pop_{size} = 100$, $g_{\max} = 250$, $p_{cro} = 0.9$, $p_{mut} = 0.1$, and $runmax = 10$. The mentioned three algorithms are implemented ten times for each instance in the datasets, and the optimal results are compared in Table 6.

In Table 6, the designed Im-NSGA-II outperforms the NSGA-II and the MOGA in most instances. In Instances 1 to 7, the Im-NSGA-II can obtain the minimum operating costs. Despite that the computation time by the Im-NSGA-II is longer than that by the NSGA-II and the MOGA, the minimum operating cost and number of vehicles are guaranteed in Instance 3. In Instances 8 to 14, excepting for

Table 5: Parameters in the datasets.

Instance	Solomon datasets	MDVRPTW datasets	Number of customers	Number of delivery customers	Number of delivery customers	Number of depots
1	—	pr_01	48	24	24	4
2	C101	—	50	25	25	3
3	C202	—	50	25	25	3
4	R101	—	50	25	25	3
5	R204	—	50	25	25	3
6	RC104	—	50	25	25	3
7	—	pr_07	72	36	36	6
8	—	pr_02	96	48	48	4
9	C102	—	100	50	50	4
10	C201	—	100	50	50	4
11	R102	—	100	50	50	4
12	RC105	—	100	50	50	4
13	RC203	—	100	50	50	4
14	—	pr_03	144	72	72	4
15	C1_2_1	—	200	100	100	5
16	R1_2_2	—	200	100	100	5
17	RC1_2_4	—	200	100	100	6
18	—	pr_09	216	108	108	6
19	—	pr_05	240	120	120	4
20	—	pr_06	288	144	144	4

Table 6: Comparative results of three algorithms.

Instance	Im-NSGA-II			NSGA-II			MOGA		
	Cost	Vehicle	Time	Cost	Vehicle	Time	Cost	Vehicle	Time
1	1,383	8	60	1,406	9	67	1,442	8	67
2	685	6	64	712	7	69	719	6	70
3	681	5	66	765	6	63	757	6	65
4	1,372	12	65	1,482	12	74	1,476	12	79
5	836	4	66	874	4	77	901	4	76
6	869	6	63	907	6	75	981	7	76
7	1,629	10	68	1,724	10	79	1,689	11	82
8	2,064	12	77	2,195	12	86	2,136	12	84
9	1,368	11	83	1,465	10	96	1,325	10	98
10	970	5	85	992	5	99	989	7	92
11	1,674	7	84	1,798	8	97	1,776	7	95
12	1,938	15	88	2,015	16	98	2,033	18	95
13	1,380	6	86	1,442	9	97	1,436	7	106
14	2,687	17	108	2,793	19	112	2,822	18	117
15	11,954	25	142	12,103	27	151	12,262	27	153
16	10,139	25	145	10,657	26	153	10,263	29	167
17	14,204	23	149	14,109	24	153	14,779	24	157
18	3,053	24	164	3,266	27	172	3,447	25	171
19	3,335	24	208	3,472	26	221	3,591	26	222
20	4,057	29	293	4,285	32	315	4,290	33	312
Average	3,314	14	108	3,423	15	118	3,456	15	119
t test				$-4.283E+00$		$-8.942E+00$	$-4.574E+00$		$-9.401E+00$
p value				$3.626E-04$		$2.004E-08$	$1.840E-04$		$8.849E-09$

Instances 9, the Im-NSGA-II can obtain the minimum values of the operating cost and the computation time, and the number of vehicles is not larger than that by the other algorithms. In Instance 9, both the operating cost and the number of vehicles by the Im-NSGA-II are higher than those in MOGA, while the computation time is minimum. In Instance 17, the Im-NSGA-II obtains a higher operating cost than NSGA-II, but both the number of vehicles and the computation time are the minimum.

5.2. Data Source. A real-world 2E-LPDN in Chongqing China is utilized to illustrate and evaluate the proposed model and solution method in the 2E-CMDPDTW. In Figure 6, a 2E-LPDN consists of one LC, two DCs (i.e., DC1 and DC2), two PCs (i.e., PC1 and PC2), 90 delivery customers, and 92 pickup customers. The logistics center labeled LC is represented by a red node with the letter *L*. The customers served by the LC are divided into two parts: the delivery customers are represented by red nodes with a dot

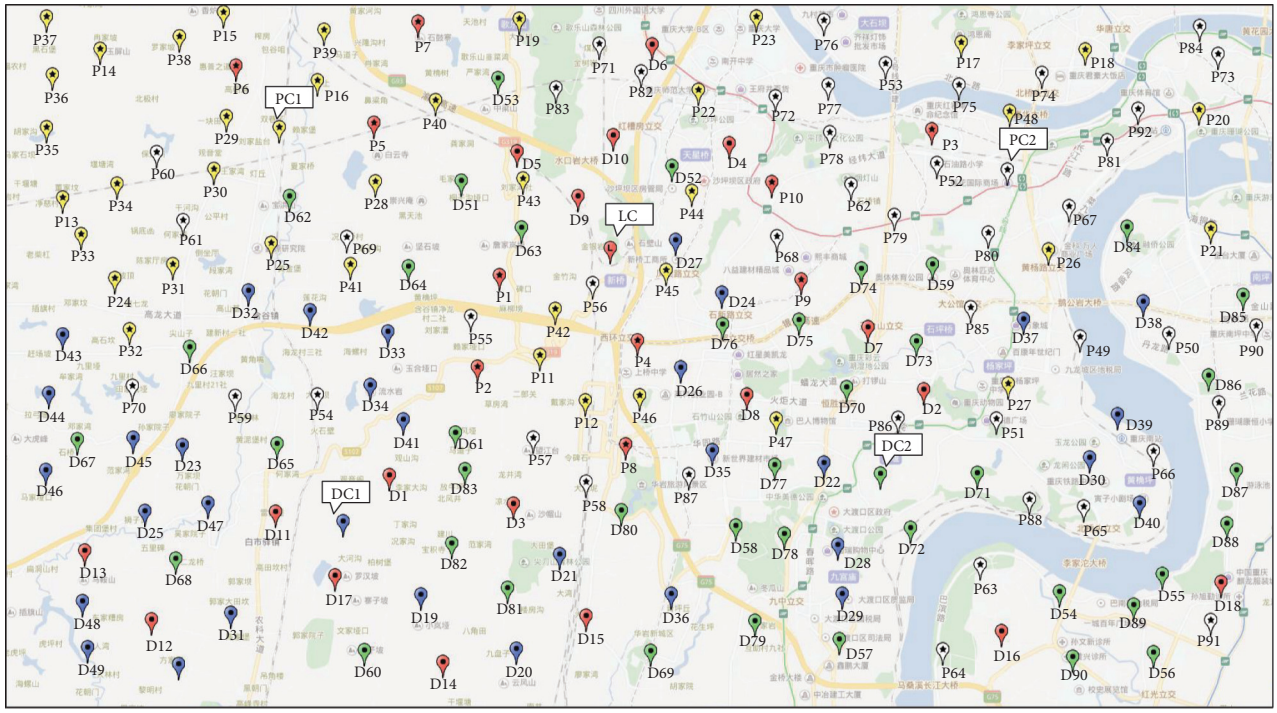


Figure 6: Geographic layouts of the LC, DCs, PCs, and customers.

symbol, and the pickup customers are represented by red nodes with a star symbol. DC1 and DC2 along with their customers are represented by blue and green nodes with a dot symbol, respectively. Similarly, nodes in yellow and white with a star symbol indicate PC1 and PC2 along with their customers, respectively. In addition, the four facilities are labeled DC1, DC2, PC1, and PC2, respectively, and distinguished from their customers. The characteristics of the LC, DCs, PCs, and their customers are presented in Table 7.

5.3. Parameter Setting and Optimization Results. Using the proposed model and solution method, the logistics network is redesigned with the optimal route in terms of operating cost savings and transportation resource reduction. First of all, the customer clustering algorithm is implemented to reassign the 182 customers, and sequentially, the greed algorithm is conducted to generate initial feasible solutions. The results of customer clustering are shown in Figure 7. Finally, the Im-NSGA-II is run to pursue optimal Pareto solutions. The parameters in this case study are based on literature [7, 25, 46, 69] and practical investigations. The parameters referring to the operation of a semitrailer truck and a vehicle are set as $K_s = 1700$, $K_v = 200$, $f_s = 6.6$, $f_v = 3.4$, $M_s = 15000$, $M_v = 1500$, $p_s = 3.4$, and $p_v = 3.2$. The parameters related to the operation and collaboration of logistics facilities are set as $F_1 = 1200$, $F_2 = 800$, $F_3 = 750$, $F_4 = 850$, $F_5 = 900$, $c_1 = 400$, $c_2 = 550$, $c_3 = 450$, $c_4 = 600$, $c_5 = 600$, $w_e = 0.2$, $w_d = 0.3$, and $B = 52$. Besides, the parameters involved in the Im-NSGA-II are $pop_{size} = 100$, $t_{\max} = 500$, $p_{cro} = 0.9$, $p_{mut} = 0.1$, and $runmax = 2$. The established alliances are organized by the LC in the logistics network. Therefore, at most 20 types of alliances are available. According to the proposed model and hybrid heuristic algorithm, we calculate the comparison results of optimization schemes for all forms of alliances shown in Table 8. Figure 8 illustrates the comparison of the optimization schemes.

In Table 8 and Figure 8, the total operating cost, the number of trucks, and the number of vehicles are reduced compared with the initial 2E-LPDN. For example, the decreased percentages of the total cost, the number of vehicles, and the number of semitrailer trucks can reach 36.9%, 11.5%, and 50%, respectively. In addition, an increasing trend of the reduction of the cost savings and the transportation resource is observed as the alliance members increase. For example, the decreased percentage of the total cost increases from 30.86% to 32.89% as well as the number of vehicles from 12.5% to 14.29%, when PC1 has been a member of the alliance {LC, DC1, DC2}. Therefore, establishing the grand alliance must be most effective to redesign and optimize the logistics network in this case. Table 9 shows the comparative results of the initial and the optimized 2E-CMDPDTW. Figure 9 illustrates the comparative results of the network before and after establishing the grand alliance.

In Table 9 and Figure 9, the total cost of the grand alliance includes the transportation cost, pickup and delivery cost, penalty cost, maintenance cost, and collaborative cost. Through establishing a grand alliance, the resource sharing in the 2E-CMDPDTW can be realized to effectively reduce the total cost and the number of vehicles in the 2E-LPDN. For example, the total cost of the initial 2E-CMDPDTW is 31,954, while the total cost of the optimized 2E-CMDPDTW is 20,161 with a decrease value of 11,793. Moreover, the number of trucks and vehicles is 4 and 26 in the initial

Table 7: The characteristics of the LC, DCs, PCs, and customers.

Facilities	Symbol	Number of customers	Customer type	Served customers	Customer symbol
LC	L	28	Delivery customers	*D*1~*D*18	
			Pickup customers	*P*1~*P*10	
DC1		32	Delivery customers	*D*19~*D*50	
DC2		40	Delivery customers	*D*51~*D*90	
PC1		38	Pickup customers	*P*11~*P*48	
PC2		44	Pickup customers	*P*49~*P*90	

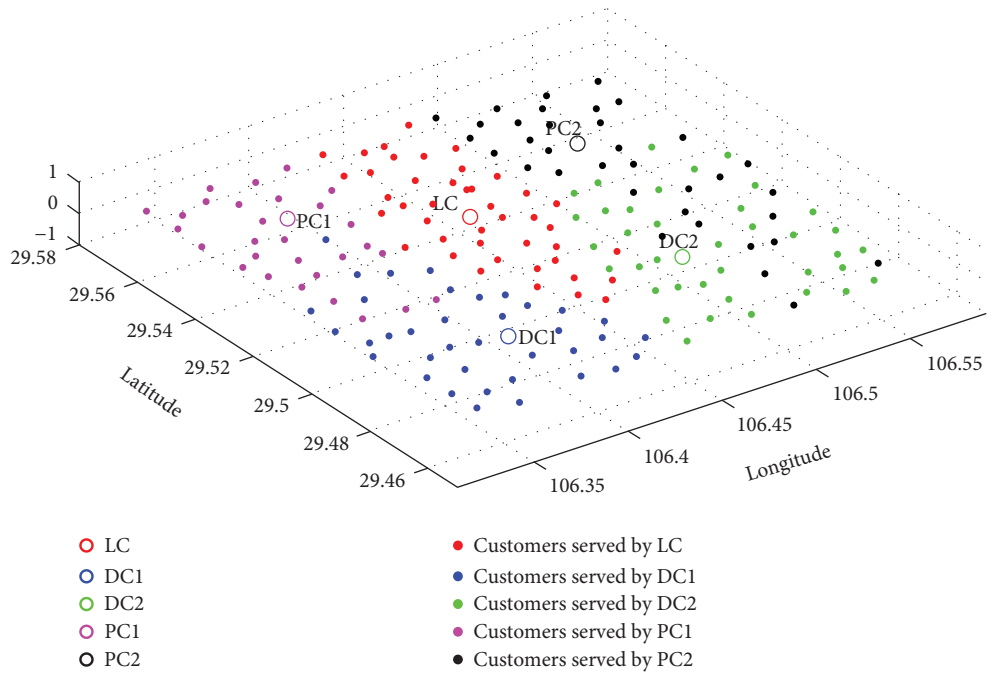


Figure 7: Results of customer clustering.

Table 8: Comparative results of the costs, the number of trucks, and the number of vehicles.

Alliance	Initial network			Optimized network			Savings and reduction		
	Cost	Vehicle	Truck	Cost	Vehicle	Truck	Cost	Vehicle	Truck
{LC}	6,492	5	—	5,793	5	—	699	0	—
{DC1}	5,414	5	—	4,951	4	—	463	1	—
{DC2}	7,057	6	—	6,579	5	—	478	1	—
{PC1}	5,723	5	—	5,056	4	—	667	1	—
{PC2}	7,269	5	—	6,690	5	—	579	0	—
{LC, DC1}	11,905	10	1	8,827	9	1	3,078	1	0
{LC, DC2}	13,548	11	1	10,359	10	1	3,189	1	0
{LC, PC1}	12,214	10	1	8,648	9	1	3,567	1	0
{LC, PC2}	13,761	10	1	10,070	9	1	3,691	1	0
{LC, DC1, DC2}	18,962	16	2	13,111	14	2	5,851	2	0
{LC, DC1, PC1}	17,628	15	2	11,682	13	1	5,947	2	1
{LC, DC1, PC2}	19,174	15	2	14,104	13	1	5,070	2	1
{LC, PC1, DC2}	19,271	16	2	13,214	14	1	6,057	2	1
{LC, PC1, PC2}	19,483	15	2	13,944	14	1	5,540	1	1
{LC, PC2, DC2}	20,817	16	2	14,636	14	1	6,181	2	1
{LC, DC1, DC2, PC1}	26,231	21	3	17,604	18	2	8,627	3	1
{LC, DC1, DC2, PC2}	24,897	21	3	16,662	18	2	8,235	3	1
{LC, PC1, PC2, DC1}	26,540	21	3	17,767	18	2	8,773	3	1
{LC, PC1, PC2, DC2}	25,463	21	3	16,538	19	2	8,925	2	1
{LC, DC1, DC2, PC1, PC2}	31,954	26	4	20,161	23	2	11,793	3	2

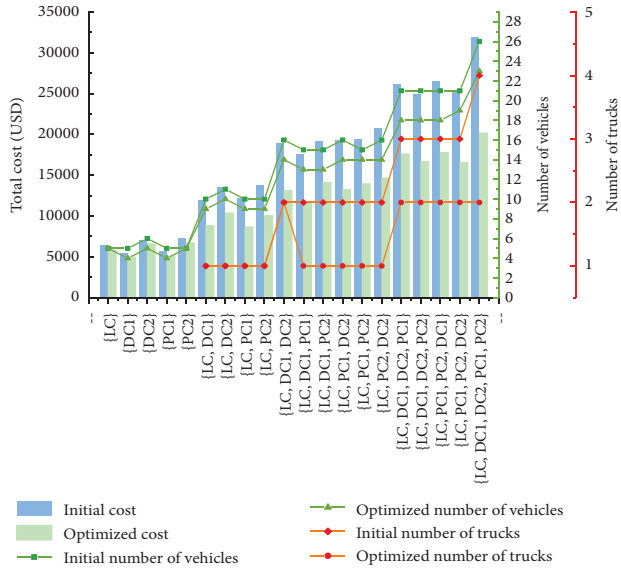


Figure 8: Comparison of the costs, number of trucks, and number of vehicles of alliances.

Table 9: Comparative results of the initial and the optimized 2E-CMDPDTW.

Case	Transportation cost ($)	Pickup and delivery cost ($)	Penalty cost ($)	Maintenance cost ($)	Collaborative cost ($)	Total operating cost ($)	Number of trucks	Number of vehicles
Initial 2E-CMDPDTW	7920	18210	3920	1904	0	31954	4	26
Optimized 2E-CMDPDTW	5859	9381	1081	1240	2600	20161	2	23

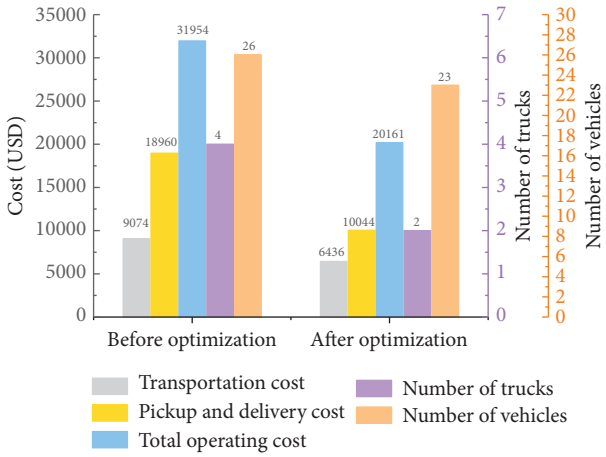


Figure 9: Comparative results of the network before and after establishing the grand alliance.

network, respectively. By contrast, with a value of 2 and 23, the numbers of trucks and numbers of vehicles are reduced by 2 and 3 in the optimized network, respectively.

5.4. Profit Allocation and Sequential Coalition Selection. It is widely accepted that economic benefits are the focus of the majority of business strategies. Similarly, potential profits earned in alliances contribute to decisions making on participation and alliance member sequence, and vice versa. According to the results in the case study, the logistics network can earn an overall profit in terms of the grand alliance. Therefore, the establishment and the stability of alliances are depending on effective profit allocations to alliance members. According to practitioners, the reward percentage is given as $\sigma = 0.12$, which means that 12% of the total profit is rewarded to the alliance provider. The results of the profit allocation in the case are presented, as shown in Table 10.

Compared to the network without optimization, the LC, DC1, DC2, PC1, and PC2 can save costs with values of 699, 463, 478, 677, and 579 in the optimized network. After optimizing the network and establishing the grand alliance, the total cost can be reduced with a value of 11,793 referring to the profit. Accordingly, these logistics facilities can acquire allocated profits with values of 2,684, 2,069, 2165, 2326, and 2,549, respectively. Figure 10 depicts the processes along with the decreased percentages of cost savings of the logistics facilities participating in the alliance. For transparency, two examples are presented: (1) the processes complying with the SMP are represented by solid lines. (2) The others are depicted by dash lines. Starting from each logistic facility, the best coalition sequences satisfying SMP are presented, as shown in Table 11. The optimal sequence {PC1, LC, DC1, PC2, DC2} is obtained based on the diagonal rule, and the decreased proportions of costs for each facility in it are illustrated in Figure 11.

Table 10: Profit of each logistics facility in the 2E-CMDPDTW.

Alliance	Initial cost	Optimized cost	$v(A_s)$	$\varphi_m(A, v)$
{LC}	6,492	5,793	699	(699, 0, 0, 0, 0)
{DC1}	5,414	4,951	463	(0, 463, 0, 0, 0)
{DC2}	7,057	6,579	478	(0, 0, 478, 0, 0)
{PC1}	5,723	5,056	667	(0, 0, 0, 667, 0)
{PC2}	7,269	6,690	579	(0, 0, 0, 0, 579)
{LC, DC1}	11,905	8,827	3,078	(1657, 1421, 0, 0, 0)
{LC, DC2}	13,548	10,359	3,189	(1705, 0, 1484, 0, 0)
{LC, PC1}	12,214	8,648	3,567	(1799, 0, 0, 1768, 0)
{LC, PC2}	13,761	10,070	3,691	(1706, 0, 0, 0, 1985)
{LC, DC1, DC2}	18,962	13,111	5,851	(2389, 1699, 1762, 0, 0)
{LC, DC1, PC1}	17,628	11,682	5,947	(2091, 1755, 0, 2101, 0)
{LC, DC1, PC2}	19,174	14,104	5,070	(2030, 1088, 0, 0, 1952)
{LC, DC2, PC1}	19,271	13,214	6,057	(2139, 0, 1817, 2101, 0)
{LC, PC1, PC2}	19,483	13,944	5,540	(2166, 0, 0, 1428, 1946)
{LC, DC2 PC2}	20,817	14,636	6,181	(2210, 0, 1886, 0, 2085)
{LC, DC1, DC2, PC1}	26,231	17,604	8,627	(2442, 1922, 1998, 2265, 0)
{LC, DC1, DC2, PC2}	24,897	16,662	8,235	(2394, 1853, 1902, 0, 2086)
{LC, DC1, PC1, PC2}	26,540	17,767	8,773	(2465, 1899, 0, 2276, 2133)
{LC, DC2, PC1, PC2}	25,463	16,538	8,925	(2471, 0, 2002, 2306, 2146)
{LC, DC1, DC2, PC1, PC2}	31,954	20,161	11,793	(2684, 2069, 2165, 2326, 2549)

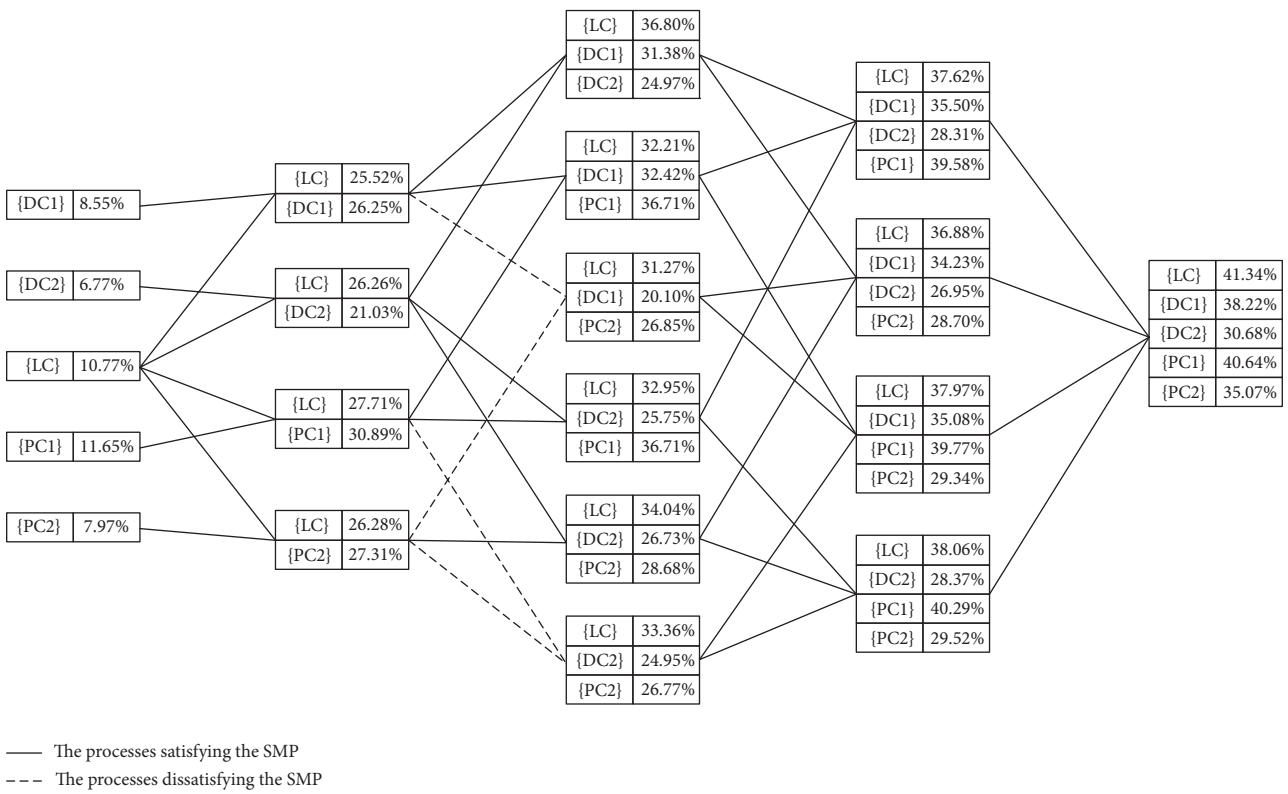


Figure 10: Cost reduction percentages when new members join the alliance.

In Figure 11, as the first member of the alliance, PC1 acquires 11.7% cost savings. The costs of PC1 and LC are decreased by 30.9% and 27.7% when the second member LC participates. Accordingly, when the grand alliance is generated, 40.6%, 41.3%, 38.2%, 36.0%, and 30.7% costs are reduced for PC1, LC, DC1, PC2, and DC2, respectively.

5.5. Alliance Stability. To demonstrate the effectiveness of the employed profit allocation method, the comparisons are conducted between the Shapley value method and other methods such as the equal profit method (EPM) [74], the game quadratic programming (GQP) [57], and the minimum costs-remaining savings (MCRS) [75]. As mentioned

Table 11: Best coalition sequences satisfying SMP from each logistics facility.

π_1 = {DC1, LC, DC2, PC1, PC2}						π_2 = {DC2, LC, PC1, DC1, PC2}					
Player i	DC1	LC	DC2	PC1	PC2	Player i	DC2	LC	PC1	DC1	PC2
η (i, π, 1)	8.6%					η (i, π, 1)	6.8%				
η (i, π, 2)	26.3%	25.5%				η (i, π, 2)	21.0%	26.3%			
η (i, π, 3)	31.4%	36.8%	25.0%			η (i, π, 3)	25.8%	33.0%	36.7%		
η (i, π, 4)	35.5%	37.6%	28.3%	39.6%		η (i, π, 4)	28.3%	37.6%	39.6%	35.5%	
η (i, π, 5)	41.3%	38.2%	30.7%	40.6%	35.1%	η (i, π, 5)	30.7%	41.3%	40.7%	38.2%	35.1%
π_3 = {PC1, LC, DC1, PC2, DC2}						π_4 = {PC2, LC, DC2, PC1, DC1}					
Player i	PC1	LC	DC1	PC2	DC2	Player i	PC2	LC	DC2	PC1	DC1
η (i, π, 1)	11.7%					η (i, π, 1)	8.0%				
η (i, π, 2)	30.9%	27.7%				η (i, π, 2)	27.3%	26.3%			
η (i, π, 3)	36.7%	32.2%	32.4%			η (i, π, 3)	28.7%	34.0%	26.7%		
η (i, π, 4)	39.8%	38.0%	35.1%	29.3%		η (i, π, 4)	29.5%	38.1%	28.4%	40.3%	
η (i, π, 5)	40.6%	41.3%	38.2%	36.0%	30.7%	η (i, π, 5)	35.1%	41.3%	30.7%	40.6%	38.2%
π_5 = {LC, PC1, DC1, PC2, DC2}											
Player i	LC	PC1	DC1	PC2	DC2						
η (i, π, 1)	10.8%										
η (i, π, 2)	27.7%	30.9%									
η (i, π, 3)	32.2%	36.7%	32.4%								
η (i, π, 4)	38.0%	39.8%	35.1%	29.3%							
η (i, π, 5)	41.3%	40.6%	38.2%	35.1%	30.7%						

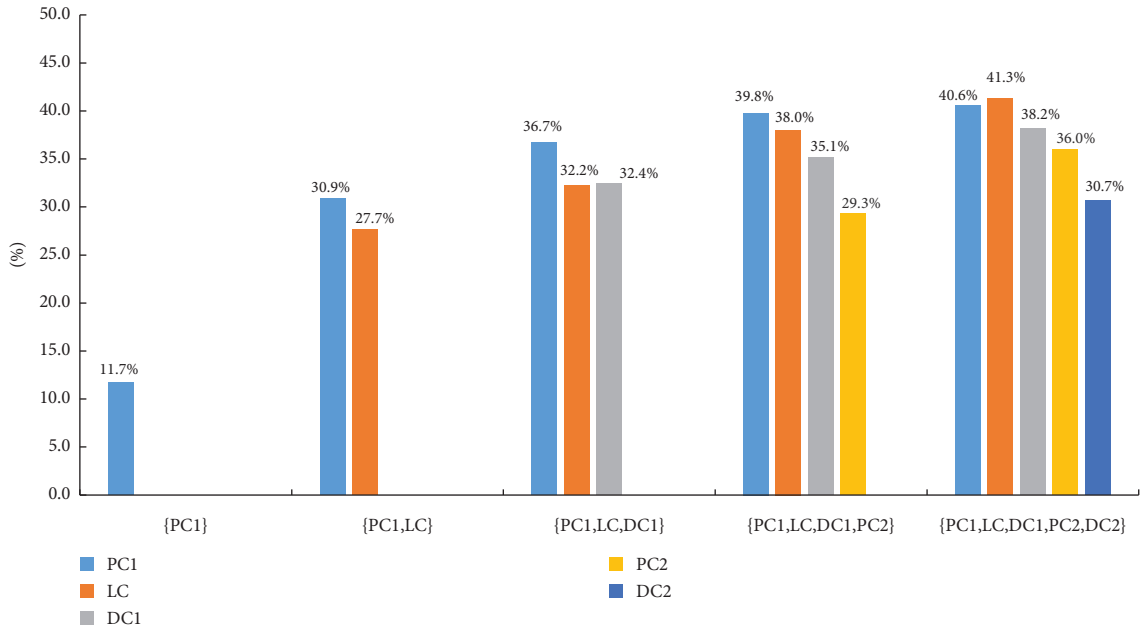


Figure 11: Cost reduction percentages of the grand alliance.

in Section 4.2.3, the comparative results of the proximity between profit allocation schemes generated by different methods and the center values are presented, as shown in Table 12 and demonstrated in Figure 12.

In Table 12 and Figure 12, the value of the core center is (2702, 2316, 2029, 2312, 2434). Comparative results between the core center value and schemes using different methods suggest that the profit allocation scheme via the MCRS method is the farthest from the core center. The profit of DC1 in MCRS is less than that in the other profit allocation schemes, and thus, DC1 is more likely to refuse to join the alliance to further improve its profit. Similarly, both the profit of LC in GQP and the profit of PC1 in EPM can be improved. The distances between profit allocation schemes using the EPM, MCRS, and GQP methods are 429, 517, and 308, which are not the closest. Therefore, the stability and development of the alliance will be threatened. It is obvious that the closest scheme is generated by the Shapley value method. Therefore, the logistics facilities will participate in the grand alliance and construct a long-term partnership

Table 12: Comparative results between different schemes and center value.

Method	Profit allocation schemes	Center	Distance
Shapley	(2684, 2069, 2165, 2326, 2549)	(2702, 2316, 2029, 2312, 2434)	305
EPM	(2913, 2132, 2277, 2118, 2353)		429
MCRS	(3113, 2032, 2077, 2218, 2353)		517
GQP	(2706, 2059, 2077, 2428, 2547)		308

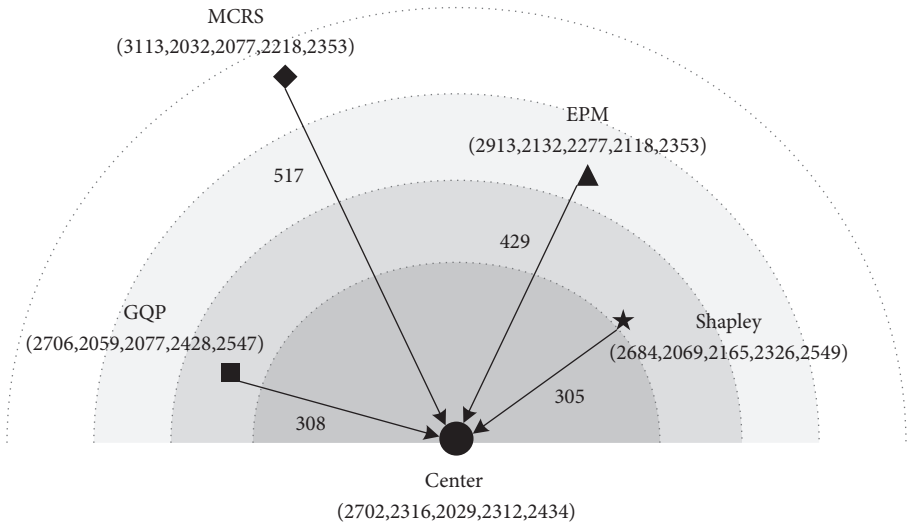


Figure 12: Comparison of distances between different methods and center value.

when the profit allocation scheme is set as (2,684, 2,069, 2,165, 2,326, 2,549).

6. Conclusions

The collaboration between logistics facilities is an effective management strategy to promote resource sharing, provide considerable profits for logistics enterprises, and improve the performance of urban pickup and delivery networks. To improve efficiency and reduce the total operating cost, the sharing of customer information and transportation resources is considered. First, a bi-objective integer programming model is developed to minimize the total operating cost and the number of vehicles in the 2E-CMDPDTW. Second, a hybrid method including the customer clustering algorithm, the greedy algorithm, and the Im-NSGA-II is developed to solve the mathematical model. Then, the Shapley value method is used for profit allocation in the alliance established in the optimized network, and the optimal sequence of joining the alliance can be generated using the SMP. Finally, the alliance stability is discussed by comparing different profit allocation methods {Fernie, 2007 #3217}.

An empirical case in Chongqing, China, is used to test the effectiveness of the proposed model and solution method for solving the 2E-CMDPDTW. Compared to the non-collaborative network, the operating cost and the number of vehicles are reduced by $11,793 and 3, respectively. The solution speed and quality of the proposed Im-NSGA-II are better than the existing multiobjective optimization algorithms, such as NSGA-II and MOGA. A reasonable profit allocation scheme for the alliance is provided by the Shapley value method. Taking into account the profit-seeking nature of the logistics facilities, the alliance organizer is more likely to select the profit allocation scheme provided by the Shapley value method rather than the GQP, MCRS, and EPM methods. The profits of $2326, $2684, $2069, $2549, and $2165 are allocated to the five logistics facilities in the optimal scheme {PC1, LC, DC1, PC2, DC2} generated by the SMP method. Establishing a grand alliance consisting of all logistics facilities in the network maximizes cost savings and transportation resource utilization. Besides, the reasonable profit distribution schemes are conducive to the stable development of the alliance, and they are determined by the appropriate distribution methods.

From the operational perspective, resource sharing has the potential to share operating risks and contributes to cost savings for logistics enterprises. Another key reason for promoting resource sharing is to achieve on-time deliveries via customer information sharing. As for urban transportation governance, resource sharing plays a significant role in optimizing logistics delivery and pickup networks and mitigating urban traffic congestions. Therefore, resource sharing among logistics facilities provides a promising way for logistics enterprises to reduce costs and improve efficiency and for the government to improve the regional logistics service capacity and build more efficient urban transportation networks.

Despite an effective approach to solving 2E-CMDPDTW is proposed in this study, the following concerns need to be addressed in future research. (1) In terms of logistics network design, customers with dynamic demands or multiple

service time windows can be modeled in a two-echelon collaborative multideport logistics network. (2) For the 2E-MDLPDN optimization, more noneconomic factors (e.g., environmental factors) can be considered. (3) In the optimization solution approach, exact algorithms can be combined with the heuristic algorithm to obtain high-quality optimal solutions and a faster convergence speed. (4) In the aspect of logistics collaboration, the contributions of different alliance members to the alliance stability can be further explored on the basis of different profit allocation methods.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Yong Wang and Xiangyang Guan conceived and designed the conceptualization and methodology; Siyu Luo and Yong Wang performed the software and validation; Yong Wang and Siyu Luo completed the formal analysis, investigation, and data curation; Yong Wang, Jinjun Tang, and Maozeng Xu contributed the resources and visualization; Siyu Luo, Yong Wang, and Xiangyang Guan completed writing the original draft and review and did editing. Yong Wang, Jinjun Tang, and Maozeng Xu implemented the supervision. Yong Wang and Jinjun Tang carried out the project administration and funding acquisition. All the authors approved the final manuscript.

Acknowledgments

This research was supported by National Natural Science Foundation of China (Project nos. 71871035 and 41977337), Key Science and Technology Research Project of Chongqing Municipal Education Commission (KJZD-K202000702), Key Project of Human Social Science of Chongqing Municipal Education Commission (no. 20SKGH079), Team Building Project for Graduate Tutors in Chongqing (no. JDDSTD2019008), and Research and Innovation Program for Graduate Students in Chongqing (CYB21219).

References

[1] G. D. Konstantakopoulos, S. P. Gayialis, E. P. Kechagias, G. A. Papadopoulos, and I. P. Tatsiopoulos, "An algorithmic approach for sustainable and collaborative logistics: a case study in Greece," *International Journal of Information Management Data Insights*, vol. 1, no. 1, Article ID 100010, 2021.

[2] H. Farvaresh and S. Shahmansouri, "A coalition structure algorithm for large-scale collaborative pickup and delivery problem," *Computers & Industrial Engineering*, vol. 149, Article ID 106737, 2020.

[3] C. B. I. R. Institute, "Market analysis of China's logistics industry in 2019 and prospects for 2020," 2020, https://www.askci.com/news/chanye/20200509/1542541160124.shtml.

[4] M. Soysal, J. M. Bloemhof-Ruwaard, and T. Bektaş, "The time-dependent two-echelon capacitated vehicle routing problem with environmental considerations," *International Journal of Production Economics*, vol. 164, pp. 366–378, 2015.

[5] S. K. Curtis and O. Mont, "Sharing economy business models for sustainability," *Journal of Cleaner Production*, vol. 266, Article ID 121519, 2020.

[6] Z. Tao, Q. Nie, and W. Zhang, "Research on travel behavior with car sharing under smart city conditions," *Journal of Advanced Transportation*, vol. 2021, Article ID 8879908, 13 pages, 2021.

[7] Y. Wang, X. Ma, M. Liu et al., "Cooperation and profit allocation in two-echelon logistics joint distribution network optimization," *Applied Soft Computing*, vol. 56, pp. 143–157, 2017.

[8] Y. Shi, N. Lin, Q. Han, T. Zhang, and W. Shen, "A method for transportation planning and profit sharing in collaborative multi-carrier vehicle routing," *Mathematics*, vol. 8, no. 10, p. 1788, 2020.

[9] Y. Wang, S. Peng, X. Zhou, M. Mahmoudi, and L. Zhen, "Green logistics location-routing problem with eco-packages," *Transportation Research Part E: Logistics and Transportation Review*, vol. 143, Article ID 102118, 2020.

[10] H. Bae and I. Moon, "Multi-depot vehicle routing problem with time windows considering delivery and installation vehicles," *Applied Mathematical Modelling*, vol. 40, no. 13-14, pp. 6536–6549, 2016.

[11] B. Rabbouch, R. Mraihi, and F. Saâdaoui, "A recent brief survey for the multi depot heterogenous vehicle routing problem with time windows," in *Proceedings of the International Conference on Hybrid Intelligent Systems*, pp. 147–157, Springer, Delhi, India, December 2017.

[12] J. R. Montoya-Torres, J. López Franco, S. Nieto Isaza, H. Felizzola Jiménez, and N. Herazo-Padilla, "A literature review on the vehicle routing problem with multiple depots," *Computers & Industrial Engineering*, vol. 79, pp. 115–129, 2015.

[13] P. Surekha and S. Sumathi, "Solution to multi-depot vehicle routing problem using genetic algorithms," *World Applied Programming*, vol. 1, no. 3, pp. 118–131, 2011.

[14] S. Karakatič and V. Podgorelec, "A survey of genetic algorithms for solving multi depot vehicle routing problem," *Applied Soft Computing*, vol. 27, pp. 519–532, 2015.

[15] J.-F. Cordeau, G. Laporte, and A. Mercier, "A unified tabu search heuristic for vehicle routing problems with time windows," *Journal of the Operational Research Society*, vol. 52, no. 8, pp. 928–936, 2001.

[16] M. Schneider, "The vehicle-routing problem with time windows and driver-specific times," *European Journal of Operational Research*, vol. 250, no. 1, pp. 101–119, 2016.

[17] X. Xiang, J. Qiu, J. Xiao, and X. Zhang, "Demand coverage diversity based ant colony optimization for dynamic vehicle routing problems," *Engineering Applications of Artificial Intelligence*, vol. 91, Article ID 103582, 2020.

[18] P. Stodola, "Hybrid ant colony optimization algorithm applied to the multi-depot vehicle routing problem," *Natural Computing*, vol. 19, no. 2, pp. 463–475, 2020.

[19] A. M. Nogareda, J. Del Ser, E. Osaba, and D. Camacho, "On the design of hybrid bio-inspired meta-heuristics for complex multiattribute vehicle routing problems," *Expert Systems*, vol. 37, no. 6, Article ID e12528, 2020.

[20] J. R. Montoya-Torres, A. Muñoz-Villamizar, and C. A. Vega-Mejía, "On the impact of collaborative strategies for goods delivery in city logistics," *Production Planning & Control*, vol. 27, no. 6, pp. 443–455, 2016.

[21] G. Xu, M. Xu, Y. Wang, Y. Liu, and Q. Lv, "Collaborative multidepot petrol station replenishment problem with multicompartments and time window assignment," *Journal of Advanced Transportation*, vol. 2020, Article ID 8843397, 22 pages, 2020.

[22] A. Muñoz-Villamizar, C. L. Quintero-Araújo, J. R. Montoya-Torres, and J. Faulin, "Short- and mid-term evaluation of the use of electric vehicles in urban freight transport collaborative networks: a case study," *International Journal of Logistics Research and Applications*, vol. 22, no. 3, pp. 229–252, 2019.

[23] Y. Wang, S. Zhang, K. Assogba, J. Fan, M. Xu, and Y. Wang, "Economic and environmental evaluations in the two-echelon collaborative multiple centers vehicle routing optimization," *Journal of Cleaner Production*, vol. 197, pp. 443–461, 2018.

[24] S. Mancini, M. Gansterer, and R. F. Hartl, "The collaborative consistent vehicle routing problem with workload balance," *European Journal of Operational Research*, vol. 293, no. 3, pp. 955–965, 2021.

[25] Y. Wang, Q. Li, X. Guan, J. Fan, M. Xu, and H. Wang, "Collaborative multi-depot pickup and delivery vehicle routing problem with split loads and time windows," *Knowledge-Based Systems*, vol. 131, Article ID 107412, 2021.

[26] Y. Li, H. Soleimani, and M. Zohal, "An improved ant colony optimization algorithm for the multi-depot green vehicle routing problem with multiple objectives," *Journal of Cleaner Production*, vol. 227, pp. 1161–1172, 2019.

[27] C. L. Quintero-Araujo, A. Gruler, A. A. Juan, and J. Faulin, "Using horizontal cooperation concepts in integrated routing and facility-location decisions," *International Transactions in Operational Research*, vol. 26, no. 2, pp. 551–576, 2019.

[28] B. Vahdani, F. Mansour, M. Soltani, and D. Veysmoradi, "Bi-objective optimization for integrating quay crane and internal truck assignment with challenges of trucks sharing," *Knowledge-Based Systems*, vol. 163, pp. 675–692, 2019.

[29] C. K. Y. Lin, "A vehicle routing problem with pickup and delivery time windows, and coordination of transportable resources," *Computers & Operations Research*, vol. 38, no. 11, pp. 1596–1609, 2011.

[30] A. Bhasker, S. P. Sarmah, and T. Kim, "Collaborative last-mile delivery and pick-up in city logistics," *International Journal of Logistics Systems and Management*, vol. 34, no. 4, pp. 533–553, 2019.

[31] A. Subramanian, L. M. A. Drummond, C. Bentes, L. S. Ochi, and R. Farias, "A parallel heuristic for the vehicle routing problem with simultaneous pickup and delivery," *Computers & Operations Research*, vol. 37, no. 11, pp. 1899–1911, 2010.

[32] F. P. Goksal, I. Karaoglan, and F. Altiparmak, "A hybrid discrete particle swarm optimization for vehicle routing problem with simultaneous pickup and delivery," *Computers & Industrial Engineering*, vol. 65, no. 1, pp. 39–53, 2013.

[33] H.-F. Wang and Y.-Y. Chen, "A genetic algorithm for the simultaneous delivery and pickup problems with time window," *Computers & Industrial Engineering*, vol. 62, no. 1, pp. 84–95, 2012.

[34] M. Mahmoudi and X. Zhou, "Finding optimal solutions for vehicle routing problem with pickup and delivery services with time windows: a dynamic programming approach based on state-space-time network representations," *Transportation Research Part B: Methodological*, vol. 89, pp. 19–42, 2016.

[35] K. Govindan, A. Jafarian, and V. Nourbakhsh, "Designing a sustainable supply chain network integrated with vehicle routing: a comparison of hybrid swarm intelligence meta-heuristics," *Computers & Operations Research*, vol. 110, pp. 220–235, 2019.

[36] L. Zhou, R. Baldacci, D. Vigo, and X. Wang, "A multi-depot two-echelon vehicle routing problem with delivery options arising in the last mile distribution," *European Journal of Operational Research*, vol. 265, no. 2, pp. 765–778, 2018.

[37] H. Li, H. Wang, J. Chen, and M. Bai, "Two-echelon vehicle routing problem with time windows and mobile satellites," *Transportation Research Part B: Methodological*, vol. 138, pp. 179–201, 2020.

[38] I. H. Dridi, E. Ben Alaïa, P. Borne, and H. Bouchriha, "Optimisation of the multi-depots pick-up and delivery problems with time windows and multi-vehicles using PSO algorithm," *International Journal of Production Research*, vol. 58, no. 14, pp. 4201–4214, 2020.

[39] A. Bettinelli, A. Ceselli, and G. Righini, "A branch-and-price algorithm for the multi-depot heterogeneous-fleet pickup and delivery problem with soft time windows," *Mathematical Programming Computation*, vol. 6, no. 2, pp. 171–197, 2014.

[40] V. Kachitvichyanukul, P. Sombuntham, and S. Kunnapapdeelert, "Two solution representations for solving multi-depot vehicle routing problem with multiple pickup and delivery requests via PSO," *Computers & Industrial Engineering*, vol. 89, pp. 125–136, 2015.

[41] A. Soriano, M. Gansterer, and R. F. Hartl, "The two-region multi-depot pickup and delivery problem," *OR Spectrum*, vol. 40, no. 4, pp. 1077–1108, 2018.

[42] L. D. C. Martins, P. Hirsch, and A. A. Juan, "Agile optimization of a two-echelon vehicle routing problem with pickup and delivery," *International Transactions in Operational Research*, vol. 28, no. 1, pp. 201–221, 2021.

[43] O. Belgin, I. Karaoglan, and F. Altiparmak, "Two-echelon vehicle routing problem with simultaneous pickup and delivery: mathematical model and heuristic approach," *Computers & Industrial Engineering*, vol. 115, pp. 1–16, 2018.

[44] E. Pérez-Bernabeu, A. A. Juan, J. Faulin, and B. B. Barrios, "Horizontal cooperation in road transportation: a case illustrating savings in distances and greenhouse gas emissions," *International Transactions in Operational Research*, vol. 22, no. 3, pp. 585–606, 2015.

[45] Y. Wang, X. Ma, Z. Li, Y. Liu, M. Xu, and Y. Wang, "Profit distribution in collaborative multiple centers vehicle routing problem," *Journal of Cleaner Production*, vol. 144, pp. 203–219, 2017.

[46] Y. Wang, S. Peng, C. Xu et al., "Two-echelon logistics delivery and pickup network optimization based on integrated cooperation and transportation fleet sharing," *Expert Systems with Applications*, vol. 113, pp. 44–65, 2018.

[47] M. Adelzadeh, V. M. Asl, and M. Koosha, "A mathematical model and a solving procedure for multi-depot vehicle routing problem with fuzzy time window and heterogeneous vehicle," *International Journal of Advanced Manufacturing Technology*, vol. 75, no. 5–8, pp. 793–802, 2014.

[48] T. G. Crainic, S. Mancini, G. Perboli, and R. Tadei, *Clustering-based Heuristics for the Two-Echelon Vehicle Routing Problem*, CIRRELT, Montréal, Canada, 2008.

[49] W. Meihua, T. Xuhong, C. Shan, and W. Shumin, "Hybrid ant colony optimization algorithm for two echelon vehicle routing problem," *Procedia Engineering*, vol. 15, pp. 3361–3365, 2011.

[50] L. Liu and W. Liao, "Optimization and profit distribution in a two-echelon collaborative waste collection routing problem from economic and environmental perspective," *Waste Management*, vol. 120, pp. 400–414, 2021.

[51] A. K. Paul and P. C. Shill, "New automatic fuzzy relational clustering algorithms using multi-objective NSGA-II," *Information Sciences*, vol. 448-449, pp. 112–133, 2018.

[52] V. L. Vachhani, V. K. Dabhi, and H. B. Prajapati, "Improving NSGA-II for solving multi objective function optimization problems," in *Proceedings of the 2016 International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1–6, IEEE, Coimbatore, India, January 2016.

[53] S. Bandyopadhyay and R. Bhattacharya, "Solving a tri-objective supply chain problem with modified NSGA-II algorithm," *Journal of Manufacturing Systems*, vol. 33, no. 1, pp. 41–50, 2014.

[54] S. Li, N. Wang, T. Jia, Z. He, and H. Liang, "Multiobjective optimization for multiperiod reverse logistics network design," *IEEE Transactions on Engineering Management*, vol. 63, no. 2, pp. 223–236, 2016.

[55] J. Wen and Y. Li, "Profit distribution model of green building supply chain with fairness preferences and cap-and-trade policy," *IOP Conference Series: Earth and Environmental Science*, IOP Publishing, vol. 237, Bristol, UK, Article ID 052043, 2019.

[56] J.-C. Liu, J.-B. Sheu, D.-F. Li, and Y.-W. Dai, "Collaborative profit allocation schemes for logistics enterprise coalitions with incomplete information," *Omega*, vol. 101, Article ID 102237, 2021.

[57] Y. Wang, S. Zhang, X. Guan et al., "Collaborative multi-depot logistics network design with time window assignment," *Expert Systems with Applications*, vol. 140, Article ID 112910, 2020.

[58] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[59] T. Wei, W. Fan, and H. Xu, "Greedy non-dominated sorting in genetic algorithm-II for vehicle routing problem in distribution," *Chinese Journal of Mechanical Engineering*, vol. 21, no. 6, pp. 18–24, 2008.

[60] B. Aerts, T. Cornelissens, and K. Sörensen, "The joint order batching and picker routing problem: modelled and solved as a clustered vehicle routing problem," *Computers & Operations Research*, vol. 129, Article ID 105168, 2021.

[61] V. C. Hemmelmayr, J.-F. Cordeau, and T. G. Crainic, "An adaptive large neighborhood search heuristic for two-echelon vehicle routing problems arising in city logistics," *Computers & Operations Research*, vol. 39, no. 12, pp. 3215–3228, 2012.

[62] S. R. A. Haddadene, N. Labadie, and C. Prodhon, "NSGAII enhanced with a local search for the vehicle routing problem with time windows and synchronization constraints," *IFAC-PapersOnLine*, vol. 49, no. 12, pp. 1198–1203, 2016.

[63] M. Xu, Z. Mei, S. Luo, and Y. Tan, "Optimization algorithms for construction site layout planning: a systematic literature review," *Engineering, Construction and Architectural Management*, vol. 27, no. 8, pp. 1913–1938, 2020.

[64] A. Aguilar-Rivera, "A GPU fully vectorized approach to accelerate performance of NSGA-2 based on stochastic non-domination sorting and grid-crowding," *Applied Soft Computing*, vol. 88, Article ID 106047, 2020.

[65] W. Peng, J. Mu, L. Chen, and J. Lin, "A novel non-dominated sorting genetic algorithm for solving the triple objective project scheduling problem," *Memetic Computing*, vol. 13, no. 2, pp. 271–284, 2021.

[66] J. C. Ferreira and M. T. A. Steiner, "A Bi-objective green vehicle routing problem: a new hybrid optimization algorithm applied to a newspaper distribution," *Journal of Geographic Information System*, vol. 13, no. 4, pp. 410–433, 2021.

[67] A. Haghrah, M. Nekoui, M. Nazari-Heris, and B. Mohammadi-ivatloo, "An improved real-coded genetic algorithm with random walk based mutation for solving combined heat and power economic dispatch," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 8, pp. 8561–8584, 2021.

[68] A. Kimms and I. Kozeletskyi, "Core-based cost allocation in the cooperative traveling salesman problem," *European Journal of Operational Research*, vol. 248, no. 3, pp. 910–916, 2016.

[69] Y. Wang, J. Zhang, K. Assogba, Y. Liu, M. Xu, and Y. Wang, "Collaboration and transportation resource sharing in multiple centers vehicle routing optimization with delivery and pickup," *Knowledge-Based Systems*, vol. 160, pp. 296–310, 2018.

[70] B. Niu, F. Xie, L. Chen, and X. Xu, "Join logistics sharing alliance or not? Incentive analysis of competing E-commerce firms with promised-delivery-time," *International Journal of Production Economics*, vol. 224, Article ID 107553, 2020.

[71] M. Rabani, D. Abdolhamidi, M. Mokhtarzadeh, and S. Fatemi-Anaraki, "Solving a bi-objective medicine distribution problem considering delivery to waste center using a hybrid clustering, mathematical modeling and NSGA-II approach," *Journal of Industrial and Systems Engineering*, vol. 13, no. 2, pp. 245–263, 2021.

[72] S. F. Ghannadpour and F. Zandiyeh, "An adapted multi-objective genetic algorithm for solving the cash in transit vehicle routing problem with vulnerability estimation for risk quantification," *Engineering Applications of Artificial Intelligence*, vol. 96, Article ID 103964, 2020.

[73] M. M. Solomon, "Algorithms for the vehicle routing and scheduling problems with time window constraints," *Operations Research*, vol. 35, no. 2, pp. 254–265, 1987.

[74] M. Frisk, M. Göthe-Lundgren, K. Jörnsten, and M. Rönnqvist, "Cost allocation in collaborative forest transportation," *European Journal of Operational Research*, vol. 205, no. 2, pp. 448–458, 2010.

[75] Y. Wang, J. Zhang, X. Guan, M. Xu, Z. Wang, and H. Wang, "Collaborative multiple centers fresh logistics distribution network optimization with resource sharing and temperature control constraints," *Expert Systems with Applications*, vol. 165, Article ID 113838, 2021.