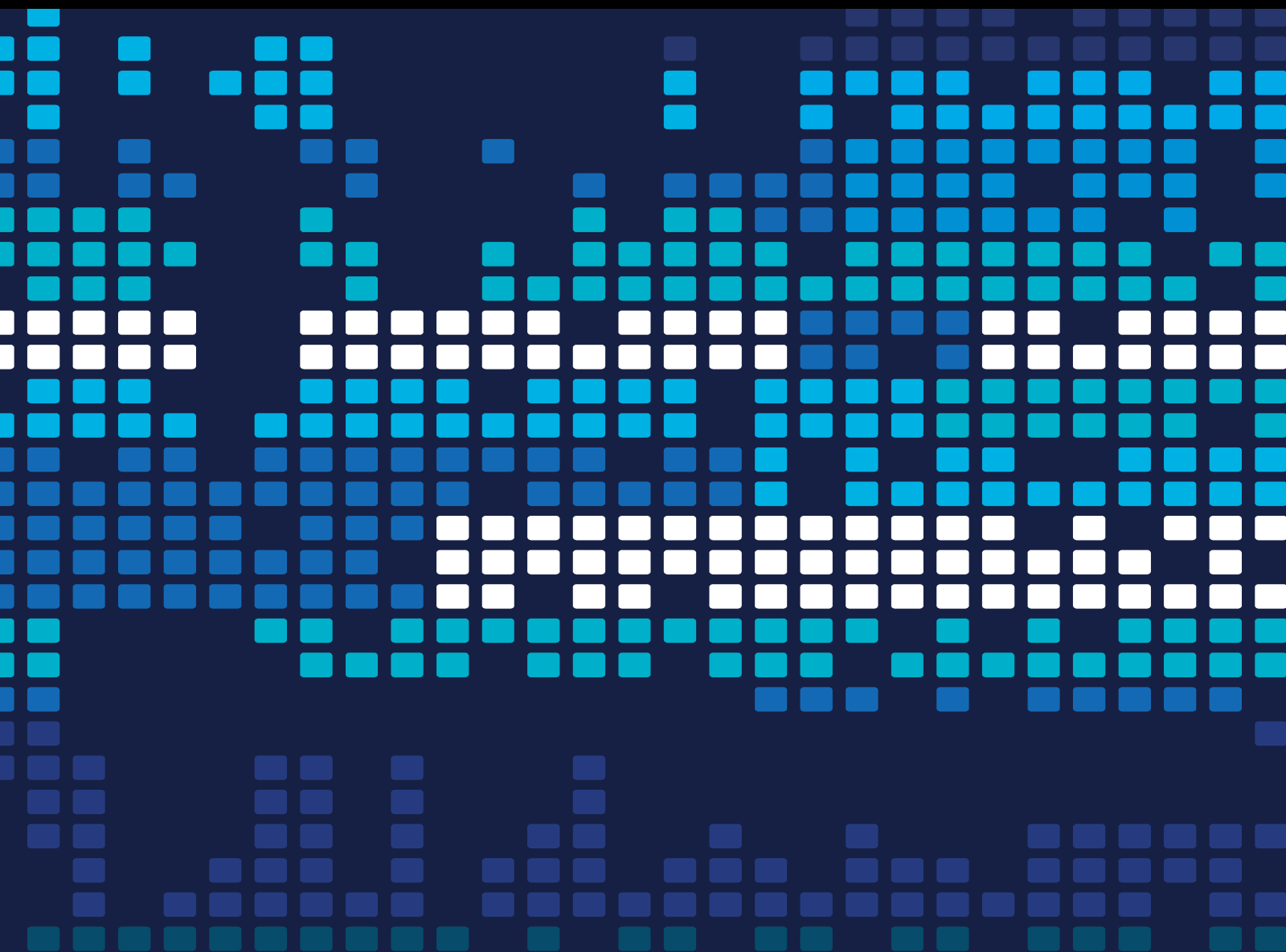


Optimization Models and Algorithms for Services and Operations Management 2020

Lead Guest Editor: Lu Zhen

Guest Editors: Xiaobo Qu and Tingsong Wang





**Optimization Models and Algorithms for
Services and Operations Management 2020**

Scientific Programming

**Optimization Models and Algorithms
for Services and Operations
Management 2020**

Lead Guest Editor: Lu Zhen


Guest Editors: Xiaobo Qu and Tingsong Wang



Copyright © 2021 Hindawi Limited. All rights reserved.

This is a special issue published in "Scientific Programming." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Chief Editor

Emiliano Tramontana , Italy

Academic Editors

Marco Aldinucci , Italy
Daniela Briola, Italy
Debo Cheng , Australia
Ferruccio Damiani , Italy
Sergio Di Martino , Italy
Sheng Du , China
Basilio B. Fraguela , Spain
Jianping Gou , China
Jiwei Huang , China
Sadiq Hussain , India
Shujuan Jiang , China
Oscar Karnalim, Indonesia
José E. Labra, Spain
Maurizio Leotta , Italy
Zhihan Liu , China
Piotr Luszczek, USA
Tomàs Margalef , Spain
Cristian Mateos , Argentina
Zahid Mehmood , Pakistan
Roberto Natella , Italy
Diego Oliva, Mexico
Antonio J. Peña , Spain
Danilo Pianini , Italy
Jiangbo Qian , China
David Ruano-Ordás , Spain
Željko Stević , Bosnia and Herzegovina
Kangkang Sun , China
Zhiri Tang , Hong Kong
Autilia Vitiello , Italy
Pengwei Wang , China
Jan Weglarz, Poland
Hong Wenxing , China
Dongpo Xu , China
Tolga Zaman, Turkey

Contents

A Dual-Channel Sale System in Financially Constrained Supply Chain

Xin Li , Zhang Tao , and Nana Feng 

Research Article (10 pages), Article ID 8872728, Volume 2021 (2021)

Gaussian Kernel Fuzzy C-Means Algorithm for Service Resource Allocation

Wei Jiang, Xi Fang , and Jianmei Ding


Research Article (6 pages), Article ID 8889480, Volume 2020 (2020)

Operation Efficiency Evaluation of the China-Europe Freight Train Based on Grey Cross-Efficiency DEA

Jun Zhang, Jianpeng Chang, Ping Lin, Minzi Song , and Yanqiu Gong


Research Article (12 pages), Article ID 8843733, Volume 2020 (2020)

Fuzzy Principal Component Analysis Model on Evaluating Innovation Service Capability

Hui Wu and Xiao-min Gu 


Research Article (9 pages), Article ID 8834901, Volume 2020 (2020)

Optimization of Continuous Berth Scheduling by Taking into Account Double-Line Ship Mooring

Cheng Luo, Hongying Fei , Dana Sailike, Tingyi Xu, and Fuzhi Huang


Research Article (11 pages), Article ID 8863994, Volume 2020 (2020)

A Comparative Performance Investigation of Swarm Optimisers on the Design of Hydrostatic Thrust Bearing

Ismail Sahin 


Research Article (14 pages), Article ID 8856770, Volume 2020 (2020)

Model and Hybrid Algorithm of Collaborative Distribution System with Multiple Drones and a Truck

Min Lin, Jun-Yan Lyu, Jia-Jing Gao, and Ling-Yu Li 

Research Article (16 pages), Article ID 8887057, Volume 2020 (2020)

Multidepot Heterogeneous Vehicle Routing Problem for a Variety of Hazardous Materials with Risk Analysis

Bochen Wang, Qiyuan Qian , Zheyi Tan, Peng Zhang, Aizhi Wu, and Yi Zhou


Research Article (11 pages), Article ID 8839628, Volume 2020 (2020)

Model for Design of Portfolio Venture Investment Contract When Taking Moral Hazards into Account

L. Yin, Y. Liu , and Z. Wang


Research Article (6 pages), Article ID 8821371, Volume 2020 (2020)

The Model of the Relationship between Urban Rail Transit and Residential Location

Lin Zhang 


Research Article (15 pages), Article ID 8851637, Volume 2020 (2020)

Analysis of the Impact of Interest Rate Liberalization on Financial Services Management in Chinese Commercial Banks

Tian Meng, Mengnan Sun, Yixuan Zhao, and Bo Zhu 


Research Article (11 pages), Article ID 8860076, Volume 2020 (2020)

Evolutionary Game Models on Multiagent Collaborative Mechanism in Responsible Innovation

Kun Yang, Wan Wang , and Bin Hu



Research Article (11 pages), Article ID 8875099, Volume 2020 (2020)

Integrated Pricing and Distribution Planning for Community Group Purchase of Fresh Agricultural Products

Wenbing Shui and Mengxia Li 


Research Article (15 pages), Article ID 8839398, Volume 2020 (2020)

Study on Coordination and Optimization of Contract Farming Supply Chain Based on Uncertain Conditions

Xinquan Liu , Xiaojing Shen , and Ming You

Research Article (9 pages), Article ID 8858812, Volume 2020 (2020)

Heuristic Algorithms for MapReduce Scheduling Problem with Open-Map Task and Series-Reduce Tasks

Feifeng Zheng, Zhaojie Wang, Yinfeng Xu, and Ming Liu 

Research Article (10 pages), Article ID 8810215, Volume 2020 (2020)

A DE-LS Metaheuristic Algorithm for Hybrid Flow-Shop Scheduling Problem considering Multiple Requirements of Customers

Yingjia Sun  and Xin Qi


Research Article (14 pages), Article ID 8811391, Volume 2020 (2020)

Cost-Benefit Models on Integrating Information Technology Services in Automotive Production Management

Bin Hu, Jianlin Lv , and Kun Yang

Research Article (9 pages), Article ID 8877780, Volume 2020 (2020)

Intelligent Differential Evolution Scheme for Network Resources in IoT

Huu Dang Quoc , Loc Nguyen The, Cuong Nguyen Doan, Toan Phan Thanh, and Neal N. Xiong

Research Article (12 pages), Article ID 8860384, Volume 2020 (2020)

Addressing the Bike Repositioning Problem in Bike Sharing System: A Two-Stage Stochastic Programming Model

Qiong Tang, Zhuo Fu , Dezhi Zhang, Hao Guo , and Minyi Li

Research Article (12 pages), Article ID 8868892, Volume 2020 (2020)

Research Article

A Dual-Channel Sale System in Financially Constrained Supply Chain

Xin Li ¹, Zhang Tao ², and Nana Feng ²

¹School of Business, Macau University of Science and Technology, Macau SAR, China

²Nanning Normal University, Nanning, Guangxi, China

Correspondence should be addressed to Nana Feng; violet3676@163.com

Received 27 July 2020; Accepted 17 May 2021; Published 2 June 2021

Academic Editor: Tingsong Wang

Copyright © 2021 Xin Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to realize the sustainable development of the communication industry, many banks have developed some new projects to provide loan to the mobile phone brand companies (MPBCs). This paper studies a perfect information game among three parties: telecom operator (TO), MPBC, and bank. In the first stage, the bank decides an interest rate and shows it to the TO and the competitive MPBC. In the second stage, the TO and MPBC are engaged in Cournot games: a simultaneous subgame, a TO-as-leader sequential subgame, and an MPBC-as-leader sequential subgame. The TO's and the MPBC's decisions and the production/sale quantities are investigated. The impacts of the interest rate and the substitute factor on the TO's and the MPBC's optimal decisions are analyzed. When the substitute factor is high, at a low interest rate, the total sales in the simultaneous subgame is higher than those in the other two subgames; at a high interest rate, the total sales in the MPBC-as-leader subgame is higher than those in the other two subgames. However, when the substitute factor is low, at a low enough interest rate, the total sales in the simultaneous subgame is higher than those in the other two subgames; at a high enough interest rate, the total sales in the MPBC-as-leader subgame is higher than those in the other two subgames; at a moderate interest rate, the total sales in the TO-as-leader subgame is higher than those in the other two subgames. Besides, the optimal interest rate of the bank is investigated and the impact of the substitute factor on the optimal interest rate is analyzed. The bank sets a higher interest rate in the MPBC-as-leader subgame than those in the other two subgames. Besides, when the substitute factor is low, the bank sets a lower interest rate in the TO-as-leader subgame than that in the simultaneous subgame; however, when the substitute factor is high, the bank sets a higher interest rate in the TO-as-leader subgame than that in the simultaneous subgame.

1. Introduction

Recently, in order to increase market share and realize the sustainable development of the TOs, more and more TOs start the customized-phone business and launch a wide range of customized packages, in which telecom operators purchase cell phones from the MPBCs, and the phones can only fit in their own network modes, such as GSM (Global System for Mobile Communications) and CDMA (Code Division Multiple Access). In addition, the customized phones are sold with telecom service provided by the TOs. Usually, the clients sign contracts with the TOs, under which the TOs sell the phones and service to the clients at a discount with a long term, usually from 1 to 3 years. Then, the customized phone is also called contract phone, otherwise

named noncontract phone. In 2016, the mobile phone users in China reached 1,316 billion and the penetration rate was 95.5 mobile phones per 100 persons. (the source is from the web of Ministry of Industry and Information Technology of the People's Republic of China (<http://www.miiit.gov.cn/>)). The mobile phone has become an essential tool in our daily life.

The demand for the better functions and service of mobile phone leads to the cooperation between the TOs and the MPBCs. By selling their cell phones with the service through the TOs, the MPBCs widen their sale channels and motive their existing sale systems, so as to increase their market share and realize the sustainable development of the MPBCs. The TOs obtain market information and feedback timely, which accelerates the phone's upgrading speed.

Recently, because of the increasing usage of the intelligent mobiles, the demanding for the network and connection quality is higher, which makes the TOs earn larger profit and realize the sustainable development of TOs.

There is not only cooperations but also competitions between the TOs and the MPBCs. Specially, on the one hand, the TOs purchase the mobile phones from the MPBCs; on the other hand, both of the TOs and the MPBCs sell mobile phones to customers. The customers' decisions are affected by two main factors: the price of the phone and the service fee of the TOs. Buying the contracted phone, customers can obtain a discount of the mobile phone price, but must pay the additional fee for the service set provided by the TOs. Usually, customers need to pay a minimum consumption for the service set, even if they do not use all of the service in the service set. Many TOs propagate the slogan "take the mobile phone for free." This kind of advertisement really attracts a lot of customers to buy the contract phone.

The TOs realize that the most of the profit is occupied by the MPBCs and then begin to change their operation strategies, even beginning to design and to manufacture their own brand mobile phone actively. The MPBCs also do not consider the TOs as their main sales channel. The MPBCs change their sale strategy from selling cheap phones to accessing high-level market; meanwhile, they start to launch new sale channels like e-commerce. Some of the MPBCs expand the foreign market actively to enhance their market share, such as Huawei. At the same time, they start to focus on the customers' requirements for the mobile phone functions, rather than on reducing the manufacturing cost.

The expansion of either TOs or MPBCs require a large amount of fund. In China, the short-term fund relies on loan from bank and the long-term fund depends mainly on equity financing. However, in Britain and America, companies' external financing relies on issuing bonds instead of stock, such that they are cautious to use bank loans. They choose the bank loans only when they need middle- or long-term fund. The MPBCs need to manufacture all of the phones, including the orders from the TOs. Consequently, the MPBCs need to borrow money from the bank. In order to realize the sustainable development of the communication industry, and many banks have developed some projects to provide loan to the MPBCs, such as "Helidai" is provided by Shanghai Pudong Development Bank.

In order to make the problem trackable, a perfect information game among three parties is designed, in which a TO purchases cell phones from a MPBC, the TO sells the cell phones with its telecom service to customers, and the MPBC manufactures and sells cell phones to both the TO and customers. The TO and MPBC are engaged in three Cournot games (the production quantity is assumed to be the sale quantity), and their products are substitutable for the customers. Besides, the MPBC has the financial constraint and need to borrow money from the bank. The interest rate is set by the bank (the detailed description of the game can be found in Section 3). Based on the analysis of this game, our study makes a number of contributions to the current literature. The contributions are as follows:

- (i) This paper studies a perfect information game among three parties: the bank, the TO, and the competitive MPBC. In the first stage, the bank decides an interest rate and shows it to the TO and the competitive MPBC. In the second stage, this work separates the subgames into three cases: a simultaneous subgame, a TO-as-leader sequential subgame, and an MPBC-as-leader sequential subgame. The TO's and the MPBC's optimal decisions and the optimal production/sale quantities are investigated.
- (ii) The impacts of the interest rate and the substitute factor on the TO's and the MPBC's optimal decisions are analyzed. When the substitute factor is high, at a low interest rate, the total sales in the simultaneous subgame is higher than those in the other two subgames; at a high interest rate, the total sales in the MPBC-as-leader subgame is higher than those in the other two subgames. However, when the substitute factor is low, at a low interest rate, the total sales in the simultaneous subgame is higher than those in the other two subgames; at a high interest rate, the total sales in the MPBC-as-leader subgame is higher than those in the other two subgames; at a moderate interest rate, the total sales in the TO-as-leader subgame is higher than those in the other two subgames. Besides, all the three slopes of the total sales are both negative and the slope of the total sales in the MPBC-as-leader subgame is higher than those in the other subgames, and the slope of the total sales in the TO-as-leader subgame is higher than that in the simultaneous subgames.
- (iii) The bank's decision and the optimal interest rate are investigated, and the impact of the substitute factor on the optimal interest rate is analyzed. The bank can set a higher interest rate in the MPBC-as-leader subgame than those in the other two subgames. Besides, when the substitute factor is low, the bank sets a low interest rate in the TO-as-leader subgame than that in the simultaneous subgame; however, when the substitute factor is high, the bank sets a higher interest rate in the TO-as-leader subgame than that in the simultaneous subgame.

The remainder of this paper is organized as follows. A literature review is provided in Section 2, and the model and preliminary results is presented in Section 3. In Section 4, the bank's decision is studied and the impact of the substitute factor on the bank's decision is analyzed. The conclusion is shown in Section 5. The proofs of the propositions are relegated to the appendix.

2. Literature Review

This work is closely related to the field of supply chain finance. Those works mainly focus on the financially constrained supply chain in which a supplier or/and a retailer need(s) to borrow money from a bank. Some works consider the presence of bankruptcy risks. Kouvelis and Zhao [1]

model a supply chain with a supplier and a retailer. Their strategic interaction is a Stackelberg game with the supplier as the leader. Both of them are capital constrained and in need of short-term financing. Kouvelis and Zhao [1] conclude that, under optimal trade credit contracts, both the supplier's profit and supply chain efficiency improve, and the retailer might improve his profits relative to under bank financing, depending on his current working capital and collateral. Kouvelis and Zhao [2] study a supply chain, in which a supplier sells a wholesale price contract to a financially constrained retailer who faces stochastic demand, so the retailer might need to borrow money from a bank to execute his order. Kouvelis and Zhao [2] find that with the presence of the retailer's bankruptcy risks, the increase of the retailer's wealth leads to the increase of the supplier's wholesale prices, but without the retailer's bankruptcy risks, the supplier's wholesale price stays the same or decreases in retailer's wealth. Xiao et al. [3] consider a financially constrained supply chain in which a supplier (leader) sells products to a retailer (follower) who has no access to bank financing due to her low credit rating. However, the supplier can borrow from a bank and offer trade credit to the retailer to alleviate her financial constraint. Xiao et al. [3] find that the revenue-sharing and buyback contracts can coordinate the supply chain only when the supply chain has a sufficient total working capital. More classical works on this stream can be found in [4–8].

This work is also related to studies of multichannel distribution and dual sales. Chiang et al. [9] build a model in which the manufacturer can open a direct channel to compete with its retailers. They show that the channel can benefit the manufacturer even when no direct sales occur. Tsay and Agrawal [10] study the channel conflict issue between direct sales and existing reseller partners and find that the addition of a direct channel is not necessarily detrimental to the reseller. Arya et al. [11] consider a dual distribution channel in which the manufacturer sells a product to a retailer and also competes with that retailer in the retail market. More works in this area can be found in the survey carried out by Tsay and Agrawal [12]. Recently, Wang et al. [13] build a supply chain, in which a contract manufacturer (CM) acts as both upstream partner and downstream competitor to an original equipment manufacturer (OEM). The two parties can engage in one of three Cournot competition games: a simultaneous game, a sequential game with the OEM as the Stackelberg leader, and a sequential game with the CM as the Stackelberg leader. On the basis of these three basic games, this study investigates the two parties' Stackelberg leadership/followership decisions. Matsui [14] applies an observable delay game framework developed in noncooperative game theory, in which a manufacturer manages dual-channel supply chains consisting of a retail channel and a direct channel. Matsui [14] investigates the timing problem concerning when the manufacturer should post its wholesale price and direct price, and they find that the manufacturer should post the direct price before or upon, but not after, setting the wholesale price for the retailer and this upfront posting of the direct price maximizes the

profits for a manufacturer employing multichannel sales strategies. Niu et al. [15] incorporate the concepts of channel power and fairness concern in a two-stage supply chain comprising a supplier and a retailer. With an online channel, the supplier competes with its retailer in a dual-channel system, but the retailer may shift part of or all orders to another supplier as the counteraction. They analyze the suppliers' decision on whether to open an online direct channel, and they find that the suppliers' fairness concern may effectively reduce its incentives to open an online channel.

Similar with the literature above, we also consider a dual-channel system, in which the MPBC not only sells the TO but also competes with the TO by the direct sales to the customers. Usually, the TO sells the phone with its telecommunication service. The MPBC is assumed to be financially constraint such that a commercial bank as the third party exists in this system. Consequently, we study a game among the three parties. In the first stage, the bank decides an interest rate and shows it to the TO and the competitive MPBC. In the second stage, the TO and MPBC are engaged in Cournot subgames: a simultaneous subgame, a TO-as-leader sequential subgame, and an MPBC-as-leader sequential subgame. On the basis of these three basic subgames, we investigate the TO's and the MPBC's optimal decisions and the optimal sales. We also analyze the impacts of the interest rate and the substitute factor on the TO's and the MPBC's optimal decisions.

3. Model

We consider that a TO (labeled t) purchases phones from a MPBC (labeled c). The TO sells the cell phones with its telecom service to customers. The MPBC manufactures and sells its products to both the TO and customers directly. The two parties' products are substitutable for the customers. To simplify the problem, the MPBC incurs the same production cost in producing its own and the TO's products.

The TO and the competitive MPBC engage quantity setting Cournot competition for customers (in the Cournot game, the production quantity is assumed to be the sale quantity). Thus, the market price of their products is jointly determined by their respective production quantities, that is, via inverse demand function. To make the model trackable, we adopt the commonly used inverse demand function for the differentiated product of firm c and t :

$$p_c(q_c, q_t) = m - q_c - b_c q_t, \quad (1)$$

where p_c is firm's market price, m is the upper bound on market size, q_c is its production quantity of the MPBC, q_t is its production quantity of the TO, and b_c is a parameter that measures the cross-effect of the change in firm c 's product demand caused by a change in that of firm t and is interpreted as the substitution rate of firm t 's product over that of firm c . Note that the limiting values $b_c = 0$ and $b_c = 1$ correspond to the cases of independent products and perfect substitutes, respectively. Similarly, the firm t 's market price is given as follows:

$$p_t(q_c, q_t) = m - q_t - b_t q_c. \quad (2)$$

The TO's products are always with the telecom service and promotion and are usually regarded as superior to those of the MPBC, and the former are assumed to be perfect substitutes for the latter, but the reverse is not true, that is, $b_c = 1$. Furthermore, $0 \leq b_t \leq 1$. To omit cases in which no production occurs, m is assumed to be sufficiently large relative to the wholesale price w . The MPBC's revenue consists of two parts: one is from customers and the other is from the TO. To simplify the model, the MPBC's marginal production cost is normalized to zero and the unit production cost is c_m . In this model, we assume that the production cost of the MPBC is borrowed from a bank, and the interest rate is denoted as r . Then, the MPBC's profit is

$$\pi_c = (m - q_c - q_t)q_c + wq_t - c_m(q_t + q_c)(1 + r). \quad (3)$$

The TO purchases the products (cell phones) from the MPBC and sells the products and its own telecom service to customers, and the service cost per product is denoted as c_t . Then, the TO's profit is

$$\pi_t = (m - q_t - bq_c)q_t - wq_t - c_t q_t. \quad (4)$$

Both (3) and (4) are concave and differentiable.

We denote i as the risk-free interest rate, which can be considered as the cost of the money for the bank. Then, the bank's profit is

$$\pi_b = (r - i)c_m(q_t + q_c). \quad (5)$$

We study a perfect information game among three parties: the bank, the TO, and the competitive MPBC. In the first stage, the bank decides an interest rate and shows it to the TO and the competitive MPBC. In the second stage, we separate the subgames into three cases: a simultaneous subgame, a TO-as-leader sequential subgame, and an MPBC-as-leader sequential subgame. We mainly focus on the subgames in the second stage in this section and will show the bank's decision in the second stage in the next section.

In this part, given the interest rate, the TO and the competitive MPBC can play three basic subgames: a simultaneous subgame, a TO-as-leader sequential subgame, and an MPBC-as-leader sequential subgame.

To explore the Stackelberg leadership preferences of the TO and the competitive MPBC and how those preferences affect the realization of the three aforementioned settings, we study a two-stage extended game called the endogenous timing game (see, e.g., [16, 17]). The extended game features a preplay stage in which the TO and the competitive MPBC simultaneously, though independently, choose either to move first and be the Stackelberg leader (denoted as L) or to move second and be the Stackelberg follower (denoted as F).

Denote π_i^S , $i = t, c$, as firm i 's profit when it is engaged in a simultaneous subgame, where S stands for simultaneous. Consequently, q_i^S , $i = t, c$, is the resulting production quantity. Denote π_i^L (π_i^F), $i = c, t$, as the firm i 's profit when it is the Stackelberg leader (follower). Let q_i^L (q_i^F), $i = c, t$, represent the corresponding production quantity.

3.1. Optimal Decisions and Payoffs. In this part, we compare the equilibrium sale quantity in the simultaneous and Stackelberg settings, in which the TO or the competitive MPBC would prefer Stackelberg leadership. Then, the closed-form expressions for the subgame perfect equilibrium quantities under the three setting are summarized in the following proposition.

Proposition 1. *In the simultaneous subgame, the equilibrium sale quantities are*

$$q_c^S = \frac{m + w + c_t - 2c_m(1 + r)}{4 - b}, \quad (6)$$

$$q_t^S = \frac{2(m - w - c_t) - b[m - c_m(1 + r)]}{4 - b}. \quad (7)$$

In the TO-as-leader subgame, the equilibrium sale quantities are

$$q_t^L = \frac{m - w - c_t}{2 - b} - \frac{b[m - c_m(1 + r)]}{2(2 - b)}, \quad (8)$$

$$q_c^F = \frac{m + w + c_t}{2(2 - b)} - \frac{bm + (4 - b)c_m(1 + r)}{4(2 - b)}. \quad (9)$$

In the MPBC-as-leader subgame, the equilibrium sale quantities are

$$q_c^L = \frac{m + w + c_t}{2(2 - b)} - \frac{bm + (4 - b)c_m(1 + r)}{4(2 - b)} + \frac{b[m - 2w + c_m(1 + r)]}{4(2 - b)}, \quad (10)$$

$$q_t^F = \frac{m - w - c_t}{2 - b} - \frac{b[m - c_m(1 + r)]}{2(2 - b)} - \frac{b[m - w - c_t - bw + bc_m(1 + r)]}{4(2 - b)}. \quad (11)$$

3.2. The Impact of r on the Sales Quantity. The TO's and MPBC's sale quantities decide their market share and their payoffs in the long term, so they are very important for both

the TO and the MPBC. In this section, we will analyze the impact of the interest rate on the MPBC's optimal sale, the TO's optimal sale, and the optimal total sale in the three

subgames. We find a threshold of b and five thresholds of r that affect the comparison results. The thresholds are shown as follows:

$$b_t = 1 - \frac{c_t}{m-w}, \quad (12)$$

$$r_1 = \frac{2w-m}{c_m} - 1, \quad (13)$$

$$r_2 = \frac{2w+2c_t+bm-2m}{bc_m} - 1, \quad (14)$$

$$r_3 = \frac{3w-m-c_t-bw}{(2-b)c_m} - 1, \quad (15)$$

$$r_4 = \frac{w+bw+c_t-m}{bc_m} - 1, \quad (16)$$

$$r_5 = \frac{(1-b)w-c_t}{(1-b)c_m} - 1. \quad (17)$$

Then, the MPBC's optimal sales comparison among three different subgames is shown as follows.

Proposition 2. When $b < b_t$, $r_2 < r_1 < r_3$:

$$\text{If } r \leq r_2 < r_1 < r_3, q_c^L < q_c^S \leq q_c^F$$

$$\text{If } r_2 < r \leq r_1 < r_3, q_c^L \leq q_c^F < q_c^S$$

$$\text{If } r_2 < r_1 < r \leq r_3, q_c^F < q_c^L \leq q_c^S$$

$$\text{If } r_2 < r_1 < r_3 < r, q_c^F < q_c^S < q_c^L$$

When $b > b_t$, $r_3 < r_1 < r_2$:

$$\text{If } r \leq r_3 < r_1 < r_2, q_c^L \leq q_c^S < q_c^F$$

$$\text{If } r_3 < r \leq r_1 < r_2, q_c^S < q_c^L \leq q_c^F$$

$$\text{If } r_3 < r_1 < r \leq r_2, q_c^S \leq q_c^F < q_c^L$$

$$\text{If } r_3 < r_1 < r_2 < r, q_c^F < q_c^S < q_c^L$$

When $b = b_t$, $r_2 = r_1 = r_3$; if $r \leq r_1$, $q_c^L \leq q_c^S \leq q_c^F$; if $r > r_1$, $q_c^F < q_c^S < q_c^L$.

When the substitute factor is high ($b \geq b_t$), only one threshold of the interest rate (r_1) affects the strategy of the MPBC. At a low interest rate ($r \leq r_1$), as a follower in the

sequential subgame, the MPBC can sell more products; at a high interest rate ($r_1 < r$), as a leader, the MPBC can sell more products.

When the substitute factor is low ($b < b_t$), there are two thresholds of the interest rate (r_2, r_3) that affect the strategy of the MPBC. At a low enough interest rate ($r \leq r_2$), as a follower, the MPBC can sell more products; at a high enough interest rate ($r_3 < r$), as a leader, the MPBC can sell more products; at a moderate interest rate ($r_2 < r \leq r_3$), a simultaneous subgame can help the MPBC sell more products.

The TO's optimal sales comparison among three different subgames is shown as follows.

Proposition 3. When $b < b_t$, $r_2 < r_4 < r_3$:

$$\text{If } r \leq r_2 < r_4 < r_3, q_t^L \leq q_t^S < q_t^F$$

$$\text{If } r_2 < r \leq r_4 < r_3, q_t^S < q_t^L \leq q_t^F$$

$$\text{If } r_2 < r_4 < r \leq r_3, q_t^S \leq q_t^F < q_t^L$$

$$\text{If } r_2 < r_4 < r_3 < r, q_t^F < q_t^S < q_t^L$$

When $b > b_t$, $r_3 < r_4 < r_2$:

$$\text{If } r \leq r_3 < r_4 < r_2, q_t^L < q_t^S \leq q_t^F$$

$$\text{If } r_3 < r \leq r_4 < r_2, q_t^L \leq q_t^F < q_t^S$$

$$\text{If } r_3 < r_4 < r \leq r_2, q_t^F < q_t^L \leq q_t^S$$

$$\text{If } r_3 < r_4 < r_2 < r, q_t^F < q_t^S < q_t^L$$

When $b = b_t$, $r_2 = r_4 = r_3$; if $r \leq r_4$, $q_t^L \leq q_t^S \leq q_t^F$; if $r > r_4$, $q_t^F < q_t^S < q_t^L$.

When the substitute factor is low ($b \leq b_t$), only one threshold of the interest rate (r_4) affects the strategy of the TO. At a low interest rate ($r \leq r_4$), as a follower, the TO can sell more products; at a high interest rate ($r_4 < r$), as a leader, the TO can sell more products.

When the substitute factor is high ($b > b_t$), there are two thresholds of the interest rate (r_2, r_3) that affect the strategy of the TO. At a low enough interest rate ($r \leq r_3$), as a follower, the TO can sell more products; at a high enough interest rate ($r_2 < r$), as a leader, the TO can sell more products; at a moderate interest rate ($r_3 < r \leq r_2$), simultaneous subgame can help the TO sell more products.

The total sales is $q = q_t + q_c$. Then, the total sales in the MPBC-as-leader subgame is

$$q^C = q_c^L + q_t^F = \frac{6m-2w-2c_t-3bm-(4-3b)c_m(1+r)}{4(2-b)} + \frac{b[c_t+bw-w+(1-b)c_m(1+r)]}{4(2-b)}, \quad (18)$$

the total sales in the TO-as-leader subgame is

$$q^T = q_c^F + q_t^L = \frac{6m-2w-2c_t-3bm-(4-3b)c_m(1+r)}{4(2-b)}, \quad (19)$$

and the total sales in the simultaneous subgame is

$$q^S = q_c^S + q_t^S = \frac{3m-w-c_t-bm-(2-b)c_m(1+r)}{4-b}. \quad (20)$$

Then, we analyze the impact of r on the total sales in the three subgames. By comparing the three slopes of the total sales, we have the following proposition.

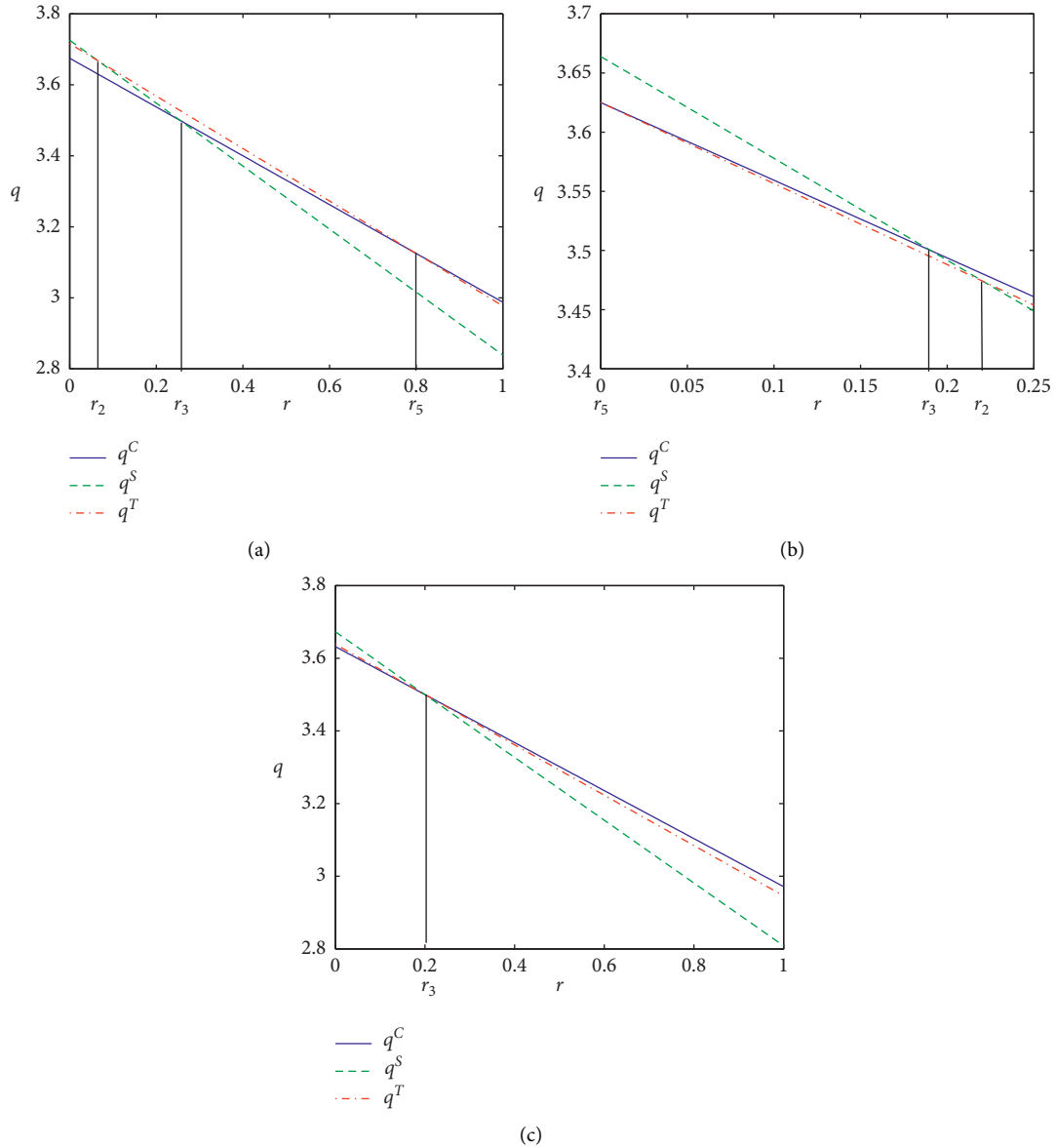


FIGURE 1: Illustration of the effects of r and b on the total sales: $m = 10$, $w = 6.5$, $c_t = 0.2$, $c_m = 2.5$, $b_t = 0.9429$, and $b = [0.90, 0.9429, 0.95]$. (a) $b < b_t$. (b) $b > b_t$. (c) $b = b_t$.

Proposition 4. $\partial q^S / \partial r < \partial q^T / \partial r < \partial q^C / \partial r < 0$.

q^S , q^C , and q^T linearly decrease in r and the slope of the total sales in the MPBC-as-leader subgame is lower than those in the other subgames, and the slope of the total sales in the TO-as-leader subgame is lower than that in the simultaneous subgames.

In order to analyze the impact of r , we also show impact of r on the total optimal sales (the sum of the TO's optimal sales and the MPBC's optimal sales) comparison among the three different subgames.

Proposition 5. When $b < b_t$, $r_2 < r_3 < r_5$:

$$\text{If } r \leq r_2 < r_3 < r_5, q^C < q^T \leq q^S$$

$$\text{If } r_2 < r \leq r_3 < r_5, q^C \leq q^S < q^T$$

$$\text{If } r_2 < r_3 < r \leq r_5, q^S < q^C \leq q^T$$

$$\text{If } r_2 < r_3 < r_5 < r, q^S < q^T < q^C$$

When $b > b_t$, $r_5 < r_3 < r_2$:

$$\text{If } r \leq r_5 < r_3 < r_2, q^C \leq q^T < q^S$$

$$\text{If } r_5 < r \leq r_3 < r_2, q^T < q^C \leq q^S$$

If $r_5 < r_3 < r \leq r_2$, $q^T \leq q^S < q^C$

If $r_5 < r_3 < r_2 < r$, $q^S < q^T < q^C$

When $b = b_t$, $r_2 = r_3 = r_5$; if $r \leq r_3$, $q^C \leq q^T \leq q^S$; if $r > r_3$, $q^S < q^T < q^C$.

We summarize the results in Propositions 4 and 5 in Figure 1. As shown in Figure 1, we set $m = 10$, $w = 6.5$, $c_t = 0.2$, and $c_m = 2.5$ and then obtain $b_t = 0.9429$. Obviously, all the three figures in Figure 1 show that $\partial q^S / \partial r < \partial q^T / \partial r < \partial q^C / \partial r < 0$. When $b = 0.9 < b_t$, as shown in Figure 1(a), $r_2 = 0.0667$, $r_3 = 0.2545$, and $r_5 = 0.8$. If $r \leq 0.0667$, $q^C < q^T \leq q^S$; if $0.0667 < r \leq 0.2545$, $q^C \leq q^S < q^T$; if $0.2545 < r \leq 0.8$, $q^S < q^C \leq q^T$; if $r > 0.8$, $q^S < q^T < q^C$. When $b = 0.95 > b_t$, as shown in Figure 1(b), $r_2 = 0.2211$, $r_3 = 0.1905$, and $r_5 = 0$. If $0 < r \leq 0.1905$, $q^T < q^C \leq q^S$; if $0.1905 < r \leq 0.2211$, $q^T \leq q^S < q^C$; if $r > 0.2211$, $q^S < q^T < q^C$. When $b = 0.9429 = b_t$, as shown in Figure 1(c), $r_2 = r_3 = r_5 = 0.2$. If $r \leq 0.2$, $q^C \leq q^T \leq q^S$; in contrast, if $r > 0.2$, $q^S < q^T < q^C$.

When $b < b_t$, as shown in Figure 1(a), there are two thresholds of the interest rate (r_2, r_5) that affect the total sales. At a low enough interest rate ($r \leq r_2$), the total sales in the simultaneous subgame is higher than those in the other two subgames; at a high enough interest rate ($r_5 < r$), the total sales in the MPBC-as-leader subgame is higher than those in the other two subgames; at a moderate interest rate ($r_2 < r \leq r_5$), the total sales in the TO-as-leader subgame is higher than those in the other two subgames.

When the substitute factor is high ($b \geq b_t$), as shown in Figure 1(b), only one threshold of the interest rate (r_3) affects the total sales. At a low interest rate ($r \leq r_3$), the total sales in the simultaneous subgame is higher than those in the other two subgames; at a high interest rate ($r_3 < r$), the total sales in the MPBC-as-leader subgame is higher than those in the other two subgames.

4. The Decision of the Bank

In this part, we consider the bank's decision in the first stage of the game among the three parties and the loan interest rate r . Recall that the profit of the bank is

$$\pi_b(r) = (r - i)c_m(q_t + q_c). \quad (21)$$

Under the three different subgames in the second stage, we obtain the optimal interest rates and summarize the results as follows.

Proposition 6. *In the simultaneous subgame, the equilibrium interest rate is*

$$r^S = \frac{(3 - b)m - w - c_t}{2c_m(2 - b)} - \frac{1 - i}{2}. \quad (22)$$

In the TO-as-leader subgame, the equilibrium interest rate is

$$r^T = \frac{3(2 - b)m - 2w - 2c_t}{2c_m(4 - 3b)} - \frac{1 - i}{2}. \quad (23)$$

In the MPBC-as-leader subgame, the equilibrium interest rate is

$$r^C = \frac{3m - (1 + b)w - c_t}{2c_m(2 - b)} - \frac{1 - i}{2}. \quad (24)$$

Next, we will make a comparison among the three interest rates: r^S , r^T , and r^C . Let $b_1 = c_t + w/m$. We summarize the comparison result as follows.

Proposition 7. *When $b < b_1$, $r^C > r^S > r^T$; when $b \geq b_1$, $r^C > r^T \geq r^S$.*

As shown in Proposition 7, the bank sets a higher interest rate in the MPBC-as-leader subgame than those in the other two subgames. Besides, when the value of b is low ($b < b_1$), the bank sets a low interest rate in the TO-as-leader subgame than that in the simultaneous subgame; however, when the value of b is high ($b \geq b_1$), the bank sets a higher interest rate in the TO-as-leader subgame than that in the simultaneous subgame.

5. Conclusion

In order to realize the sustainable development of the communication industry, many banks have developed some new projects to provide loan to the mobile phone brand companies (MPBCs). In this work, we study a perfect information game among three parties: the bank, the TO, and the competitive MPBC. In the first stage, the bank decides an interest rate and shows it to the TO and the competitive MPBC. In the second stage, we separate the Cournot subgame into three cases: a simultaneous subgame, a TO-as-leader sequential subgame, and an MPBC-as-leader sequential subgame. We investigate the TO's and the MPBC's optimal decisions and the optimal sales. Besides, we also analyze the impacts of the interest rate and the substitute factor on the TO's and the MPBC's optimal decisions.

We analyze the impacts of substitute factor and interest rate on the MPBC's sale quantity, the TO's sale quantity, and the total sale quantity. For the total sale, when the substitute factor is high, at a low interest rate, the total sales in the simultaneous subgame is higher than those in the other two subgames; at a high interest rate, the total sales in the MPBC-as-leader subgame is higher than those in the other two subgames. However, when the substitute factor is low, at a low enough interest rate, the total sales in the simultaneous subgame is higher than those in the other two subgames; at a high enough interest rate, the total sales in the MPBC-as-leader subgame is higher than those in the other two subgames; at a moderate interest rate, the total sales in the TO-as-leader subgame is higher than those in the other two subgames. Besides, all of the three slopes of the total sales are both negative and the slope of the total sales in the MPBC-as-leader subgame is lower than those in the other subgames, and the slope of the total sales in the TO-as-leader subgame is lower than that in the simultaneous subgames.

In addition, we also investigate the bank's decision, the interest rate and analyze the impact of the substitute factor on the interest rate. We find that the bank sets a higher

interest rate in the MPBC-as-leader subgame than those in the other two subgames. Besides, when the substitute factor is low, the bank sets a low interest rate in the TO-as-leader subgame than that in the simultaneous subgame; however, when the substitute factor is high, the bank sets a higher interest rate in the TO-as-leader subgame than that in the simultaneous subgame.

Appendix

Proof of Proposition 1. We consider three different subgames: the simultaneous subgame, the TO-as-leader Stackelberg subgame, and the MPBC-as-leader Stackelberg subgame. And, the proof will be separated into three parts.

In the simultaneous subgame, the two parties make the decisions simultaneously.

Given q_t , based on (3), taking the first-order derivative of π_c with respect to q_c , we obtain

$$\frac{\partial \pi_c}{\partial q_c} = m - q_c - q_t - q_c - c_m(1+r). \quad (\text{A.1})$$

Obviously, $\partial \pi_c / \partial q_c$ decreases in q_c and π_c is concave in q_c . Making $\partial \pi_c / \partial q_c = 0$, we obtain

$$q_c = \frac{m - q_t - c_m(1+r)}{2}. \quad (\text{A.2})$$

Given q_c , based on (4), taking the first-order derivative of π_t with respect to q_t , we obtain

$$\frac{\partial \pi_t}{\partial q_t} = m - q_t - bq_c - q_t - w - c_t. \quad (\text{A.3})$$

Obviously, $\partial \pi_t / \partial q_t$ decreases in q_t and π_t is concave in q_t . Making $\partial \pi_t / \partial q_t = 0$, we obtain

$$q_t = \frac{m - bq_c - w - c_t}{2}. \quad (\text{A.4})$$

Combining (A.2) with (A.4), we obtain

$$\begin{aligned} q_c^S &= \frac{m + w + c_t - 2c_m(1+r)}{4-b}, \\ q_t^S &= \frac{2(m - w - c_t) - b[m - c_m(1+r)]}{4-b}. \end{aligned} \quad (\text{A.5})$$

In the TO-as-leader subgame, the TO is the leader and the MPBC is the follower in the Stackelberg subgame. Given q_t , we have obtained the MPBC's optimal decision as shown in (A.2). Coupled with (A.2) and (4), we obtain

$$\pi_t = \left\{ m - q_t - \frac{b[m - q_t - c_m(1+r)]}{2} \right\} q_t - wq_t - c_tq_t. \quad (\text{A.6})$$

Based on the above equation, taking the first-order derivative of π_t with respect to q_t , we obtain

$$\frac{\partial \pi_t}{\partial q_t} = m - w - c_t - \frac{b[m - c_m(1+r)]}{2} - 2q_t + bq_t. \quad (\text{A.7})$$

Clearly, $\partial \pi_t / \partial q_t$ decreases in q_t and π_t is concave in q_t . Making $\partial \pi_t / \partial q_t = 0$, we obtain

$$q_t^L = \frac{m - w - c_t}{2-b} - \frac{b[m - c_m(1+r)]}{2(2-b)}. \quad (\text{A.8})$$

Coupled with (A.2), we have

$$q_c^F = \frac{m + w + c_t}{2(2-b)} - \frac{bm + (4-b)c_m(1+r)}{4(2-b)}. \quad (\text{A.9})$$

In the MPBC-as-leader subgame, the MPBC is the leader and the TO is the follower in the Stackelberg subgame. Similarly, based on (A.4) and (3), we can obtain

$$\begin{aligned} q_c^L &= \frac{m + w + c_t}{2(2-b)} - \frac{bm + (4-b)c_m(1+r)}{4(2-b)} + \frac{b[m - 2w + c_m(1+r)]}{4(2-b)}, \\ q_t^F &= \frac{m - w - c_t}{2-b} - \frac{b[m - c_m(1+r)]}{2(2-b)} - \frac{b[m - w - c_t - bw + bc_m(1+r)]}{4(2-b)}. \end{aligned} \quad (\text{A.10})$$

Proof of Proposition 2. Based on (6), (9), and (10), we obtain

$$q_c^L - q_c^F = \frac{b[m + c_m(1+r) - 2w]}{4(2-b)}, \quad (\text{A.11})$$

$$q_c^S - q_c^F = \frac{b[2m - 2w - 2c_t - bm + bc_m(1+r)]}{4(4-b)(2-b)}, \quad (\text{A.12})$$

$$q_c^L - q_c^S = \frac{b[m + c_t + bw - 3w + (2-b)c_m(1+r)]}{2(4-b)(2-b)}. \quad (\text{A.13})$$

where $q_c^L > q_c^F$ leads to $r > 2w - m/c_m - 1$, $q_c^S > q_c^F$ leads to $r > 2w + 2c_t + bm - 2m/bc_m - 1$, and $q_c^L > q_c^S$ leads to $r > 3w - m - c_t - bw/(2-b)c_m - 1$.

Let $r_1 = 2w - m/c_m - 1$, $r_2 = 2w + 2c_t + bm - 2m/bc_m - 1$, and $r_3 = 3w - m - c_t - bw/(2-b)c_m - 1$. Now, we will

compare the values of $r_i, i = 1, 2, 3$, with each other, and by doing so, can we compare the values of q_c^L, q_c^S , and q_c^F :

$$r_1 - r_2 = \frac{2bw - 2w - 2c_t - 2bm + 2m}{bc_m} = \frac{2[(1-b)(m-w) - c_t]}{bc_m}. \quad (\text{A.14})$$

As a result, $r_1 > r_2$ leads to $b < 1 - c_t/m - w$. Let $b_t = 1 - c_t/m - w$. Similarly, $r_2 > r_3$ leads to $b > b_t$ and $r_1 > r_3$ leads to $b > b_t$.

Therefore, when $b < b_t, r_2 < r_1 < r_3$. If $r \leq r_2 < r_1 < r_3$ (or $r \leq r_2, r < r_1$, and $r < r_3$), $q_c^L < q_c^S \leq q_c^F$ (or $q_c^S \geq q_c^F, q_c^L > q_c^F$, and $q_c^L > q_c^S$).

Similarly, if $r_2 < r \leq r_1 < r_3, q_c^L \leq q_c^F < q_c^S$; if $r_2 < r_1 < r \leq r_3, q_c^F < q_c^L \leq q_c^S$; if $r_2 < r_1 < r_3 < r, q_c^L < q_c^S < q_c^F$.

When $b > b_t, r_3 < r_1 < r_2$. Similarly, if $r \leq r_3 < r_1 < r_2, q_c^L \leq q_c^S < q_c^F$; if $r_3 < r \leq r_1 < r_2, q_c^S < q_c^L \leq q_c^F$; if $r_3 < r_1 < r \leq r_2, q_c^S \leq q_c^F < q_c^L$; if $r_3 < r_1 < r_2 < r, q_c^F < q_c^S < q_c^L$.

When $b = b_t, r_2 = r_1 = r_3$; if $r \leq r_1, q_c^L \leq q_c^S \leq q_c^F$; if $r > r_1, q_c^F < q_c^S < q_c^L$. \square

Proof of Proposition 3. The proof is similar to the proof of Proposition 2. \square

Proof of Proposition 4. Taking the first-order derivative of q with respect to r in the three subgames, we have

$$\frac{\partial q^C}{\partial r} = -c_m \frac{(2-b)}{4}, \quad (\text{A.15})$$

$$\frac{\partial q^T}{\partial r} = -c_m \frac{(4-3b)}{(8-2b)}, \quad (\text{A.16})$$

$$\frac{\partial q^S}{\partial r} = -c_m \frac{(2-b)}{(4-b)}, \quad (\text{A.17})$$

$$\begin{aligned} \frac{\partial q^T}{\partial r} - \frac{\partial q^C}{\partial r} &= \frac{c_m b(b-1)}{8-4b} < 0, \\ \frac{\partial q^S}{\partial r} - \frac{\partial q^C}{\partial r} &= \frac{-b(2-b)}{4(4-b)} < 0, \\ \frac{\partial q^T}{\partial r} - \frac{\partial q^S}{\partial r} &= \frac{b^2}{(4-b)(8-4b)} > 0. \end{aligned} \quad (\text{A.18})$$

Proof of Proposition 5. The Proof is similar to the proof of Proposition 2. \square

Proof of Proposition 6. We also separate the proof into three parts: the simultaneous subgame, the TO-as-leader Stackelberg subgame, and the MPBC-as-leader Stackelberg subgame.

In the simultaneous subgame, based on (21), coupled with (6) and (7), we obtain

$$\begin{aligned} \pi_b^S &= (r-i)c_m(q_t^S + q_c^S) \\ &= (r-i)c_m \left\{ \frac{m+w+c_t-2c_m(1+r)}{4-b} \right. \\ &\quad \left. + \frac{2(m-w-c_t)-b[m-c_m(1+r)]}{4-b} \right\} \\ &= \frac{c_m}{4-b} [(3-b)m-w-c_t-(2-b)c_m(1+r)](r-i). \end{aligned} \quad (\text{A.19})$$

Based on (A.19), taking the first-order derivative of π_b^S with respect to r , we obtain

$$\frac{\partial \pi_b^S}{\partial r} = \frac{c_m}{4-b} \{(3-b)m-w-c_t-(2-b)c_m(1-i+2r)\}. \quad (\text{A.20})$$

Since $0 \leq b \leq 1, \partial \pi_b^S / \partial r$ decreases in r . Making $\partial \pi_b^S / \partial r = 0$, we obtain

$$r^S = \frac{(3-b)m-w-c_t}{2c_m(2-b)} - \frac{1-i}{2}. \quad (\text{A.21})$$

In the TO-as-leader subgame, based on (21), coupled with (8) and (9), we obtain

$$\pi_b^T = \frac{c_m}{4(2-b)} [6m-2w-2c_t-3bm-(4-3b)c_m(1+r)](r-i). \quad (\text{A.22})$$

Similarly, taking the first-order derivative of π_b^T with respect to r and making $\partial \pi_b^T / \partial r = 0$, we obtain

$$r^T = \frac{3(2-b)m-2w-2c_t}{2c_m(4-3b)} - \frac{1-i}{2}. \quad (\text{A.23})$$

In the MPBC-as-leader subgame, based on (21), coupled with (10) and (11), we obtain

$$\pi_b^C = \frac{c_m}{4} [3m-w-c_t-bw-(2-b)c_m(1+r)](r-i). \quad (\text{A.24})$$

Similarly, taking the first-order derivative of π_b^C with respect to r and making $\partial \pi_b^C / \partial r = 0$, we obtain

$$r^C = \frac{3m-(1+b)w-c_t}{2c_m(2-b)} - \frac{1-i}{2}. \quad (\text{A.25})$$

Proof of Proposition 7. In this part, we will make an comparison among r^S, r^T , and r^C . Based on (22) and (23),

$$\begin{aligned} r^S - r^T &= \frac{(3-b)m-w-c_t}{2c_m(2-b)} - \frac{3(2-b)m-2w-2c_t}{2c_m(4-3b)} \\ &= \frac{w+c_t-bm}{2c_m(2-b)(4-3b)}. \end{aligned} \quad (\text{A.26})$$

As a result, $r^S > r^T$ when $b < c_t + w/m$; $r^S \leq r^T$ when $b \geq c_t + w/m$. Let $b_1 = c_t + w/m$. Based on (22) and (24),

$$r^C - r^S = \frac{3m - (1+b)w - c_t}{2c_m(2-b)} - \frac{(3-b)m - w - c_t}{2c_m(2-b)} = m - w. \quad (\text{A.27})$$

where $m > w$; then, $r^C > r^S$.

Based on (23) and (24),

$$\begin{aligned} r^C - r^T &= \frac{3m - (1+b)w - c_t}{2c_m(2-b)} - \frac{3(2-b)m - 2w - 2c_t}{2c_m(4-3b)} \\ &= \frac{3b(1-b)[m - w + c_t/3(1-b)]}{2c_m(2-b)(4-3b)}. \end{aligned} \quad (\text{A.28})$$

Because $b < 1$ and $m > w$, $m - w + c_t/3(1-b) > 0$. As a result, $r^C > r^T$.

Combining the results above, we summarize the conclusion as follows.

When $b < b_1$, $r^C > r^S > r^T$; when $b \geq b_1$, $r^C > r^T \geq r^S$. \square

Data Availability

The data used to support the findings of this study are included within the article.

Disclosure

The part of the manuscript was presented in 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This project was supported in part by the National Natural Science Foundation of China under Grant no. 71801233, the National Social Science Foundation of China under Grant no. 19BJY184, and the Philosophy and Social Sciences Foundation of Guangxi under Grant no. 20FGL006.

References

- [1] P. Kouvelis and W. Zhao, "Financing the newsvendor: supplier vs. bank, and the structure of optimal trade credit contracts," *Operations Research*, vol. 60, no. 3, pp. 566–580, 2012.
- [2] P. Kouvelis and W. Zhao, "The newsvendor problem and price-only contract when bankruptcy costs exist," *Production and Operations Management*, vol. 20, no. 6, pp. 921–936, 2011.
- [3] S. Xiao, S. P. Sethi, M. Liu, and S. Ma, "Coordinating contracts for a financially constrained supply chain," *Omega*, vol. 72, 2016.
- [4] J. A. Buzacott and R. Q. Zhang, "Inventory management with asset-based financing," *Management Science*, vol. 50, no. 9, pp. 1274–1292, 2004.
- [5] M. Dada and Q. Hu, "Financing newsvendor inventory," *Operations Research Letters*, vol. 36, no. 5, pp. 569–573, 2008.
- [6] C. H. Lee and B.-D. Rhee, "Trade credit for supply chain coordination," *European Journal of Operational Research*, vol. 214, no. 1, pp. 136–146, 2011.
- [7] X. Xu and J. R. Birge, *Joint Production and Financing Decisions: Modeling and Analysis*, SSRN, Rochester, NY, USA, 2004.
- [8] S. A. Yang and J. R. Birge, "Trade credit, risk sharing, and inventory financing portfolios," *Management Science*, vol. 64, no. 8, pp. 3667–3689, 2018.
- [9] W.-Y. K. Chiang, D. Chhajed, and J. D. Hess, "Direct marketing, indirect profits: a strategic analysis of dual-channel supply-chain design," *Management Science*, vol. 49, no. 1, pp. 1–20, 2003.
- [10] A. A. Tsay and N. Agrawal, "Channel conflict and coordination in the e-commerce age," *Production and Operations Management*, vol. 13, no. 1, pp. 93–110, 2004.
- [11] A. Arya, B. Mittendorf, and D.-H. Yoon, "Friction in related-party trade when a rival is also a customer," *Management Science*, vol. 54, no. 11, pp. 1850–1860, 2008.
- [12] A. A. Tsay and N. Agrawal, "Modeling conflict and coordination in multi-channel distribution systems: a review," in *Handbook of Quantitative Supply Chain Analysis*, pp. 557–606, Springer US, New York, NY, USA, 2004.
- [13] Y. Wang, B. Niu, and P. Guo, "On the advantage of quantity leadership when outsourcing production to a competitive contract manufacturer," *Production and Operations Management*, vol. 22, no. 1, pp. 104–119, 2013.
- [14] K. Matsui, "When should a manufacturer set its direct price and wholesale price in dual-channel supply chains?" *European Journal of Operational Research*, vol. 258, no. 2, pp. 501–511, 2017.
- [15] B. Niu, Q. Cui, and J. Zhang, "Impact of channel power and fairness concern on suppliers market entry decision," *Journal of the Operational Research Society*, vol. 68, pp. 1–12, 2017.
- [16] R. Amir and A. Stepanova, "Second-mover advantage and price leadership in bertrand duopoly," *Games and Economic Behavior*, vol. 55, no. 1, pp. 1–20, 2006.
- [17] J. H. Hamilton and S. M. Slutsky, "Endogenous timing in duopoly games: Stackelberg or cournot equilibria," *Games and Economic Behavior*, vol. 2, no. 1, pp. 29–46, 1990.

Research Article

Gaussian Kernel Fuzzy C-Means Algorithm for Service Resource Allocation

Wei Jiang,¹ Xi Fang ,² and Jianmei Ding³

¹College of Economics and Management, Xi'an University of Posts and Telecommunications, Xi'an 710061, China

²School of Economics and Management, Shanghai Institute of Technology, Shanghai 200235, China

³School of Economics and Management, Nanjing University of Science and Technology, Nanjing 210094, China

Correspondence should be addressed to Xi Fang; fangxi@sit.edu.cn

Received 30 June 2020; Revised 20 November 2020; Accepted 5 December 2020; Published 19 December 2020

Academic Editor: Xiaobo Qu

Copyright © 2020 Wei Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With respect to the cluster problem of the evaluation information of mass customers in service management, a cluster algorithm of new Gaussian kernel FCM (fuzzy C-means) is proposed based on the idea of FCM. First, the paper defines a Euclidean distance formula between two data points and makes them cluster adaptively based on the distance classification approach and nearest neighbors in deleting relative data. Second, the defects of the FCM algorithm are analyzed, and a solution algorithm is designed based on the dual goals of obtaining a short distance between whole classes and long distances between different classes. Finally, an example is given to illustrate the results compared with the existing FCM algorithm.

1. Introduction

Clustering is an unsupervised learning method that is not reliant on predefined classes and training datasets with class labels. Clustering objects are divided into classes or clusters on the basis of feature similarity measurement. Hence, the same clusters share high similarities within the same cluster but largely differ from each other between different clusters. Traditional clustering methods are primarily based on the partition, hierarchy, grid, density, and model. As data mining in rapid development necessitates higher requirements for clustering, a clustering algorithm based on sample attribution, preprocessing, similarity measurement, allocation and scheduling, update strategy, and measurement [1, 2] have been advanced and applied to data mining [3, 4]. Considering the fuzziness of membership between sample points and cluster centers, the objective function-based fuzzy c-means (FCM) algorithm still prevails in theory and practice.

The core of an FCM algorithm is to design and determine a clustering center. The design mainly consists of quantifying cluster centers, locating them, and scheming an objective function accordingly. The cluster centers are quantified

manually in most cases, or their optimal amount is determined in a given range using information entropy and other methods. For example, Duan and Wang [5] indicated that the clustering center was acquired by multiattribute information with broken-line fuzzy numbers. A novel clustering algorithm, Nei Mu, was proposed in [6] based on which datasets are converted into data points of attribute space to construct a directed graph of K-nearest neighbors. This algorithm contributes to upgrading the clustering of data with large density fluctuation and an arbitrary distribution, but not all data points have K-nearest neighbors. Xue and Sha [7] initiated a coordinate-based density method using a gray prediction model of a clustering algorithm to determine the initial clustering center.

A clustering center should be determined with modifications in a dynamic process. The existing determination methods largely include K-means clustering algorithms, partition- and density-based clustering algorithms, clustering algorithms based on the local density of data points, and KZZ algorithms. For these, the K-means algorithm is used with a given initial center, whereas partition- and density-based clustering algorithms are used to determine the initial clustering center by a density function of sample

points using max-min distance means or the maximum distance product method. Zhang and Wang [8] pointed out the nearest data points were bracketed to facilitate location of other clustering centers at the same time that a resolution with high constraints was added to the objective function. Chiu Stephen [9] defined measures for each data point to identify the initial clustering center. Agustin et al. [10] studied a group genetic algorithm, aiming to improve the performance of group clustering by coding and defining fitness functions. A semisupervised clustering algorithm was put forward in [11] via the kernel FCM clustering algorithm with clustering errors containing labeled and unlabeled data to design the objective function. Since FCM fails to deal with noise, an efficient kernel-induced FCM based on a Gaussian function was presented in [12] to improve the objective function.

The following cases are some of the existing FCM studies. Qian and Yao [13] focused on high sensitivity to the initial center point and introduced three incremental fuzzy clustering algorithms for large-scale sparse high-dimensional datasets. Niu and She [14] proposed a fast parallel clustering algorithm based on cluster initialization. By generating a hierarchical K-means clustering tree to auto-select the number of clusters, Hu [15] obtained better clustering results. Aiming at the high time complexity of traditional FCM algorithms, a single-pass Bayesian fuzzy clustering algorithm was advocated for large-scale data in [16], which boosted its performance in time complexity and convergence. Zhou et al. [17] introduced the neighborhood information of multidimensional data to improve the clustering algorithm, increasing the robustness of outliers and noise points. Chen and Liu [18] designed a clustering algorithm on the minimum connected dominating set to remedy the defect that common algorithms easily fall into local minimum points. Xie et al. [19] combined the GWO algorithm with the principle of maximum entropy in a multidimensional big data environment. Duan and Wang [5] described multiple attributes of the objects to be clustered as polygonal fuzzy numbers, and a clustering algorithm was designed accordingly. By advancing an adaptive algorithm for the entropy weight of the feature weight of FCM, Huang et al. [20] focused on the influence of the feature weight on a clustering algorithm. Taking the preference vectors' clustering degree as a neighborhood similarity, Xu and Fan [21] aimed at constructing a heuristic clustering algorithm for multiattribute complex large group clustering and decision.

These documents focus on FCM algorithm-associated issues, but few achievements have been made in big data scenarios. In view of the differences between large data point clustering and small sample point clustering, the sample points of big data were simplified in this paper, making FCM more applicable for big data scenarios. Next, an FCM algorithm was designed by taking both long between-class distances and short inner-class distances into consideration,

which traditional FCM algorithms failed to do. This study thus provides theoretical and practical guidance for data clustering in a big data environment.

2. Gaussian Kernel FCM Clustering Algorithm

Since service resources are generally allocated in multiple ways, and there is a reciprocal relationship between the limited resources in one channel of allocation and those in another in terms of resource quantity, group consistency is beyond reach in which different resource consumers prefer different channels, leading to changing evaluation data. If the price mechanism fails to optimize the service resource allocation, consumer demands should be considered while pursuing social benefits to attain higher efficiency of resource allocation. Consumers primarily feature heterogeneity, conflicts of interest, and differences in evaluation forms, which necessitate decomposition of the customer group to divide the large-scale consumer groups into several small clusters, thus simplifying resource coordination.

Suppose that the consumer subject of a service resource is expressed as $X = (x_1, \dots, x_n)$, individual consumer as x_i , $i \in (1, \dots, n)$, number of channels (data dimension) as p , evaluation data as x_{ij} , $j \in (1, \dots, p)$, U_{ik} is the membership of sample i in Class k , with fuzzy matrix $U = [u_{ik}]_{n \times c}$ provided that there are c classes, and y_k , $k \in (1, \dots, c)$ represents cluster centers. Its objective function can be represented by the Gaussian kernel FCM clustering algorithm [8]:

$$J(U, V) = 2 \sum_{i=1}^n \sum_{k=1}^c u_{ik}^m (\beta - \beta \phi(x_i, y_k)), \quad (1)$$

where $\phi(x_i, y_k) = \exp(-\|x_i - y_k\|^2 / \sigma^2)$, β is a characteristic constant of the Gaussian function, and m is a fuzzy index used to control the fuzzy degree of classification. The higher the index, the higher the fuzzy degree. σ^2 is the variance of the given data. Hence,

$$u_{ik} = \frac{(1/(\beta - \beta \phi(x_i, y_k)))^{(1/m-1)}}{\sum_{k=1}^c (1/(\beta - \beta \phi(x_i, y_k)))^{(1/m-1)}}, \quad (2)$$

$$y_k = \frac{\sum_{i=1}^n u_{ik}^m \beta \phi(x_i, y_k) x_i}{\sum_{i=1}^n u_{ik}^m \beta \phi(x_i, y_k)}. \quad (3)$$

If $\|V^{\text{present}} - V^{\text{previous}}\| \leq \varepsilon$, then the iteration is discontinued, at which there is optimum classification. Both traditional FCM algorithms and the Gaussian kernel FCM clustering algorithm focus on the inner-class distance instead of between-class distance. Result-oriented, both values should be considered in order to obtain better clustering. Due to the large number of consumers to whom service resources are allocated, direct computing of membership will lead to problems such as high computing complexity and slow convergence of the optimal solution, resulting in a

decline in clustering efficiency. Therefore, preprocessing of data points should occur prior to clustering in order to reduce the number of data points that need clustering and enhance the scalability of the clustering algorithm.

3. Preprocessing of Evaluation Information of Consumers

$\forall x_i, x_j$ can be considered as a constraint data pair. The Euclidean distance formula is deployed to calculate its distance:

$$d(x_i, x_j) = \|x_i - x_j\| = \sqrt{(x_{i1} - x_{j1})^2 + \dots + (x_{ip} - x_{jp})^2}. \quad (4)$$

Set ε and γ in advance for $d(x_i, x_j)$ (both ε and γ can take lower values for more accurate classification).

- (i) If $d(x_i, x_j) \leq \varepsilon$, then it is considered that x_i and x_j are extremely close and can be placed into one class
- (ii) If $d(x_i, x_j) \geq \gamma$, it is considered that x_i is far from x_j , and bracketing them together is next to impossible
- (iii) Data points between ε and γ cannot be effectively identified

To delete data points quickly, the characteristics of distances between data points and the possibility of clustering different data points should be considered and investigated in the preprocessing procedure. Deletion should be done via the following steps:

- (i) Step 1: take data points x_i and x_j with the smallest distance in X to meet $d(x_i, x_j) \leq \varepsilon$, and combine $F_1 = \phi$ with x_i and x_j . Then, $F_1 + \{x_i, x_j\} \rightarrow F_1$, and $X - \{x_i, x_j\} \rightarrow X$.
- (ii) Step 2: take the mean value of x_i and x_j by $(x_i + x_j)/2$ as a new data point, and identify data point x_l in Set X whose mean value is less than ε . Then, $F_1 + \{x_l\} \rightarrow F_1$ and $X - \{x_l\} \rightarrow X$.
- (iii) Step 3: take $\{x_i, x_j, x_l\}$ as a new data point with a value of $((x_i + x_j)/2 + x_l)/2$. Repeat Step 2 until no new data points can be found, and form new sets F_1 and X .
- (iv) Step 4: repeat Steps 1 to 3 for Set X form Set Y , including F_1, \dots, F_m and X , wherein the final mean values of data points in F_1, \dots, F_m are taken as new data points, respectively.
- (v) Step 5: let data points in Y be nodes in the graph based on graph theory, and the connecting line of nodes be the distance. If the distance is greater than γ , then the connecting line is deleted, thus forming a connected network graph. Assuming that points a, b , and c make a circle, and a is the farthest from b , it can be considered that there is a higher probability of forming a cluster by a and b than by a and c , or b and c , so the connecting line can be

deleted. Here, Graph G without cycles is the connected network graph.

- (vi) Step 6: in Graph G with a plurality of nodes, the nodes are sorted by the number of adjacent points. Each node has a plurality of nondominated nodes to make a cluster.
- (vii) Step 7: since cluster sides vary in length, it is difficult to generate an effective cluster set for clusters with their average value as its sides. Deletion leads to basically equal distance from each data point in a cluster to the center of the cluster. Thus, point estimation may be adopted to work out its expected value. Given that Cluster C_l has q neighbors, its sample variance is s_l , $(x_i - \bar{x}) / (s_l / \sqrt{q}) \leq t_{1-\alpha}$. An adaptive k-nearest neighbor algorithm is used to search a data point closest to or within a set distance from the given data point and merge it into clustering components for clustering fusion. Therefore, x_i that satisfies the formula will be included in Cluster C_l ; otherwise, the data point is deleted from C_l . If included in multiple classes, the data point will enter the cluster as a matter of priority, where it is the nearest to the cluster center, so the average value of all data points in Cluster C_l is data point z_l .
- (viii) Step 8: data points are downsized using the approach mentioned above, and the pertinence of clustering by Set Z is strengthened. The original dataset X becomes Set $Z = (z_1, \dots, z_m)$.

4. Clustering Algorithm of Consumer Evaluation

The number of clusters and the initial cluster center must be determined first to cluster by the FCM clustering algorithm. The former may be obtained by manually determining or defining an interval range and giving preference to the best cluster number. From the perspective of consumer clustering, the number of clusters is that of the evaluation channels, which is expressed as P , since clustering better coordinates the needs of consumers that prefer different service channels.

The initial cluster center will change with the objective function value of the optimal fuzzy classification. However, it is difficult to meet the difference requirements between classes. In general, scholars add a penalty function to the objective function of the existing model (1) or similar models to maximize the between-class distance $(1/p(p-1)) \sum_{q=1}^p \sum_{k=1}^p d(y_q, y_k)$. However, the following problems are encountered:

- (i) The inner-class distance is a function of the distances between each data point and the cluster center and the power of membership, while the between-class distance is the average value of distance differences between cluster centers. Both vary widely in value and thus are beyond comparison. Their

incorporation into an objective function (minimum value) may fail to accommodate maximizing the between-class distance and minimizing the inner-class distance in an iteration, but may focus on the former.

- (ii) Iteration termination occurs on the basis that understanding the difference of objective functions is within a specific range, with which the optimal cluster center and membership function can be obtained. As a likely nonconvex function with a local optimal solution, the objective function at the end of the iteration may not be solved, resulting in a small value, and there may be a small difference in the objective function values of two iterations and large values in another two. The algorithm cannot prove its convergence.

To maintain a short inner-class distance and long between-class distance, partitioning the two indexes and setting a more appropriate iteration termination condition should be necessary based on the above considerations. Then, determination of an optimized cluster center can proceed. The steps are as follows:

- (i) Step 1: as the clustering result is sensitive to the selection of the initial cluster centers, the distance between cluster centers should be increased as much as possible. The dominant point with the most neighbors in dataset Z is taken as the first cluster center, the data point farthest from the dominant point as the second, the data point with the largest product of the distance to the two cluster centers as the third, and so on, until P initial clustering centers are solved.
- (ii) Step 2: calculate u_{ik} and y_k , respectively, by equations (2) and (3). β can be given or estimated by a sample variance: $(1/p)\sum_{j=1}^p ((1/n-1)\sum_{i=1}^n (x_{ij} - (1/n)\sum_{i=1}^n x_{ij})^2)$.
- (iii) Step 3: set the threshold α of the inner-class distance and β of the between-class distance. Variance $s = \sqrt{(1/p-1)\sum_{k=1}^p (y_k - \bar{y}_k)^2}$ is used to characterize between-class differences, where $\bar{y}_k = (1/p)\sum_{k=1}^p y_k$.
- (iv) Step 4: if $J(U, V) \leq \alpha$ and $s \geq \beta$, then the iteration is terminated, and the obtained u_{ik} and y_k are the most suitable membership function and cluster center, respectively.
- (v) Step 5 (sample classification): work out the distance $\phi(x_i, y_k)$ between data points $z_j, j \in (1, \dots, m)$ and each cluster center, and classify the minimum values.

5. Simulation Research

Given that a service resource targets a large number of consumers and may be allocated via five channels, a random sample survey was conducted on 100 consumers to seek their

service evaluation data on each channel. The consumer group is clustered to pursue a more effective allocation of resources.

By the given steps, the possibility of clustering each data point (consumer) is preprocessed based on the evaluation data. Calculate the distance between two respective evaluation data points by formula (4), and $\varepsilon = 0.2$ and $\gamma = 0.70$:

- (i) Data close to each other are initially clustered by Steps 1 to 4 in Section 3 to obtain Set Y ($Y = \{F_1, F_2, \dots, F_{23}, X\}$) composed of F_1, \dots, F_m and X , where in $F_1 = \{x_1, x_3\}$ and $F_2 = \{x_{58}, x_{77}\} \dots F_i$ are regarded as new data and each data in X as separate data. Thus, the initial evaluation set is simplified to Set Y with 72 data points.
- (ii) By step 5, the data in Set Y are processed for connected graph G without circles (see Figure 1) formed by elements in Y , wherein isolated points and points not on the main connected graph are not drawn.
- (iii) Process each node on Graph G by employing steps 6 and 7 to conclude dataset Z (including 41 data points) on the basis of an adaptive nearest-neighbor classification rule. z_i represents one or several data points in X , whose partial relationship is shown in Table 1.
- (iv) Since the evaluation data involve 5 channels, find 5 initial cluster centers by Step 1 in Section 4:
 $y_1^0 = (0.50, 0.42, 0.66, 0.53, 0.56)$,
 $y_2^0 = (0.96, 0.98, 0.17, 0.87, 0.06)$,
 $y_3^0 = (0.13, 0.01, 0.11, 0.04, 0.70)$,
 $y_4^0 = (0.88, 0.06, 0.59, 0.92, 1.00)$, and
 $y_5^0 = (0.03, 1.00, 0.28, 0.26, 0.44)$.
- (v) Given $m = 2$, calculate u_{ik} and y_k with Steps 2 to 4. Suppose $\alpha = 130.5$ and $\beta = 0.3$. The iteration stops, provided that $J = 130.43$, $s = 0.3227$, and $t = 14$, so the cluster center is in its optimal state after 14 iterations:
 $y_1^{14} = (0.45, 0.32, 0.48, 0.24, 0.53)$,
 $y_2^{14} = (0.92, 0.97, 0.18, 0.85, 0.09)$,
 $y_3^{14} = (0.33, 0.19, 0.09, 0.30, 0.48)$,
 $y_4^{14} = (0.87, 0.06, 0.59, 0.92, 0.99)$, and
 $y_5^{14} = (0.27, 0.76, 0.15, 0.17, 0.27)$.
- (vi) Perform Step 5 to classify samples, procuring a clustering result of Set Z .
- (vii) Get Set X based on Table 2 and the corresponding relationship between Set Z and Set X in Table 1. This is the clustering result of the original evaluation data, as shown in Table 3.

The distance $\|V^{\text{present}} - V^{\text{previous}}\|$ (expressed by $d_{i,j}$) upon 33 iterations is shown in Table 4, according to [9].

Changes in $d_{i,j}$ are shown in Figure 2.

Figure 2 illustrates that the value of $d_{i,j}$ increases first, decreases next, and then increases and is free of monotone convergence. The clustering center may not be optimal when

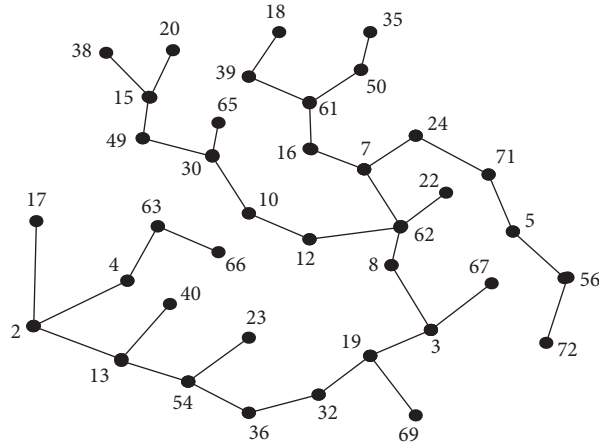


FIGURE 1: Connected graph of elements in Set Y.

TABLE 1: Correspondence between Set Z and Set X.

Z	X
z_1	x_{63}, x_{81}
z_2	$x_{58}, x_{77}, x_{92}, x_{97}$
z_3	$x_{65}, x_{76}, x_{95}, x_{96}$
z_4	x_{26}
z_5	$x_{12}, x_{19}, x_{46}, x_{90}$
\vdots	\vdots
z_{41}	x_{37}, x_{44}, x_{98}

TABLE 2: Cluster results of Set Z.

C_1	$z_8, z_9, z_{13}, z_{14}, z_{16}, z_{17}, z_{20}, z_{23}, z_{24}, z_{27}, z_{28}, z_{31}, z_{41}$
C_2	$z_{10}, z_{11}, z_{22}, z_{40}$
C_3	$z_1, z_2, z_3, z_{12}, z_{26}, z_{30}, z_{33}, z_{34}, z_{35}, z_{37}$
C_4	$z_{21}, z_{29}, z_{32}, z_{36}$
C_5	$z_4, z_5, z_6, z_7, z_{15}, z_{18}, z_{19}, z_{25}, z_{38}, z_{39}$

TABLE 3: Cluster results of Set X.

C_1 (38)	$x_2, x_4, x_8, x_{10}, x_{15}, x_{17}, x_{20}, x_{23}, x_{25}, x_{29}, x_{30}, x_{32}, x_{34}, x_{37}, x_{38}, x_{39}, x_{42}, x_{44}, x_{45}, x_{51}, x_{55}, x_{57}, x_{59}, x_{61}, x_{62}, x_{64}, x_{67}, x_{68}, x_{73}, x_{75}, x_{78}, x_{83}, x_{84}, x_{85}, x_{89}, x_{91}, x_{98}, x_{99}$
C_2 (8)	$x_7, x_{16}, x_{24}, x_{43}, x_{56}, x_{60}, x_{82}, x_{88}$
C_3 (26)	$x_1, x_3, x_6, x_9, x_{21}, x_{35}, x_{48}, x_{49}, x_{52}, x_{53}, x_{58}, x_{63}, x_{65}, x_{70}, x_{72}, x_{76}, x_{77}, x_{79}, x_{81}, x_{92}, x_{93}, x_{94}, x_{95}, x_{96}, x_{97}, x_{100}$
C_4 (10)	$x_5, x_{11}, x_{18}, x_{27}, x_{31}, x_{40}, x_{41}, x_{47}, x_{50}, x_{71}$
C_5 (18)	$x_{12}, x_{13}, x_{14}, x_{19}, x_{22}, x_{26}, x_{28}, x_{33}, x_{36}, x_{46}, x_{54}, x_{66}, x_{69}, x_{74}, x_{80}, x_{86}, x_{87}, x_{90}$

TABLE 4: $d_{i,j}$ after 33 iterations.

i, j	$d_{i,j}$	i, j	$d_{i,j}$	i, j	$d_{i,j}$	i, j	$d_{i,j}$
1, 2	0.0130	9, 10	0.0276	17, 18	0.0024	25, 26	0.0023
2, 3	0.0141	10, 11	0.0249	18, 19	0.0021	26, 27	0.0024
3, 4	0.0157	11, 12	0.0199	19, 20	0.0019	27, 28	0.0027
4, 5	0.0177	12, 13	0.0134	20, 21	0.0019	28, 29	0.0029
5, 6	0.0205	13, 14	0.0085	21, 22	0.0018	29, 30	0.0032
6, 7	0.0239	14, 15	0.0055	22, 23	0.0019	30, 31	0.0035
7, 8	0.0272	15, 16	0.0038	23, 24	0.0020	31, 32	0.0039
8, 9	0.0287	16, 17	0.0029	24, 25	0.0021	32, 33	0.0044

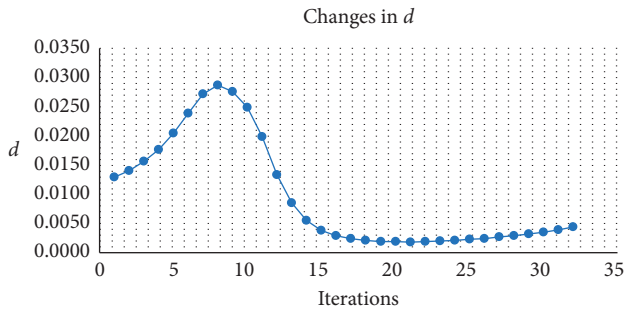


FIGURE 2: Changes in $d_{i,j}$.

the iteration discontinues when $\|V^{\text{present}} - V^{\text{previous}}\| \leq \varepsilon$, and a more suitable center satisfying the conditions may appear after n iterations with smaller values of $d_{i,j}$.

In this paper, the iteration ceased when $J(U, V) \leq \alpha$ and $s \geq \beta$, where the additional condition $s \geq \beta$ ensured an appropriate distance between different classes and made it easier to attain appropriate values for α and β .

6. Conclusion

Complex huge group clustering is the basis for the effective distribution of service resources and group coordination; nevertheless, traditional FCM and its improved versions are incapable of processing numerous data points to be clustered. In this case, the deletion of data points was studied in the current paper by using a graph-based clustering algorithm, adaptive clustering algorithm, and Gaussian kernel clustering algorithm. Meanwhile, a new Gaussian algorithm was proposed to present both the inner-class distance and between-class distance, which the objective function fails to do.

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.

References

- [1] W. Li and J. Li, "Improvement of semi-supervised kernel clustering algorithm based on multi-factor stock selection," *Statistics and Information Forum*, vol. 33, no. 3, pp. 30–36, 2018.
- [2] L. Zhen, "Tactical berth allocation under uncertainty," *European Journal of Operational Research*, vol. 247, no. 3, pp. 928–944, 2015.
- [3] Z. Zhang, J. Lin, and R. Miao, "Hesitant fuzzy language condensed hierarchical clustering algorithm and its application," *Statistics and Decision*, vol. 21, 2019.
- [4] L. Zhen, "Modeling of yard congestion and optimization of yard template in container ports," *Transportation Research Part B Methodological*, vol. 90, 2016.
- [5] Y. Duan and G. Wang, "A FCM clustering algorithm based on polygonal fuzzy numbers to describe multiple attribute index information," *Systems Engineering-Theory and Practice*, vol. 12, pp. 3220–3228, 2016.
- [6] D. Ying, X. Ying, and J. Ye, "A novel clustering algorithm based on graph theory," *Computer Engineering and Application*, vol. 45, no. 3, pp. 47–50, 2009.
- [7] Y. Xue and X. Sha, "On gray prediction model based on an improved FCM algorithm," *Statistics and Decision*, vol. 09, pp. 29–32, 2017.
- [8] H. Zhang and J. Wang, "Improved fuzzy C-means clustering algorithm based on selecting initial clustering center," *Computer Science*, vol. 36, no. 6, pp. 206–208, 2009.
- [9] L. Chiu Stephen, "A cluster estimation method with extension to fuzzy model identification," in *Proceedings of the IEEE Conference on Control Applications Part 2*, pp. 1240–1245, Orlando, FL, USA, August 1994.
- [10] L. E. Agustin-Blas, S. Salcedo-Sanz, S. Jimenez-Fernandez, L. Carro-Calvo, J. Del Ser, and J. A. Portilla-Figueras, "A new grouping genetic algorithm for clustering problems," *Expert Systems with Application*, vol. 39, pp. 9695–9703, 2012.
- [11] H. Zhang and J. Lu, "Semi-supervised fuzzy clustering: a kernel-based approach," *Knowledge-Based Systems*, vol. 22, no. 6, pp. 477–481, 2009.
- [12] S. Ramathilagam and Y.-M. Huang, "Extended Gaussian kernel version of fuzzy c-means in the problem of data analyzing," *Expert Systems with Applications*, vol. 38, no. 4, pp. 3793–3805, 2011.
- [13] X. Qian and L. Yao, "Extended incremental fuzzy clustering algorithm for sparse high-dimensional big data," *Computer Engineering*, vol. 45, no. 6, pp. 75–81, 2019.
- [14] X. Niu and K. She, "Research on fast parallel clustering algorithm for large scale data," *Computer Science*, vol. 39, no. 1, pp. 134–137, 2012.
- [15] W. Hu, "Improved hierarchical K-means clustering algorithm," *Computer Engineering and Applications*, vol. 49, no. 2, pp. 157–159, 2013.
- [16] J. Liu, Y. Jiang, J. Wang, Z. Deng, and S. Wang, "Single pass bayesian fuzzy clustering," *Journal of Software*, vol. 29, no. 9, pp. 2664–2680, 2018.
- [17] S. Zhou, W. Xu, and C. Tian, "Data-weighted fuzzy C-means clustering algorithm," *Systems Engineering and Electronics*, vol. 36, no. 11, pp. 2314–2319, 2014.
- [18] X. Chen and R. Liu, "Improved clustering algorithm and its application in complex huge group decision-making," *Systems Engineering and Electronics*, vol. 28, no. 11, pp. 1695–1699, 2006.
- [19] F. Xie, C. Lei, F. Li, D. Huang, and J. Yang, "Unsupervised hyperspectral feature selection based on fuzzy C-means and grey wolf optimizer," *International Journal of Remote Sensing*, vol. 40, no. 9, pp. 3344–3367, 2019.
- [20] H. Huang, K. Chang, H. Yu et al., "Research on adaptive entropy weight fuzzy C-means clustering algorithm," *Systems Engineering-Theory & Practice*, vol. 36, no. 1, pp. 219–223, 2016.
- [21] X. Xu and Y. Fan, "Improved ants-clustering algorithm and its application in multi-attribute large group decision making," *Systems Engineering and Electronics*, vol. 33, no. 2, pp. 346–349, 2011.

Research Article

Operation Efficiency Evaluation of the China-Europe Freight Train Based on Grey Cross-Efficiency DEA

Jun Zhang,^{1,2} Jianpeng Chang,^{1,2} Ping Lin,² Minzi Song ,³ and Yanqiu Gong⁴

¹Enterprise Management Research Center, Chongqing Technology and Business University, Chongqing 400067, China

²School of Management Science and Engineering, Chongqing Technology and Business University, Chongqing 400067, China

³School of Foreign Language, Chongqing Technology and Business University, Chongqing 400067, China

⁴College of Business Administration, Chongqing Technology and Business University, Chongqing 400067, China

Correspondence should be addressed to Minzi Song; songminzi@email.ctbu.edu.cn

Received 9 July 2020; Revised 28 August 2020; Accepted 16 September 2020; Published 31 October 2020

Academic Editor: Tingsong Wang

Copyright © 2020 Jun Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The China-Europe Freight Train (CEFT) serves as an important carrier and platform for international economic cooperation and international trade circulation between China and Europe. Since it worked, its actual operation and development have been affected by many factors, but the level of its actual operating efficiency and the main affecting factors of CEFT have been difficult to find, which has severely limited its sustainable development. Therefore, this paper scientifically selected the operation efficiency evaluation indicator system of CEFT and combined grey system theory, cross-efficiency method, and DEA to construct a new DEA evaluation model based on grey cross-efficiency, which can not only overcome the problem of ignoring the relative importance ratings of the evaluation indicator in the general DEA evaluation model and the traditional cross-efficiency DEA evaluation model but also more accurately evaluate the actual operation efficiency of CEFT. At the same time, based on the actual operating data of CEFT from 2011 to 2018 and the above new evaluation models, the CEFT's operation efficiency was evaluated and tested by examples, showing that on the one hand, the grey cross-efficiency DEA evaluation model can more accurately evaluate the actual operation efficiency of CEFT than other traditional evaluation models; on the other hand, it is found that the "overseas cities," "operating lines," and "entry-exit nodes" are currently the main factors that limit the actual operation efficiency of CEFT and indicating improvement direction for the future efficient and sustainable development of CEFT.

1. Introduction

On March 19, 2011, the "Chongqing-Sinkiang-Europe International Railway" container freight train from Chongqing, China, to Duisburg, Germany, was launched, marking the official opening of the new railway freight model between China and Europe, the China-Europe Freight Train (CEFT). The CEFT connected the two strategic spaces of the Silk Road Economic Belt and the Yangtze River Economic Belt, and it served as an important carrier and platform for the country to promote the "Belt and Road" initiative. Moreover, it also broke the foreign trade pattern that focused on the eastern coastal cities after China's reform and opening up, making the Midwest region a bridgehead that opens to the west. Since

it began to work, the CEFT has not only achieved major breakthroughs in terms of scale, coverage, and freight category but also formed a relatively clear operation model and a relatively stable operation pattern. In 2019, the CEFT ran 8,225 trains and delivered 725,000 TEUs, with a comprehensive reload rate of 94%, and has embarked on a normalized and large-scale development path. However, after several years of disorderly development, a series of problems on the operation of the CEFT have been revealed, such as high costs, disorderly competition, insufficient supply and demand connection, and low customs clearance efficiency. Therefore, the purpose of this paper is to conduct a comprehensive scientific evaluation of the operation efficiency of the CEFT, explore the essential causes, and propose targeted solutions.

Until now, no scholars have studied the operation efficiency of the CEFT, but some scholars have discussed the operation efficiency of railways. Existing studies have either focused on the operation efficiency of railway passenger transport, or on rail freight transport, or on the comprehensive efficiency of the railway system. Moreover, existing studies mostly use DEA or its improved forms to measure the operation efficiency of the railway system, mainly because specific production function forms are not needed to be set in advance. This paper used the DEA method to evaluate the annual operation efficiency of the CEFT since its inception and explore its current operation status and its shortcomings. Even though the cross-efficiency DEA model can achieve the ranking of alternatives, it cannot distinguish the relative size of the weights of inputs and outputs. Therefore, to solve this problem, this paper introduced the grey system theory, proposed a cross-efficiency DEA model based on grey correlation, and fixed the proportional relationship between input or output index weights, carrying out the efficiency evaluation of the CEFT.

2. Research Status

At present, there have been many researches on the efficiency of railway operations, and the topics covered are broad. They can be further divided into two categories according to the research theme: one is to explore the impact of railway operation efficiency on the economic and social environment, such as the impact of railway transport on green innovation efficiency [1] and the impact on energy and environmental protection efficiency [2, 3], while the other is to explore how to improve railway operation efficiency, namely, attempt to identify the problems in the development of railway operations by evaluating the efficiency of railway operations, and then put forward targeted recommendations [4, 5], and in this paper, the second category has been drawn attention. Initially, relevant studies treated railway passenger transport and freight transport equally and explored the comprehensive railway operation efficiency [6, 7]. However, due to the significant differences between the railway passenger transport and freight transport in the process of operation, Yu and Lin [8, 9] and Li [10] pointed out that the efficiency of the railway passenger transport and railway freight transport should be studied within a unified framework, respectively. Therefore, the research on how to improve the efficiency of railway operation can be further divided into three types, one is the research on the operation efficiency of the railway passenger transport, another is the research on the operation efficiency of the railway freight transport, and the last is the research on the comprehensive operation efficiency of the railway passenger and freight transport. The network DEA method was employed to measure the railway passenger transport operation efficiency, freight operation efficiency, service efficiency, and technical efficiency of the 20 railway systems in 2002 and further compared the

network DEA to the traditional DEA, pointing out that the network DEA owned greater advantages than the traditional DEA in the railway and could provide more information about improvements in railway efficiency [9]. Besides, Sameni et al. and Preston [11] used DEA to measure the operational efficiency of the British railway passenger transport and explored the utilization of railway capacity. Friebe et al. [12] used PMF (Production Frontier Model) to study the reform efficiency of the EU railway passenger transport system. In another paper, Kutlar et al. [13] used CCR (the abbreviation of its developers' initials, i.e., Charnes, Cooper, and Rhodes [14]) and BCC (the abbreviation of its developers' initials, i.e. Banker, Charnes, and Cooper [15]), respectively, to evaluate the technical efficiency and resource allocation efficiency of 31 railway operation companies in the world from 2000 to 2009. The two methods obtained results with obvious differences, mainly because the BBC considered the factor of scale returns compared to CCR.

Many methods can be applied to study the efficiency of railway operation and can be divided into the following four categories according to whether boundaries and parameters have been set [10]: first is the parameters-nonboundary method represented by Tobit and OLS, which is applicable to the factors analysis affecting efficiency; the second is the nonparameter and non-boundary method represented by the total productivity index, which has strict restrictions on the price index of the input and output; the third is the parameter-boundary method, which can be further divided into random boundary method and fixed boundary parameter method; the former is more sensitive to the assumption of the distribution method of random variables, while the latter choosing different production functions may lead to different results; the fourth is the nonparametric-boundary method represented by data envelopment analysis (DEA). DEA is a tool for evaluating the relative efficiency of homogeneous decision-making units with multiple inputs and outputs. Compared with other efficiency evaluation tools, it does not require the setting of specific production function forms in advance, nor does it require to determine the relevant input and output weights, only input and output data are required, and no other processing is required for the data. Because of its outstanding advantages, DEA has received extensive attention in the academic and practical fields since its introduction. In theory, DEA was first proposed by Charnes et al. [14] in 1978 and the original DEA model was called the CCR model, which aimed to quantify the relative efficiency of all decision-making units by constructing an effective production frontier. Then, in 1984, Banker et al. [15] proposed the BCC model based on CCR, which considered variable scale efficiency. Since then, DEA models with different characteristics and functions, such as the cross-efficiency model [16], confidence interval model [17], cone scale model [18], and super-efficiency model [19], have been proposed. In terms of application, DEA and its improved forms have been

widely used in efficiency research in various fields, including transport system [20], education system [21], construction field [22], tourism industry [23], environmental protection field [24], and hospital system [25].

3. Selection of Evaluation Indicators for Operation Efficiency of CEFT

This section aims to build a scientific, objective, and reasonable operation efficiency evaluation indicator system for CEFT.

The construction of the indicator system needs to strictly follow the principles of representativeness, completeness, specificity, scientificity, accessibility, and independence. We divided the construction process of the indicator system into two phases, namely, the preliminary screening phase and the key screening phase.

3.1. Phase 1: the Preliminary Indicator Screening. This paper used the analytical method to carry out the first screening of indicators. First of all, the theoretical knowledge related to the evaluation of operation efficiency has been fully familiarized with, and concurrently the background knowledge, development status, construction results, etc., of the operation of CEFT have been carefully studied. On the basis of a certain understanding of the theory and actual situation, by referring to relevant literature and documents, the development connotation of CEFT has been analyzed and determined; finally, the appeals of interest subjects have been analyzed from the perspectives of three different interest subjects. The freight owners, freight forwarding, and operating companies of CEFT have different evaluation on operation efficiency. For example, from the perspective of the freight owners, the main considerations are speed and time efficiency and transport costs, while for the freight forwarding, the primary considerations are the scale of transport and freight security, and the economic and trade contribution brought by the operation of the train is what the operating companies focus on. Through the above analysis, the operation efficiency of CEFT can be divided into six levels: scale range, speed efficiency, transport efficiency, transport cost, economic and trade contribution, and freight traffic safety.

- (i) The range of scale is an important indicator for evaluating the development volume and scale efficiency of CEFT and is also the basis for judging the different stages of development of CEFT. Within a certain period, increasing the scale of CEFT is the key way to reduce the operation costs.
- (ii) The speed efficiency is a significant manifestation of the high-quality operation of CEFT. Increasing the average delivery speed and ensuring the logistics efficiency of the whole process are also an effective way to improve the operation quality of CEFT.
- (iii) The transport efficiency indicator describes the level of transport capacity and nodes smoothness of the CEFT, the level of transport, and organization such as the scheduling, negotiation, and implementation,

and the overall efficiency of the use of transport resources such as fixed facilities and mobile equipment.

- (iv) The transport cost is the transport expenditure shared by CEFT in order to complete the transport workload, and it is an important basis for CEFT to set the price of the freight transport and the main factor for disorderly competition of freight sources.
- (v) The economic and trade contribution reflects the relations between the logistics development of CEFT and China-EU economic and trade exchanges, showing the contribution of CEFT to the development of Sino-European trade.
- (vi) The freight traffic safety statistics, such as the frequency of operational safety accidents, freight damage and freight theft, and international insurance business volume of CEFT are important indicators that can reflect the CEFT's overall freight safety guarantee and risk prevention and control capabilities.

After the evaluation indicator layer was determined, the preliminary indicator layer that met the CEFT efficiency evaluation was set according to the construction principles of the CEFT efficiency evaluation indicator. And then, after considering relevant experts' advisory opinions, the representative subindicators have been identified, forming a preliminary evaluation indicator system with 6 indicator layers and 33 subindicator layers, which are presented in Table 1.

3.2. Phase 2: the Key Indicator Screening. In the preliminary indicator system, there exist overlaps between subindicators, easily leading to distorted evaluation results. Therefore, it was necessary to further screen the 33 subindicators.

In this phase, we adopt the method of expert survey to screen the subindicators. Firstly, we selected 60 experts to conduct a questionnaire. The selecting process is presented as follows.

- (i) Step 1: defining the aims of selecting experts. In this case, the aim is to assess the importance of subindicators which are used for evaluating the operation efficiency of CEFT.
- (ii) Step 2: specifying the criteria of selection experts. In this case, the criteria include experience, knowledge, willingness, time, and independence.
- (iii) Step 3: finding expert sources. In this case, we select experts from railway administrations, logistic enterprises, governments, research institutions, and universities.
- (iv) Step 4: selecting alternatives. In this case, we identify 60 alternative experts based on the criteria specified in Step 2.

In order to ensure the professionalism, experts with a higher degree of knowledge on logistics and railway operation, government leaders with at least 10 years of related

TABLE 1: Preliminary indicator system for operation efficiency evaluation of CEFT.

Indicator layer	Subindicator layer	Indicator description
Scale range	Annual operation trains	The number of annual operation trains
	Annual outbound trains	The number of outbound trains in one year
	Annual inbound trains	The number of inbound trains in one year
	Annual containers	The number of all containers actually carried by the CEFT in one year
	Annual outbound containers	The total number of containers actually carried in the outbound CEFT in one year
	Annual inbound containers	The total number of containers actually carried in the inbound CEFT in one year
	Train return ratio	The ratio of the returned trips to the total trips, reflecting the balance of the two-way trips between China and Europe
	Ratio of returned containers	The ratio of the returned containers to the total number of freight containers transported by the train, reflecting the organization of the returned freight of CEFT
	Domestic cities	The number of cities with entry-exit nodes of CEFT in China
Overseas cities	The number of overseas cities with entry-exit nodes of CEFT	
Speed efficiency	Container conversion	The actual number of containers carried by CEFT is converted into the number of trains in accordance with the standard of carrying 41 containers per train.
	Average time	The "station-to-station" transport time of CEFT, including transit time plus node stay time
Transport efficiency	Average speed	The ratio of CEFT transport mileage to transport time
	Operating lines	The number of CEFT running lines among China, Europe, and countries along the belt and road, including trunk lines and branch lines
	Entry-exit nodes	The number of entry-exit nodes of CEFT in China
	Capacity utilization	The ratio of actual train operation to the maximum train operation
	Average static load rate	The average of the static load of the loaded containers, reflecting the degree of utilization of the freight transport capacity provided by the CEFT to the freight railways
	Full-capacity rate	The ratio of the loaded containers with full capacity to the number of all loaded containers
Operation cost	Fulfilled rate	The ratio of the actual number of CEFT to the number of scheduled CEFT
	Loaded container rate	The ratio of the number of loaded containers to the total number of containers
Economic and trade	Average freight charges of operation	The cost of transporting 1 container for 1 km. The average freight charges of the train = (domestic freight + foreign freight)/total mileage
	Government subsidy	Financial support by the local government, referring to the benchmark of freight of each train
Freight traffic safety	Total value of annual freight	The value of importing-exporting freights on CEFT, reflecting the status and role of CEFT in China-EU economic and trade cooperation
	Total value of the outbound freight	The value of exported freights on CEFT in one year
	Total value of the inbound freight	The value of imported freights on CEFT in one year
	Unit TEU value	The average value per TEU of freight transported by CEFT, indicating the segmentation and target positioning of the CEFT logistics market
	Outbound unit TEU value	The average value per TEU of freights transported by outbound CEFT, indicating the segmentation and target positioning of the CEFT logistics market
	Inbound unit TEU value	The average value per TEU of freights transported by inbound CEFT, indicating the segmentation and target positioning of the CEFT logistics market
	Ratio of key categories China-EU total trade	The ratio of the number of TEU of the top 5 categories to the total number of TEU The total international trade between China and Europe
Freight traffic safety	Safety accidents	The number of accidents that affected the normal transport of CEFT due to reasons such as concealing the name of the cargo, overweight, poor loading and reinforcement, and damaged containers
	Freight damage and theft	The number of incidents of theft, damage, etc.
	Insurance claims business	The purchase of CEFT full-freight insurance and compensation for the train freight, including the number of insurance business, the proportion of insurance coverage, and the number of claims

experiences, and business executives with at least 10 years of related experiences are selected as the experts to be surveyed.

After 60 experts are selected, we design a questionnaire for assessing the importance of subindicators presented in Table 1. The questionnaire is given in Table 2. Then, we distribute online questionnaires to them through e-mail.

After experts finish them, we collect them automatically, and then, based on the quantitative statistical results of the questionnaire and the availability of data, we obtain the final CEFT's efficiency evaluation indicator system, which includes 8 key subindicators. The specific results are shown in Table 3.

TABLE 2: Questionnaire of the importance ratings of subindicators of evaluating CEFT’s operation efficiency.

Indicator layer	Subindicator layer	Extremely important	Very important	Important	General important	Unimportant
Scale range	Annual operation trains					
	Annual outbound trains					
	Annual inbound trains					
	Annual containers					
	Annual outbound containers					
	Annual inbound containers					
	Train return ratio					
	Ratio of returned containers					
	Domestic cities					
	Overseas cities					
Speed efficiency	Average time					
	Average speed					
Transport efficiency	Operating lines					
	Entry-exit nodes					
	Capacity utilization					
	Average static load rate					
	Full-capacity rate					
	Fulfilled rate					
Operation cost	Loaded container rate					
	Average freight charges of operation					
Economic and trade contribution	Government subsidy					
	Total value of annual freight					
	Total value of the outbound freight					
	Total value of the inbound freight					
	Unit TEU value					
	Outbound unit TEU value					
	Inbound unit TEU value					
	Ratio of key categories					
China-EU total trade						
Freight traffic safety	Safety accidents					
	Freight damage and theft					
	Insurance claims business					

TABLE 3: Evaluation indicator system of CEFT’s overall operation efficiency.

Indicator layer	Subindicator layer	Indicator description
Scale range	Annual operation trains	The number of annual operation trains
	Overseas cities	The number of overseas cities with entry-exit nodes of CEFT
	Domestic cities	The number of cities with entry-exit nodes of CEFT in China
	Annual containers	The number of all containers actually carried by the CEFT in one year
Transport efficiency	Train return ratio	The ratio of the returned trips to the total trips, reflecting the balance of the two-way trips between China and Europe
	Operating lines	The number of CEFT running lines among China, Europe, and countries along the belt and road, including trunk lines and branch lines
	Entry-exit nodes	The number of entry-exit nodes of CEFT in China
Economic and trade contribution	Total value of annual freight	The value of importing and exporting freights on CEFT, reflecting the status and role of CEFT in China-EU economic and trade cooperation

In order to meet the demand for the evaluation indicator in the DEA evaluation model, the evaluation indicator system of CEFT’s overall operation efficiency was further classified based

on the input-output perspective, choosing the annual operation trains (number), overseas cities (number), domestic cities (number), operating lines (number), and entry-exit nodes

TABLE 4: Input and output indicator system on efficiency evaluation of CEFT's overall operation efficiency.

Dimension	Indicator	Indicator description
Inputs	Annual operation trains, X_1	The number of annual operation trains
	Overseas cities, X_2	The number of overseas cities with entry-exit nodes of CEFT
	Domestic cities, X_3	The number of cities with entry-exit nodes of CEFT in China
	Operating lines, X_4	The number of CEFT operating lines among China, Europe, and countries along the belt and road, including trunk lines and branch lines
	Entry-exit nodes, X_5	The number of entry-exit nodes of CEFT in China
Outputs	Annual containers, Y_1	The number of all containers actually carried by the CEFT in one year
	Total value of annual freights, Y_2	The value of importing and exporting freights on CEFT, reflecting the status and role of CEFT in China-EU economic and trade cooperation
	Train return ratio, Y_3	The ratio of the returned trips to the total trips, reflecting the balance of the two-way trips between China and Europe

(number) as input index X_1, X_2, X_3, X_4, X_5 , respectively and choosing the annual containers (standard container), total value of annual freights (100 million US dollars), and train return ratio (%) as output index Y_1, Y_2, Y_3 . The input and output indicator system on efficiency evaluation of CEFT's overall operation efficiency is shown in Table 4.

4. Design of Evaluation Method for CEFT's Operation Efficiency

In this paper, we utilized a novel cross-efficiency DEA model embedded with grey correlation analysis to evaluate the operation efficiency of CEFT.

4.1. CCR Model. CCR is the most basic DEA model, which is developed under the assumption that returns to scale is fixed. Suppose there are n decision-making units (DMUs), which are depicted as $\{DMU_j | j = 1, 2, \dots, n\}$. For each DMU_j , m input items and s output items are selected, which are represented as $\{x_{ij} | i = 1, 2, \dots, m\}$ and $\{y_{rj} | r = 1, 2, \dots, s\}$, respectively. Besides, let w_i be the weight assigned to the i^{th} input item and μ_r the weight of r^{th} output item, $i = 1, 2, \dots, m$ and $r = 1, 2, \dots, s$. Then, the efficiency of DMU_j , which can be denoted as h_j , can be determined by the following equation:

$$h_j = \frac{\sum_{r=1}^s \mu_r y_{rj}}{\sum_{i=1}^m \omega_i x_{ij}}. \quad (1)$$

In the CCR model, the aim is to maximize the efficiency of each DMU denoted as h_d under the conditions that the efficiency of any DMU is not greater than 1, and the weights of all inputs and outputs must be greater than 0. The model can be depicted as follows:

$$\max \frac{\sum_{r=1}^s \mu_r y_{rd}}{\sum_{i=1}^m \omega_i x_{id}} = h_d,$$

$$\text{s.t.} \begin{cases} \sum_{r=1}^s \mu_r y_{rj} \leq 1, & j = 1, 2, \dots, n, \\ \sum_{i=1}^m \omega_i x_{ij} \end{cases}$$

$$\omega_i \geq 0, \mu_r \geq 0. \quad (2)$$

The above model can be further transformed to a linear programming model, which is presented as follows:

$$\max \sum_{r=1}^s \mu_r y_{rd} = h_d$$

$$\text{s.t.} \begin{cases} \frac{\sum_{r=1}^s \mu_r y_{rj}}{\sum_{i=1}^m \omega_i x_{ij}} \leq 1, & j = 1, 2, \dots, n \\ \sum_{i=1}^m \omega_i x_{id} = 1 \\ \omega_i \geq 0, \mu_r \geq 0, & i = 1, 2, \dots, m, r = 1, 2, \dots, s. \end{cases} \quad (3)$$

The CCR model seeks a set of values for input and output weights to maximize h_d . Solving the CCR model, if $h_d = 1$, then DMU_d is weakly efficient; if $h_d = 1$ and $\omega_i \geq 0, \mu_j \geq 0$, then DMU_d is efficient; if $h_d < 1$, then DMU_d is inefficient.

4.2. Cross-Efficiency DEA. Cross-efficiency DEA, proposed by Sexton et al., is an extension to the traditional DEA model. Compared with the CCR model, the cross-efficiency DEA model is developed based on the combination of peer-appraisal and self-appraisal and can provide a ranking among the DMUs.

Let DMU_{j_0} be the target DMU. Using model (3), we can obtain the optimal values including the weights of input items and output items, which are represented as $\{\mu_{r_{j_0}}^* | r = 1, 2, \dots, s\}$ and $\{\omega_{i_{j_0}}^* | i = 1, 2, \dots, m\}$,

respectively. Then, the cross-efficiency of DMU_j can be determined by the following equation

$$E_{j_0j} = \frac{\sum_{r=1}^s \mu_r^* \gamma_{rj}}{\sum_{i=1}^m \omega_i^* x_{ij}}, \quad (4)$$

in which $j_0, j = 1, 2, \dots, n$. Using the above equation, we can construct a cross-efficiency matrix, which is depicted in Table 5.

Then, we can calculate the value of average cross-efficiency of DMU_j:

$$E_j = \frac{1}{n} \sum_{j_0}^n E_{j_0j}, \quad j = 1, 2, \dots, n. \quad (5)$$

Then, we can rank all the DMUs based on the values of their cross-efficiency $E_j (j = 1, 2, \dots, n)$. Besides, we can further compare the efficiency E_{jj} obtained by self-appraisal with 1 to determine whether or not DMU_j is efficient. If $E_{jj} = 1$, then DMU_j is efficient; Else, DMU_j is inefficient.

4.3. Cross-Efficiency DEA Embedded with Grey Correlation Analysis. This paper incorporated grey correlation analysis into the traditional cross-efficiency DEA model to enable the novel model to be solved under the condition that the weights of inputs or outputs had fixed relative values.

Firstly, the weights of input and output items are determined by grey correlation analysis. Taking determining the weights of input items as an example, we let the reference values denoted as $X_0 = \{X_0(1), X_0(2), \dots, X_0(m)\}$ and let the input values of the target DMU, DMU_j, as $\{x_{ij} | i = 1, 2, \dots, m\}$, $j = 1, 2, \dots, n$. Then, the correlation coefficient between input values of DMU_j and reference values can be determined as follows:

$$r_{ij} = \frac{\Delta \min + \zeta \Delta \max}{|X_0(i) - x_{ij}| + \zeta \Delta \max}, \quad (6)$$

where $\zeta \in [0, 1]$ is the resolution coefficient and ζ typically takes the value of 0.5. In addition, $\Delta \min = \min_j \min_i |X_0(i) - x_{ij}|$ and $\Delta \max = \max_j \max_i |X_0(i) - x_{ij}|$.

Then, for target DMU DMU_j ($j = 1, 2, \dots, n$), we can obtain the correlation weights of inputs $\{\tau_{1j}, \tau_{2j}, \dots, \tau_{mj}\}$, in which τ_i is determined as follows:

$$\tau_{ij} = \frac{r_{ij}}{\sum_{i=1}^m r_{ij}}, \quad i = 1, 2, \dots, m. \quad (7)$$

Similarly, we can calculate the correlation weights of outputs, which are denoted as $\{\pi_{1j}, \pi_{2j}, \dots, \pi_{sj}\}$.

Then, we extract the relative size relationship from the obtained weights as the conditions and develop a novel cross-efficiency DEA model, which can be depicted as follows:

TABLE 5: Cross-efficiency matrix for DMUs.

Target DMU	DMU			
	1	2	...	n
1	E_{11}	E_{12}	...	E_{1n}
2	E_{21}	E_{22}	...	E_{2n}
3	E_{31}	E_{32}	...	E_{3n}
...
n	E_{n1}	E_{n2}	...	E_{nn}
Average cross-efficiency	E_1	E_2	...	E_n

$$\begin{aligned} \max \quad & \sum_{r=1}^s \mu_r \gamma_{rd} = h_d \\ \text{s.t.} \quad & \begin{cases} \frac{\sum_{r=1}^s \mu_r \gamma_{rj}}{\sum_{i=1}^m \omega_i x_{ij}} \leq 1, & j = 1, 2, \dots, n \\ \sum_{i=1}^m \omega_i x_{id} = 1 \\ \omega_i = \gamma \tau_{id}, & i = 1, 2, \dots, m \\ \mu_r = \sigma \pi_{rd}, & r = 1, 2, \dots, s \\ \gamma \geq 0, \sigma \geq 0, \end{cases} \end{aligned} \quad (8)$$

where γ and σ are the independent variables and ω_i and μ_r are the interdependent variables. Solving model (8), we can obtain the efficient values of DMU_d, $d = 1, 2, \dots, n$.

5. Empirical Research

5.1. Data Sources. The paper collected the data from the official website of China-Europe Freight Train and other official sources. The data of input and output indicators are presented in Table 6.

5.2. Data Analysis. Firstly, we should normalize the collected data. We could classify indicators into two classes, i.e., cost indicators and benefit indicators.

Taking the values of input indicators as an example, for cost indicators, the normalized value \bar{x}_{ij} could be determined by the following formula:

$$\bar{x}_{ij} = \frac{\max_j x_{ij} - x_{ij}}{\max_j x_{ij} - \min_j x_{ij}}. \quad (9)$$

For benefit indicators, the normalized value \bar{x}_{ij} was determined as follows:

$$\bar{x}_{ij} = \frac{x_{ij} - \min_j x_{ij}}{\max_j x_{ij} - \min_j x_{ij}}. \quad (10)$$

Following the above two equations, we could obtain the normalized data, which is presented in Table 7.

Then, we calculate the reference values (0.2000, 0.2000, 0.4000, 0.4000, 0.4643, 0.8000, 0.8769, and 1.0000) by

TABLE 6: Annual data of input and output indicators.

Year	Input					Output		
	Annual operation trains	Overseas cities	Domestic cities	Operating lines	Entry-exit nodes	Annual containers	Total value of annual freights	Train return ratio
2011	17	1	1	1	1	1404	0.7	0
2012	42	2	3	3	1	3674	1.84	0
2013	80	8	6	8	2	6960	4.61	0
2014	308	9	7	10	2	26078	13.04	9.1
2015	815	11	10	21	2	68902	34.45	32.5
2016	1702	13	15	26	4	145794	72.9	33.6
2017	3673	34	35	57	4	317930	145	34.7
2018	6363	49	56	65	5	533268	248.11	42.3

maximizing the data in each row of the input part of Table 7 to determine the weights of input criteria. Following the calculation process described in the above section, we could determine the correlation coefficients and correlation weights, which are presented in Table 8 and Figure 1.

In order to validate the advantages of the proposed model, we used the traditional DEA model and cross-efficiency DEA to calculate the efficiencies of CEFT, respectively. The results obtained by the three models are presented in Table 9 and Figure 2.

5.3. Analysis of the Effectiveness of Evaluation Methods. It could be seen from Table 9 and Figure 2 that for the traditional DEA model, it was difficult to completely distinguish all decision-making units and could not guarantee the uniqueness of the evaluation results. The efficiency value of the cross-efficiency DEA model has exceeded 0.7 in 2014, which was slightly inconsistent with the actual operating situation, and due to the limitation of the number of operating years of CEFT in the research object of this paper, there were fewer decision-making units and the crossefficiency DEA model has completely distinguished the decision-making units. However, with the increase in the number of operation trains, the cross-efficiency model was not suitable for the evaluation of the operating efficiency of CEFT. Compared with the traditional DEA model and the cross-efficiency DEA model, the efficiency of the crossefficiency evaluation method with the grey relation constraints was closer to the actual operating situation after considering the correlation of each indicator, and it was also easy to completely distinguish the decision-making units. Besides, the discriminating ability has been enhanced, and the evaluation results tended to be stable. Therefore, the results of the grey cross-efficiency DEA were used as examples for analysis. An efficiency value greater than 0.6 and less than 1 was called effective and if it was less than 0.6, it was called ineffective. Based on the results in Table 9, the CEFT from 2011 to 2018 can be divided into two echelons; the first echelon which included the years of 2011 and 2012 was ineffective; the second echelon which included the years of 2013, 2014, 2015, 2016, 2017, and 2018 was effective.

5.4. Analysis of Evaluation Results. Based on the evaluation results, the operating efficiency of CEFT in the first two years was ineffective, and the operating efficiency in the next six years was effective, of which the efficiency in 2017–2018 was better. The year of 2013 was a turning point in the operation of CEFT, because the operating efficiency began to be effective, and the efficiency value gradually increased and tended to be stable. Until 2016, the operating efficiency value of CEFT showed a downward trend, while the efficiency value began to rebound in 2017, and the efficiency value reached the highest in 2018.

On the whole, 2011 and 2012 were non-DEA effective evaluation units, and there were obvious problems in both inputs and outputs. 2011 and 2012 were the stages of exploration and development of CEFT where the number of operation trains was less than 50, the number of domestic cities was 3, the number of overseas cities was 2, the number of operating lines was 3, the number of entry-exit nodes was only one, the number of containers was less than 5,000 TEU, and there was no returned train. Compared with 2011, the growth rate of the number of overseas cities and operating lines reached 200% in 2012, and the inputs increase in key indicators resulted in a corresponding increase of 162% and 163% in the containers and the total value of freights. From the perspective of all input and output indicators, in the invalid decision-making units, insufficient outputs were entirely due to insufficient inputs; so in the next few years, the inputs of various indicators should be increased, especially the number of overseas cities and operating lines.

From 2013 to 2016, the period was a low-efficiency decision-making unit, and there were unreasonable structural problems in inputs and outputs. These four years served as the rapid development stage of CEFT. In 2014, the containers and the total value of freights far exceeded the sum of the previous three years, with growth rates as high as 275% and 183%, respectively. In addition, inbound trains began to operate this year and the main benefit was that the two input indicators of the number of operation trains and the number of operating lines were inputted at a growth rate of 285% and 125%, respectively. However, the other 2 indicators had insufficient input, such as entry-exit nodes and the overseas cities. Therefore, in 2014, the inputs in operation trains and operating lines had a further rise, while the inputs in entry-exit nodes and overseas cities relatively decreased. In 2015, the two indicators

TABLE 7: Normalized data.

Year	X_1	X_2	X_3	X_4	X_5	Y_1	Y_2	Y_3
2011	0.0027	0.0204	0.0179	0.0154	0.2000	0.0026	0.0028	0.0000
2012	0.0066	0.0408	0.0536	0.0462	0.2000	0.0069	0.0074	0.0000
2013	0.0126	0.1224	0.1250	0.1231	0.4000	0.0131	0.0186	0.0000
2014	0.0484	0.2653	0.2679	0.2769	0.4000	0.0489	0.0526	0.2151
2015	0.1281	0.4490	0.4643	0.4615	0.4000	0.1292	0.1388	0.7683
2016	0.2675	0.5714	0.5536	0.6000	0.8000	0.2734	0.2938	0.7943
2017	0.5772	0.6939	0.6250	0.8769	0.8000	0.5962	0.5844	0.8203
2018	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

TABLE 8: Correlation coefficients and correlation weights of indicators.

Indicators	X_1	X_2	X_3	X_4	X_5	Y_1	Y_2	Y_3
Correlation coefficients	0.5297	0.6816	0.6790	0.7449	0.9477	0.7425	0.7456	0.9892
Correlation weights	0.1478	0.1902	0.1895	0.2079	0.2645	0.2997	0.3010	0.3993

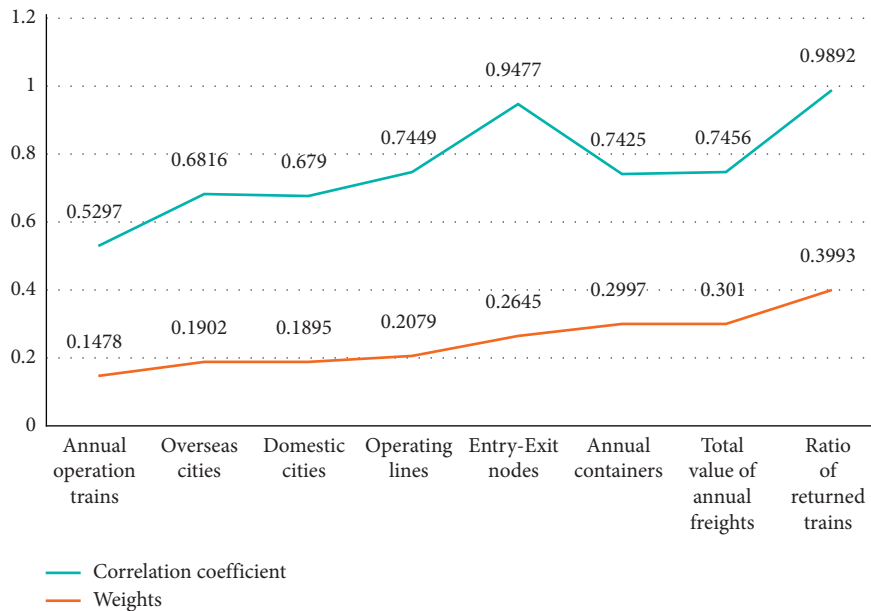


FIGURE 1: Correlation coefficients and correlation weights of input and output indicators.

TABLE 9: Efficiency values of each DMU.

Year	DEA model	Cross-efficiency DEA model	The propose dmodel
2011	0.9469	0.5581	0.5401
2012	1.0000	0.5882	0.5692
2013	1.0000	0.6405	0.6295
2014	0.9931	0.7134	0.6934
2015	1.0000	0.8066	0.7906
2016	1.0000	0.7705	0.7595
2017	1.0000	0.8224	0.8094
2018	1.0000	0.8305	0.8145

of operation trains and overseas cities maintained a relatively high growth, making the ratio of returned trains increase 23%, while entry-exit nodes were still insufficient. In 2016, the

efficiency value showed a downward trend. Although the growth rate of operation trains was still as high as 109% and the total number of operation trains exceeded 1500 which was twice the number of operation trains in 2015, the containers and the total value of freights decreased by 52% year-on-year, and the ratio of returned trains increased only 1%. This showed that at this stage, the operation trains had been inputted too much. Although the containers and the total value of freights have been improved to a certain extent, however, when looking at the ratio of returned trains which reflected the balanced degree of the two-way trains between China and Europe, there were great shortcomings. In terms of inputs in CEFT, simply increasing the number of operation trains would not improve the overall efficiency value. On the contrary, increasing the inputs of unnecessary trains would waste resources and reduce the efficiency value.

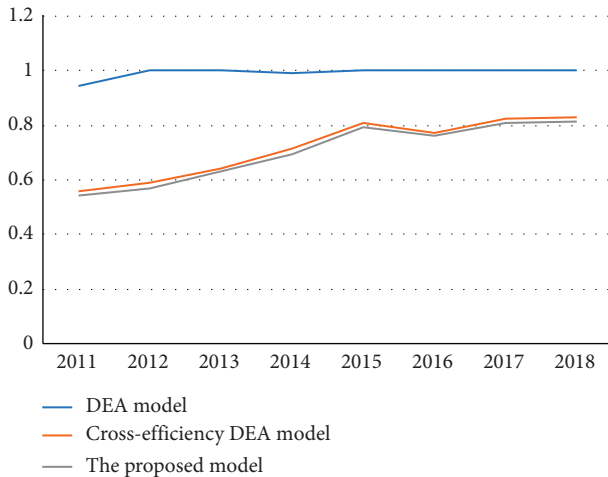


FIGURE 2: Comparison of efficiency values of inputs and outputs of the three models.

The period of 2017 to 2018 was a decision-making unit with better effectiveness, and the CEFT had entered the stage of improving quality and efficiency. In 2017, the operation trains exceeded the total of the previous six years' operation trains and the operating lines increased to 57 with a growth rate of 46%. Correspondingly, the growth rate of the containers also increased by 6%, and the ratio of returned trains also increased. Besides, the increase in the inputs in operating lines has expanded the radiation range of CEFT and enhanced the ability of CEFT to gather supplies. In 2018, the growth of inputs in the number of operation trains, overseas cities, domestic cities, and operating lines dropped sharply, and the growth rates fell by 43%, 117%, 73%, and 105%, respectively. However, under these circumstances, the total value of freights decreased by only 28%, while the ratio of returned trains increased by 19%, because of the fact that the Suifenhe railway port had already opened the CEFT, and the increase in entry-exit nodes had improved the efficiency of CEFT. The growth rate of the input indicators of various scales in 2018 tended to decrease, indicating that the input and output indicators of CEFT were more rational in structure and this period was in a transition period. The development of CEFT should change from simply pursuing the number of operation trains to improving the quality of trains, and it was not allowed to continue blindly increasing the number of operation trains. On the contrary, the three indicators could be viewed as the focus of future development of CEFT, and they were overseas cities, operating lines, and entry-exit nodes.

6. Recommendations for Future Efficient and Sustainable Development of China-Europe Freight Train

Based on the above comprehensive evaluation results of the CEFT's operation efficiency, countermeasures for improving CEFT's operation efficiency in the future will be presented as follows:

- (1) To expand CEFT overseas service coverage and increase the number of overseas cities that CEFT connects, CEFT should attach great importance to exploring European market and expanding the coverage of trains in European cities and appropriately increasing the number of overseas cities that CEFT can connect. Specifically, first is that CEFT can work closely with third-party international logistics companies, international freight forwarding companies, and foreign trade manufacturing companies in Europe to give full play to its faster transport than international ocean freight transport and lower cost than international air cargo transport, attracting more suitable sources of international railway transport with better quality and more accurate professional services. The second is to jointly formulate marketing plans and actively carry out logistics business promotion based on different regional characteristics along the CEFT lines and its radiating regions. The third is to steadily promote the layout of cities which CEFT can connect and freight hubs in Europe. In important node cities and cities where freight business is relatively concentrated, CEFT can cultivate and build a number of international railway transport hubs and transit distribution centers, warehouse distribution centers, overseas warehouses, etc., to improve the gathering and transport capacity of the nodes.
- (2) To optimize the domestic and international freight hubs of CEFT and moderately supplement the operation lines, CEFT should integrate, expand, and optimize the distribution of freights in both the Chinese market and European market while continuously expanding its service market and service scope in China and Europe. The details are as follows: first, it needs to moderately integrate and merge the distribution and supply of export-oriented freight sources in the Chinese market, rationally divide the service and radiation scope of the CEFT lines in China, and moderately source the freight and form some freight hubs with a certain size in China. The second is to properly integrate and merge the distribution and supply of import-oriented freight sources in the European market, to cooperate in depth with various international freight forwarding companies in Europe, and to reasonably divide the service and radiation range of the European and Chinese CEFT lines in the European regions so that it can form some European freight hubs with a certain scale. The third is to integrate and optimize all the existing CEFT lines, which means CEFT should moderately reduce train lines with low transport demand and low transport efficiency, moderately increase CEFT lines with relatively high transport demand and relatively high transportation efficiency, and moderately

expand or add some new lines to meet with new demands.

- (3) To integrate entry-exit node resources of CEFT and to enhance entry-exit nodes' passing capacity, since the establishment of international logistics transport entry-exit nodes is mainly affected by international transport infrastructure, customs, and other fixed facilities, it is difficult to add entry-exit nodes of CEFT on the Chinese border. Therefore, under the current circumstances where the number of entry-exit nodes listed by the CEFT is relatively limited, in order to effectively improve the passing capacity of nodes, it can only be turned to integrate the relevant resources of the existing nodes to improve customs clearance efficiency. At the same time, according to statistics, the average detention time of CEFT at border nodes currently accounts for about 30% of the total time. There are three main reasons. First is the delay in customs clearance due to documents and customs inspections; the second reason is when passing through the nodes, two railway track changes are required, and the efficiency of track change at each node is relatively low. Third, the operational capacity of customs clearance and inspection at border nodes are relatively inadequate, making each node a time lag in customs clearance and inspection. Therefore, on the one hand, it is necessary to coordinate Chinese and foreign railway transport enterprises, customs, and inspection and quarantine departments, systematically sort out, integrate, and simplify the entry-exit customs clearance process of CEFT, eliminate invalid customs clearance links, and improve inefficient customs clearance links and entry-exit customs clearance effectiveness. On the other hand, CEFT can integrate and optimize existing node resources, dig up and make effective use of idle resources, and moderately supplement or add some basic facilities such as customs clearance, inspection, and track changes at various nodes to effectively expand the nodes' customs passing capabilities.

7. Conclusion

This paper mainly adopted a new DEA model based on grey cross-efficiency to study the operation efficiency evaluation of CEFT, accurately identified the fundamental problems affecting the operation efficiency of CEFT, and put forward effective future sustainable development countermeasures. The main contributions of this paper were as follows.

Firstly, this paper used the grey correlation weight calculation method to accurately find the relative optimal weight ratio of the input and output indicator of the CEFT operation and then introduced this optimal ratio relationship as a limiting condition to the cross-efficiency DEA evaluation model to solve the problem of ignoring the relative size of the input and output indicator weights.

Secondly, the new grey cross-efficiency DEA evaluation model was applied to quantitatively and accurately evaluate the operation efficiency of CEFT. Through accurate evaluation and analysis, it found that the "overseas cities," "operating lines," and "entry-exit nodes" were currently the main influencing factors that restricted the actual operation efficiency of CEFT so that it further indicated the improvement direction for CEFT's future efficient and sustainable development.

However, there still exist some deficiencies waiting to be solved; for example, we do not take the operation efficiency of different trains and further do not account for the interdependent relationship of different trains, which are both our future research directions.

The questionnaire for evaluating the importance ratings of subindicators of assessing CEFT's operation efficiency is presented in Table 2.

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 has been financed by the National Social Sciences Funding Project (Grant no. 18BGL007); Chongqing Social Sciences Planning Major Project (Grant no. 2018TBWT-ZD07); Chongqing Municipal Education Commission Humanities and Social Sciences Research General Project (Grant no. 18SKGH069); Chongqing Science and Technology Bureau Technology Foresight and Institutional Innovation Project (Grant no. cstc2019jsyj-zzysbAX0047); Chongqing Doctoral Program of Social and Scientific Planning (Grant no. 2018BS79); and Scientific Research Start-Up Foundation of Chongqing Technology and Business University (Grant no. 1855016).

References

- [1] Y. Huang and Y. Wang, "How does high-speed railway affect green innovation efficiency? a perspective of innovation factor mobility," *Journal of Cleaner Production*, vol. 265, 2020.
- [2] M. Song, G. Zhang, W. Zeng, J. Liu, and K. Fang, "Railway transportation and environmental efficiency in China," *Transportation Research Part D-transport and Environment*, vol. 48, pp. 488–498, 2016.
- [3] Z. Liu, C.-X. Qin, and Y.-J. Zhang, "The energy-environment efficiency of road and railway sectors in China: evidence from the provincial level," *Ecological Indicators*, vol. 69, pp. 559–570, 2016.
- [4] A. Beck, H. Bente, and M. Schilling, "Railway efficiency: an overview and a look at opportunities for improvement," *International Transport Forum Discussion Paper*, Organisation for Economic Co-operation and Development (OECD), Paris, France, 2013.

- [5] Y. C. Lai, "Increasing railway efficiency and capacity through improved operations," *Control and Planning*, vol. 69, no. 5, 2008.
- [6] M. M. Movahedi, S. Saati, and A. R. Vahidi, "Iranian railway efficiency (1971-2004): an application of DEA," *International Journal of Contemporary Mathematical Sciences*, vol. 2, no. 32, pp. 1569–1579, 2007.
- [7] M. M. Movahedi, S. Y. Abtahi, and M. Motamedi, "Iran railway efficiency analysis, using DEA: an international comparison," *International Journal of Applied Operational Research-An Open Access Journal*, vol. 1, no. 1, pp. 1–7, 2011.
- [8] M.-M. Yu and E. T. J. Lin, "Efficiency and effectiveness in railway performance using a multi-activity network DEA model," *Omega*, vol. 36, no. 6, pp. 1005–1017, 2008.
- [9] M.-M. Yu, "Assessing the technical efficiency, service effectiveness, and technical effectiveness of the world's railways through NDEA analysis," *Transportation Research Part A: Policy and Practice*, vol. 42, no. 10, pp. 1283–1294, 2008.
- [10] L. Li, "Comprehensive efficiency evaluation of Chinese railway systems: based on double-stage and double-activity analysis framework," *Nankai Economic Studies*, vol. 5, pp. 95–110, 2010.
- [11] M. K. Sameni and J. Preston, "Value for railway capacity: assessing efficiency of operators in great Britain," *Transportation Research Record*, vol. 2289, no. 2289, pp. 134–144, 2012.
- [12] G. Friebel, M. Ivaldi, and C. Vibes, "Railway (De) Regulation: a European efficiency comparison," *Economica*, vol. 77, no. 305, pp. 77–91, 2010.
- [13] A. Kutlar, A. Kabasakal, and M. Sarikaya, "Determination of the efficiency of the world railway companies by method of DEA and comparison of their efficiency by Tobit analysis," *Quality & Quantity*, vol. 47, no. 6, pp. 3575–3602, 2013.
- [14] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units," *European Journal of Operational Research*, vol. 2, no. 6, pp. 429–444, 1978.
- [15] R. D. Banker, A. Charnes, and W. W. Cooper, "Some models for estimating technical and scale inefficiencies in data envelopment analysis," *Management Science*, vol. 30, no. 9, pp. 1078–1092, 1984.
- [16] T. R. Sexton, R. H. Silkman, and A. J. Hogan, "Data envelopment analysis: critique and extensions," *New Directions for Program Evaluation*, vol. 1986, no. 32, pp. 73–105, 1986.
- [17] R. G. Thompson, F. D. Singleton, R. M. Thrall, and B. A. Smith, "Comparative site evaluations for locating a high-energy physics lab in Texas," *Interfaces*, vol. 16, no. 6, pp. 35–49, 1986.
- [18] A. Charnes, W. W. Cooper, Q. L. Wei, and Z. M. Huang, "Cone ratio data envelopment analysis and multi-objective programming," *International Journal of Systems Science*, vol. 20, no. 7, pp. 1099–1118, 1989.
- [19] L. M. Seiford and J. Zhu, "Infeasibility of super-efficiency data envelopment analysis models," *INFOR: Information Systems and Operational Research*, vol. 37, no. 2, pp. 174–187, 1999.
- [20] R. Mahmoudi, A. Emrouznejad, S.-N. Shetab-Boushehri, and S. R. Hejazi, "The origins, development and future directions of data envelopment analysis approach in transportation systems," *Socio-Economic Planning Sciences*, vol. 69, p. 100672, 2020.
- [21] M. Torres-Samuel, C. L. Vásquez, M. Luna et al., "Performance of education and research in Latin American countries through data envelopment analysis (DEA)," *Procedia Computer Science*, vol. 170, pp. 1023–1028, 2020.
- [22] M. Nahangi, Y. Chen, and B. McCabe, "Safety-based efficiency evaluation of construction sites using data envelopment analysis (DEA)," *Safety Science*, vol. 113, pp. 382–388, 2019.
- [23] S. Chaabouni, "China's regional tourism efficiency: a two-stage double bootstrap data envelopment analysis," *Journal of Destination Marketing & Management*, vol. 11, pp. 183–191, 2019.
- [24] X. Liu, P. Guo, and S. Guo, "Assessing the eco-efficiency of a circular economy system in China's coal mining areas: emergy and data envelopment analysis," *Journal of Cleaner Production*, vol. 206, pp. 1101–1109, 2019.
- [25] S. Kohl, J. Schoenfelder, A. Fügner, and J. O. Brunner, "The use of Data Envelopment Analysis (DEA) in healthcare with a focus on hospitals," *Health Care Management Science*, vol. 22, no. 2, pp. 245–286, 2019.

Research Article

Fuzzy Principal Component Analysis Model on Evaluating Innovation Service Capability

Hui Wu and Xiao-min Gu 

School of Financial Technology, Shanghai Lixin University of Accounting and Finance, 995 Shangchuan Road, Pudong New Area, Shanghai 201209, China

Correspondence should be addressed to Xiao-min Gu; guxiaomin@lixin.edu.cn

Received 9 June 2020; Revised 23 June 2020; Accepted 26 July 2020; Published 30 September 2020

Academic Editor: Lu Zhen

Copyright © 2020 Hui Wu and Xiao-min Gu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article mainly evaluates the regional innovation service capacity through the TOPSIS method. Firstly, a regional collaborative innovation network is constructed and the Yangtze River Delta region is selected for analysis. Secondly, an evaluation index is constructed for innovation service capability, fuzzy principal component analysis is used to refine quantitative and qualitative index data of innovation service capability, and the index weight is calculated. Then the region of the Yangtze River Delta is selected and TOPSIS method is used to assist in the effective decision-making process of the evaluation of innovative service capabilities. Due to the large amount of data in this article, MATLAB programming is used. Finally, through the comparative analysis of the results, countermeasures and suggestions are put forward from the perspective of the improvement of collaborative innovation service capabilities.

1. Introduction

The development of regional collaborative innovation is conducive to promoting the rapid development of the regional economy, optimizing the efficiency of innovation resource allocation, transforming factor-driven to innovation-driven, and promoting the overall improvement of regional economic competitiveness. The collaborative innovation model with scientific structure and active communication is conducive to the establishment of a long-term and effective scientific and technological cooperation mechanism between regions. However, due to the imbalance of resource allocation, disparity in development levels, simple repeated construction, and disorderly vicious competition among regions, the development of regional collaborative innovation has seriously hindered the promotion of regional collaborative innovation.

Collaborative innovation theory is the theoretical basis of cross-regional collaborative innovation. After Haken proposed synergetics in 1976, different scholars elaborated collaborative innovation from different angles. Kim [1]

analysed the ways in which companies enhance international competitiveness by synergizing internal and external information and other innovative elements. Huber [2] believes that innovation promotes the spiral of organizational learning through process practice, so that the innovation ability in collaborative innovation is endogenous in the innovation elements. Radosevic [3] conducted research on the synergy between technological innovation and institutional change and believed that the key to determining the success of corporate innovation lies in the coordinated development of technology and many nontechnical elements. Power [4] believes that innovative ideas, processes, groups, and organizations are important elements of collaborative innovation. Morris [5] empirically studied the important impact of innovation factors on high-tech R&D through social networks and further enriched the means for achieving collaborative innovation.

In the current research trends of regional collaborative innovation, foreign scholars' research on cross-regional collaborative innovation focuses on the research perspective, the relationship with innovation performance, and

influencing factors. In terms of research perspectives, the perspectives of papers and patents, strategic emerging industries, and industrial ecosystems are introduced [6]. The research on the relationship between regional collaborative innovation and innovation performance focuses on the evaluation of cross-regional system innovation capability [7]. The research on the influencing factors of regional collaborative innovation is carried out from the perspective of collaborative cooperation tendency, innovation driving factors, and the cooperation relationship of innovation subjects [8, 9]. The domestic scholars' research on cross-regional collaborative innovation ability mainly focuses on the evaluation of innovation ability, the research on the influencing factors, and mechanism and mode of innovation ability. The evaluation of cross-regional collaborative innovation capability provides different insights from different perspectives, covering research methods such as dynamic evaluation methods, comparative analysis methods, DEA, and Malmquist index [10, 11]. The research on the influencing factors of cross-regional collaborative innovation capability has analysed its influencing factors from different research perspectives, including policy conditions, network development, resource investment, external learning, and communication. Some scholars have also explored the mechanism and model of cross-regional collaborative innovation, starting from the perspective of synergetic, undertaking industrial transfer, and technology diffusion.

Innovative service capability refers to the comprehensive evaluation of the effect of a service innovation activity by an enterprise. The innovative service capability not only reflects the effectiveness of the organization's operations and market competition but also reflects the projected development level and actual effect of the organization's projects. In recent years, research on innovative service capabilities has mostly focused on service-oriented enterprises and manufacturing industries that provide products and services at the same time [12, 13]. In order to achieve effective market competition, not only service-oriented enterprises and manufacturing-oriented enterprises but also high-tech enterprises also strive to provide customers with perfect services and constantly carry out service innovation activities in order to improve the service quality of enterprises, thereby attracting and retaining customers, and stabilize or expand the market [14, 15]. In addition, under the leadership of regional collaborative innovation, whether an enterprise's innovation service capability can be improved is crucial to the overall performance of the enterprise.

It can be seen by combing the relevant research at home and abroad that the existing literature provides a basic analysis framework for the expansion of innovation service capabilities and sustainable development at the practical level, but there are relatively few studies on regional innovation service capabilities, and the existing evaluation indicators. The synergy between innovation elements in the system has not been considered. One of the purposes of the measurement and evaluation of innovation service capacity is to form a more effective governance mechanism and model to promote the development of interregional system

innovation, so it is necessary to study the evaluation of innovation service capacity.

Based on the current status of regional collaborative innovation development, this study is dedicated to building an evaluation model for innovation service capabilities. First, this study uses fuzzy principal component analysis to refine quantitative and qualitative indicator data for innovation service capability selection and calculates index weights. Second, this study uses TOPSIS to assist in the evaluation of innovative service capabilities. Due to the large amount of data in this article, this study was implemented using MATLAB programming. Finally, through a comparative analysis of the results, this article obtained an optimization plan for improving innovation service capabilities and proposed countermeasures and suggestions on how to guide the flow of innovation elements and improve regional collaborative innovation.

2. Evaluation Index System Construction and Evaluation Method

2.1. Regional Collaborative Innovation Network. Citing the ARA business network model proposed by Håkansson and Snehota [16] (Figure 1, combined with the relationship between actors and social networks involved in social network analysis, a regional collaborative innovation network is defined. The commercial network contains three elements: participants, resources, and activities. It can be seen that if you want to research a network, you must first define the network and secondly define the participants (or actors), resources, and activities (social relations). The resources and activities here can be collectively referred to as relationships, and they have an inseparable relationship with the participants. In addition, there are relationships among participants, resources, and activities. In order to facilitate measurement, they are called relationships here. As a result, the participants and relationships of a network are defined, and the boundaries of the network are clearly defined.

Participants in collaborative innovation networks usually cover enterprises, universities, scientific research institutions, intermediary services, governments, etc., but the core network of cooperative innovation networks mainly includes enterprises, universities, and scientific research institutions; networks composed of governments, intermediary services, etc., are marginal network, which can also be called a related network. The main research of this study is the core network, that is, the main body of the network is enterprises, universities, and scientific research institutions.

The regional collaborative innovation network studied in this paper selects the biopharmaceutical industry in the high-tech industry and builds a collaboration based on the relationship between companies, universities, and scientific research institutions in the biopharmaceutical industry and the application of patent (national invention patents and utility models) innovation network. The results of the visual analysis of the development of the biological medicine industry in the Yangtze River Delta are shown in Figure 2.

It can be seen from Figure 2 that the innovation ability of universities in the collaborative innovation network of the

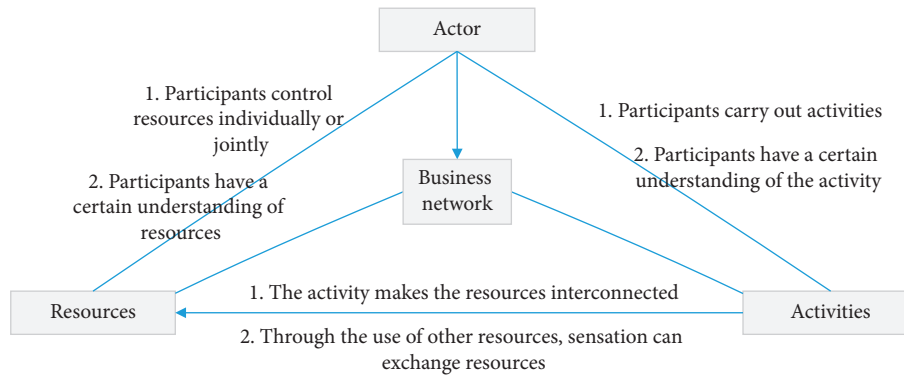


FIGURE 1: ARA business network model [16].

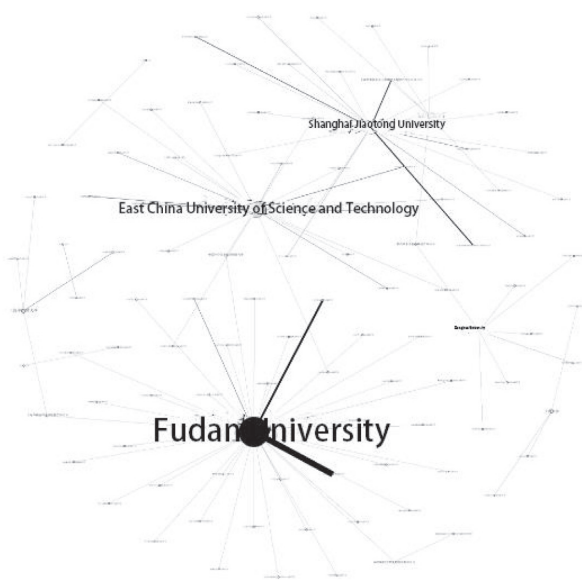


FIGURE 2: Visual analysis of the regional coordinated innovation network of the biomedical industry in the Yangtze River Delta.

biomedical industry in the Yangtze River Delta is obvious. It can be seen from the figure that the more prominent nodes are Fudan University, East China University of Science and Technology, and Shanghai Jiaotong University, which shows that the role of universities in regional collaborative innovation is far greater than that of enterprises and scientific research institutions. Moreover, the collaborative innovation cooperation between enterprises and scientific research institutions also relies on colleges and universities, and colleges and universities have outstanding effects in regional system innovation.

2.2. Construction of Evaluation Index System for Innovation Service Capability. This article takes the “China Regional Innovation Capacity Report” undertaken by the China Science and Technology Development Strategy Research Group as an example to describe the overview of the innovation capacity evaluation index system. First of all, the innovation subjects of the regional innovation capability

framework are universities, R&D institutions, enterprises, intermediary institutions, and governments. Secondly, the framework takes into account the flow of innovation elements among innovation entities and the innovation capability of the innovation entities themselves. Thirdly, the framework emphasizes the importance of the construction of an innovation environment and points out that the government needs to carry out function transformation to promote technological innovation of enterprises. Finally, the innovation capability framework takes into account the three dimensions of regional development stock, relative level, and growth rate. On the basis of the above analysis framework, we constructed five first-level indicators, 20 second-level indicators, 40 third-level indicators, and 137 fourth-level indicators including knowledge creation, knowledge acquisition, enterprise innovation, innovation environment, and innovation performance. Regional innovation capability index system.

On this basis, combined with the needs of innovation service capability assessment, the five indicators of knowledge creation, knowledge acquisition, enterprise innovation, innovation environment, and innovation performance are assisted in three aspects: regional collaborative innovation input, regional collaborative innovation output, and regional collaborative innovation conditions. Construction of indicators, specifically when establishing an evaluation indicator system, selects indicators from four aspects: simplicity, relativity, representativeness, comparability, and feasibility of indicators [12, 13, 15, 17]. The specific evaluation index system for innovation service capability is shown in Table 1.

2.3. Evaluation Method of Innovation Service Capability in the Yangtze River Delta

2.3.1. Selection of TOPSIS Method. Simply simplifying the data and reducing the dimensions cannot tell whether the decision is effective. The problem of innovation service capability evaluation belongs to the category of multi-objective decision-making, so the problem can only be solved effectively by using good decision-making methods. TOPSIS (Technology for Order Preference by Similarity to an Ideal Solution) method is an analysis method for solving multiobjective decision-making problems with limited

TABLE 1: Innovation service capability index system.

First-level indicators	Secondary indicators	Third-level indicators	Unit	Code	Citations	
Regional collaborative innovation investment	Enterprise	Number of enterprises	Pc	X1	[12]	
		R&D personnel	People	X2	[13]	
	Efficient	Number of universities	Pc	X3	[15]	
		R&D personnel	People	X4	[17]	
	R&D institutions	Number of R&D institutions	Pc	X5		
		R&D personnel	People	X6		
		Government finance technology grant	Billion	X7		
		Government financial project appropriation	Billion	X8		
		Government	Government science and technology project implementation funds	Billion	X9	
			Scientific papers	Article	X10	
Regional collaborative innovation output	Results	Patent application	Item	X11		
		Patent ownership transfer	Item	X12		
		Form an expert or industry standard	Item	X13		
		New product sales revenue	Billion	X14		
		Patent transfer income	Billion	X15		
Assistance in regional collaborative innovation	Economic basis	GDP	Billion	X16		
		Financial institution loan balance	Billion	X17		
	Resident life	Employed population	People	X18		
		Number of college graduates or above	People	X19		

solutions [18]. Its logical relationship is relatively simple and clear, with good interactivity and easy implementation. At present, this method has been applied to the research of innovation service capability evaluation. Therefore, the reason for choosing the TOPSIS algorithm in this study is as follows:

- (1) The indicators in this article are more objective, and it is not suitable to choose subjective evaluation methods, such as intuitive judgment method and linear weighting method
- (2) The evaluation results in this paper do not involve cost and efficiency, so the activity cost method and data envelope algorithm are not used
- (3) The indicator data in this paper are relatively small, so if the artificial neural network algorithm is not used, the results will not converge
- (4) The indicators in this article are clear and there is no greater ambiguity, so fuzzy comprehensive evaluation method and gray system theory method are not used

3. Model Principle

3.1. Principle of Fuzzy Principal Component Analysis. Principal component analysis is a statistical analysis method that converts multiple variables into a few principal components through dimensionality reduction techniques, which uses the original more evaluation indicators with simplified comprehensive principal component indicators. Instead, the comprehensive indicator retains most of the information of the original variables and is not related to each other, which can simplify complex problems. The fuzzy principal component analysis method is based on the variance contribution rate of the principal component combined with the fuzzy theory to obtain a reasonable correction weight. The specific process is as follows:

After standardizing the secondary indicators of the original data, perform weighted synthesis to obtain the evaluation value of the primary indicators:

$$\eta_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^n (a_{ij})^2}}, \quad \text{where } i = 1, 2, \dots, n; \quad (1)$$

$$j = 1, 2, \dots, k.$$

In the formula, a_{ij} represents the evaluation value of the j -th secondary index of the i -th region and η_{ij} represents the evaluation value of the i -th secondary index of the j -th region after normalization. The comprehensive weighting of the secondary indicators yields $x_{ij}^* = w_{ij}\eta_{ij}$, where w_{ij} is the weight of the secondary indicators in the corresponding indicators.

Principal component analysis of first-level indicators
Initial sample matrix $X^* = (x_{ij}^*)_{n \times p}$, where $i = 1, 2, \dots, n; j = 1, 2, \dots, p$

- (1) Standardize the index according to the $x_{ij} = x_{ij}^* - \bar{x}_j^*/s_j^*$ formula to obtain a standardized evaluation matrix $X = (x_{ij})_{n \times p}$. In the formula, \bar{x}_j^* and s_j^* are the sample mean and sample standard deviation of the j index, respectively.
- (2) Calculate the correlation coefficient matrix $R_{p \times p}$, eigenvalue $\lambda_1 \geq \dots \geq \lambda_p \geq 0$, and normalized eigenvector E_{ij} between indicators.
- (3) Using the unit feature vector corresponding to the previous eigenvalue of p , the principal component of p can be expressed as a linear combination of the original indicators, thereby obtaining the principal component equation:

$$F_{ij} = X_{ij}E_{ij}, \quad i = 1, 2, \dots, n; j = 1, 2, \dots, p. \quad (2)$$

- (4) The variance contribution rate of the j principal component is $\mu_j = \lambda_j / \sum_{j=1}^p \lambda_j$. When the cumulative variance contribution rate $\mu = \sum_{j=1}^m \mu_j$ reaches a certain value (generally not less than 85%), the first m principal components $F_{i1}, F_{i2}, \dots, F_{im}$ ($i = 1, 2, \dots, n$) are taken, that is, the m principal components are considered to reflect the information of the original p evaluation indicators with fewer indicators.
- (5) Use the variance contribution rate of each principal component as the initial weight, and modify the initial weight to obtain the final weight. The reason is that when evaluating the original indicators, each expert gives the number of weight intervals of each indicator and the optimism of the collaborative innovation ability of each candidate area according to the needs of the enterprise. For this purpose, it is necessary to calculate the weight $W_i = \{w_{ij}\}$, $i = 1, 2, \dots, n; j = 1, 2, \dots, m$ that corrects the weight of the principal component and correct the weight of the principal component. The specific correction process is as follows:

$$B([p_j, q_j]) = (1 - a_j) \frac{p_j - p_{\min}}{p_{\max} - p_{\min}} + a_j \frac{q_j - q_{\min}}{q_{\max} - q_{\min}}. \quad (3)$$

The weight fuzzy set $B\{[p_1, q_1], [p_2, q_2], \dots, [p_n, q_n]\}$ is the value of the weight interval of the evaluation index, and $B([p_n, q_n])$ indicates the relative importance of the index. In the formula, $p_{\min} = \min_{1 < j < n} p_j$, $p_{\max} = \max_{1 < j < n} p_j$, $q_{\min} = \min_{1 < j < n} q_j$, $q_{\max} = \max_{1 < j < n} q_j$, and a_j represents the optimistic coefficient for the region, $0 < a_j < 1$.

Therefore, normalizing the important membership degree can get the index weight as follows:

$$w_{ij} = \frac{B([p_j, q_j])}{\sum_{j=1}^m B([p_j, q_j])}, \quad i = 1, 2, \dots, n. \quad (4)$$

- (6) Amend the preliminary weight to obtain the final weight of the regional innovation capability evaluation index:

$$\beta_i = \frac{w_{ij}\mu_j}{\sum_{j=1}^m w_{ij}\mu_j}, \quad i = 1, 2, \dots, n. \quad (5)$$

3.2. Principle of TOPSIS Evaluation Method. Similarity to ideal solution sorting method TOPSIS is a commonly used scheme limited multiobjective decision analysis method, which has the ideal solution and negative ideal solution to sort the program set by means of a multiproperty issue. The ideal solution is a virtual best solution that does not exist in the solution set, and each of its attribute values is the best

value in the solution; the negative ideal solution is the virtual worst solution, and each attribute value is the best in the solution. Poor value. All the alternatives in the solution set are compared with the distance between the ideal solution and the negative ideal solution. The solution that is close to the ideal solution and far from the negative ideal solution is the best solution. The specific evaluation steps are as follows.

3.2.1. Construct the Initial Matrix. There are n candidate programs and m evaluation indicators, and x_{ij}^* represents the evaluation value of the j index of the i program, and the initial matrix V of the principal component obtained by $V = X^*{}^T U$ (U is the coefficient matrix) is standardized according to $v_{ij} = v_{ij}^* / \sqrt{\sum_{i=1}^n (v_{ij}^*)^2}$, $i = 1, 2, \dots, n; j = 1, 2, \dots, m$. The standardized matrix $A = (a_{ij})_{n \times m}$ is obtained.

3.2.2. Construct Weighted Normalization Matrix.

$$B = A^T W = \begin{pmatrix} w_{11}a_{11} & w_{12}a_{12} & \dots & w_{1m}a_{1m} \\ w_{21}a_{21} & w_{22}a_{22} & \dots & w_{2m}a_{2m} \\ \dots & \dots & \dots & \dots \\ w_{n1}a_{n1} & w_{n2}a_{n2} & \dots & w_{nm}a_{nm} \end{pmatrix}. \quad (6)$$

Among them $W_{n \times m}$ are the correction weights obtained by fuzzy principal component analysis.

3.2.3. Determine the Ideal Solution B^+ and Negative Ideal Solution B^- .

$$B^+ = \left[\max_i b_{ij} \mid j \in J \right] = [b_1^+, b_2^+, b_j^+, \dots, b_m^+], \quad i = 1, 2, \dots, n,$$

$$B^- = \left[\min_i b_{ij} \mid j \in J \right] = [b_1^-, b_2^-, b_j^-, \dots, b_m^-], \quad i = 1, 2, \dots, n. \quad (7)$$

3.2.4. Calculate the Distance. The distance between the evaluation value and the ideal solution and the negative ideal solution is given by

$$s_i^+ = \sqrt{\sum_{j=1}^m (b_{ij} - b_j^+)^2}, \quad s_i^- = \sqrt{\sum_{j=1}^m (b_{ij} - b_j^-)^2}. \quad (8)$$

3.2.5. Determine the Relative Proximity C_i .

$$C_i = \frac{s_i^-}{s_i^- + s_i^+}, \quad i = 1, 2, \dots, n, \quad 0 \leq C_i \leq 1. \quad (9)$$

3.2.6. Sorting Preference. The schemes are sorted and selected in the order of C_i from large to small, and the relative closeness C_i selects the optimal scheme, which is the area with the highest innovation service capability.

4. Case Study

In May 2010, the State Council formally approved the implementation of “the Regional Plan for the Yangtze River Delta Region”. The document states that the Yangtze River Delta includes Shanghai, Jiangsu, and Zhejiang. In 2014, the “Guiding Opinions of the State Council on Relying on the Golden Waterway to Promote the Development of the Yangtze River Economic Belt” also made it clear for the first time to include Anhui Province in the Yangtze River Delta to participate in development. In May 2016, the Standing Committee of the State Council formally adopted the “Development Plan for the Urban Agglomeration of the Yangtze River Delta Region.” The plan clearly mentioned that Shanghai should play a leading role to promote the common development of various metropolitan areas, including Nanjing Metropolitan Area, Hangzhou Metropolitan Area, Hefei Metropolitan Area, Suzhou Changzhou Metropolitan Area, and Ningbo Metropolitan Area. The Yangtze River Delta urban agglomeration is complete with large, medium, and small cities, with one megacity, one megacity, 13 large cities, 9 medium cities, and 42 small cities. The Yangtze River Delta is a typical region for regional

collaborative innovation development, and this study uses this as an object for case analysis.

4.1. Expert Questionnaire Survey and Data Collection. Table 2 shows the original data and initial weights for enterprises to evaluate suppliers, and Table 3 shows the result of weighted standardization of the secondary indicators of their original data. The expert survey method was used to verify the construction of innovative service capabilities in this study, and then the corresponding indicator data were selected. The data involved in this article were all from the China Statistical Yearbook, the China Science and Technology Yearbook, and the statistical yearbooks of various provinces and cities.

4.2. Fuzzy Principal Component Analysis. According to the initial evaluation weight interval of each index: [17, 23], [11, 18], [6, 9], we use the fuzzy algorithm to obtain the modified weight W_i and the final principal component weights β_j , respectively:

$$\begin{aligned} a_1 &= 0.95 \quad W_1 = (0.3339, 0, 0.4965, 0.1696) \quad \beta_1 = (0.5299, 0, 0.3836, 0.0865), \\ a_2 &= 0.76 \quad W_2 = (0.3361, 0, 0.4834, 0.1805) \quad \beta_2 = (0.5340, 0, 0.3739, 0.0921), \\ a_3 &= 0.82 \quad W_3 = (0.3354, 0, 0.4875, 0.1771) \quad \beta_3 = (0.5327, 0, 0.3770, 0.0903), \\ a_4 &= 0.65 \quad W_4 = (0.3373, 0, 0.4762, 0.1865) \quad \beta_4 = (0.5362, 0, 0.3686, 0.0952). \end{aligned} \quad (10)$$

Among them, a_1, a_2, a_3 , and a_4 are known expert's optimistic coefficients for innovation service capability.

4.3. TOPSIS Evaluation. In order to adapt to the problem of index selection and evaluation of innovation service capability with many indexes, this article uses MATLAB programming to implement the TOPSIS algorithm, which not only saves time, but also provides results more accurately. The specific calculation process and results are as follows.

4.3.1. Construct the Initial Matrix. In this paper, there are 4 candidate programs, namely, 4 regions and 3 evaluation indicators. x_{ij} represents the evaluation value of the j index of the i program. The initial evaluation matrix V is obtained by multiplying the initial matrix X by the coefficient matrix U :

$$V = \begin{pmatrix} 97.6278 & 76.3622 & 79.9790 \\ 40.1815 & 112.8293 & 37.4480 \\ 9.4977 & 41.3866 & 48.2996 \\ 89.6354 & 37.2349 & 70.5447 \end{pmatrix}. \quad (11)$$

Normalize according to $v_{ij} = v_{ij}^* / \sqrt{\sum_{i=1}^6 (v_{ij}^*)^2}$, $j = 1, 2, \dots, 4$ and get a standardized matrix $A = (a_{ij})_{4 \times 3}$.

4.3.2. Construct Weighted Normalization Matrix.

$$B = A^T W = \begin{pmatrix} 0.2873 & 0.0462 & 0.2367 \\ 0.1192 & 0.0354 & 0.1080 \\ 0.0281 & 0.0116 & 0.1405 \\ 0.2669 & 0.0416 & 0.2006 \end{pmatrix}. \quad (12)$$

Among them $W = \begin{pmatrix} 0.5299 & 0.0865 & 0.3836 \\ 0.5340 & 0.0921 & 0.3739 \\ 0.5327 & 0.0903 & 0.3770 \\ 0.5362 & 0.0952 & 0.3686 \end{pmatrix}$ are the

correction weights obtained by fuzzy principal component analysis.

4.3.3. Determine the Ideal Solution B^- and Negative Ideal Solution B^+ . From $B^+ = [\max_i b_{ij} | j \in J] = [b_1^+, b_2^+, \dots, b_3^+]$, $i = 1, 2, \dots, 4$, get

$$B^+ = [0.3356 \quad 0.2367 \quad 0.0462]. \quad (13)$$

TABLE 2: Expert survey.

First-level indicators		Secondary indicators		Shanghai	Jiangsu	Zhejiang	Anhui
Name	Weights	Name	Weights				
Regional collaborative innovation investment	[17, 23]	Enterprise	50	128	130	110	140
		Efficient	20	90	90	90	80
		R&D institutions	30	21	30	15	40
		Government	37	21	30	15	40
Regional collaborative innovation output	[11, 18]	Results	35	80	120	170	99
		Income	30	34	50	60	30
Assistance in regional collaborative innovation	[6, 9]	Economic basis	26	34	50	60	30
		Resident life	14	23	11	15	20

TABLE 3: Weighted synthesis of secondary indicators.

First-level indicators		Shanghai	Jiangsu	Zhejiang	Anhui
Name	Weights				
Regional collaborative innovation investment	[17, 23]	41.3956	40.08919	40.9547	41.86353
Regional collaborative innovation output	[11, 18]	37.08671	26.16658	39.15287	41.84066
Assistance in regional collaborative innovation	[6, 9]	40.16341	48.53234	30.64865	43.26791

From $B^- = [\min_i b_{ij} | j \in J] = [b_1^-, b_2^-, \dots, b_3^-], i = 1, 2, \dots, 4$, get

$$B^- = [0.0281 \quad -0.0132 \quad 0.0116]. \tag{14}$$

4.3.4. Calculate the Distance between the Evaluation Value and the Ideal Solution and the Negative Ideal Solution. From $s_i^+ = \sqrt{\sum_{j=1}^3 (v_{ij} - v_j^+)^2}, i = 1, 2, \dots, 4$, get the distance between the evaluation value and the ideal solution:

$$s^+ = [0.0483 \quad 0.2521 \quad 0.3241 \quad 0.0777]. \tag{15}$$

From $s_i^- = \sqrt{\sum_{j=1}^3 (v_{ij} - v_j^-)^2}, i = 1, 2, \dots, 4$, get the distance between the evaluation value and the ideal solution:

$$s^- = [0.3618 \quad 0.1535 \quad 0.1537 \quad 0.3220]. \tag{16}$$

4.3.5. Determine the Relative Proximity

C. From $C_i = s_i^- / (s_i^- + s_i^+), i = 1, 2, \dots, 4, 0 \leq C_i \leq 1$, get:

$$C = [0.8822 \quad 0.3785 \quad 0.8055 \quad 0.3217]. \tag{17}$$

Think of the value of C_i as a score for the innovation service capability of each region, and select a better region. Because of $C_1 > C_3 > C_2 > C_4$, the first area, Shanghai, has a higher innovation service capability than other areas.

5. Conclusions and Recommendations

5.1. Research Conclusion. The innovation of this study mainly lies in the innovation of methods. In the regional collaborative innovation network, the evaluation of innovation service capability is very important. At present, the most commonly used methods are fuzzy clustering, DEA method, AHP method, and fuzzy analytic hierarchy process. When analyzing these methods, it is not difficult to find that the current innovation service capacity evaluation involves

too many indicators. The calculation process becomes very complex, even difficult to calculate. The use of fuzzy principal component analysis combined with TOPSIS can objectively select regional collaborative innovation indicators, which has higher accuracy, simpler principles, and more objective conclusions, and when the indicators increase, this method performs data processing. Effective decision-making has more advantages. When there are a lot of data, use statistical software to process the data, and implement the algorithm through MATLAB programming, you can quickly get scientific results, so you can make decisions quickly. Of course, this method can be applied not only to the evaluation of innovative service capabilities, but also to other aspects, such as supplier selection, route selection, and location selection of distribution centers.

This article takes the biomedical industry as an example in the regional collaborative innovation network. It can be found that universities play an important role in the network, which shows that companies tend to cooperate with universities to innovate, and they must pay attention to and play a role in the improvement of innovation service capabilities. The important role of universities in the process is innovation creation and dissemination. Through the index construction and evaluation process of innovation service capacity, it can be found that among the four regions in the Yangtze River Delta, Shanghai, Jiangsu Province, Zhejiang Province, and Anhui Province, Shanghai has the strongest innovation service capacity, followed by Zhejiang Province, Then came Jiangsu Province, and finally Anhui Province, and the conclusions of the study are in good agreement with the actual situation in the Yangtze River Delta.

5.2. Suggestions. Finally, according to the research conclusion of this paper, the following two countermeasures and suggestions are proposed to enhance the innovation service capability.

First of all is the optimization suggestions on the innovation level of industry-university-research cooperation: the improvement of regional collaborative innovation capabilities needs to be optimized from the perspective of regional and collaborative innovation environments, strengthen regional cooperation, increase the exchange and cooperation of innovations, avoid overlapping information, and enjoy the support of multiple policies and multiple innovation environments. At the same time, Shanghai, as a leader in the economic development of the Yangtze River Delta, should play its leading role in creating a globally influential scientific and technological innovation center and realize the common development of the region. The independent innovation activities of universities, enterprises, and scientific research institutions are not preferred. The optimization of the environment requires many efforts. The core body of the industry-university-research cooperative innovation network is an important participant and an important builder of the network. Financial institutions and intermediaries should also play their roles to form an open and dynamic innovation environment; an industry-university-research cooperation innovation platform is established to maximize resource utilization efficiency. Innovation requires the cooperation and resource sharing of the core and related subjects, so a mature operating mechanism is needed to integrate the resources of universities, enterprises, scientific research institutions, governments, financial institutions, and intermediary institutions to optimize the efficiency of resource utilization, so establishing a collaborative innovation platform is crucial.

Secondly, the government should optimize from the level of policies and regulations, introduce corresponding encouragement policies, and give more preferential tax policies to promote the creation and development of innovation. In addition to supporting and coordinating functional R&D transformation platforms, in addition to optimization at the legal level, online and offline platforms can be established, combined with existing advanced information technology, so that the main body of the industry-university-research cooperation innovation network can better share resources and technology, a corresponding biomedical R&D and transformation platform is established to help companies innovate better and promote faster marketization of innovations. The government plays an important role in this process because the public service platform must protect public welfare and encourage the research and development of shared technologies, so the government plays a vital role in it.

Data Availability

The data involved in this article were all from the China Statistical Yearbook, the China Science and Technology Yearbook, and the statistical yearbooks of various provinces and cities. Among them, the data in Figure 2 come from the patent search database of the China Intellectual Property Office. The data used to support the findings of this study were supplied by author under license and so cannot be made freely available. Requests for access to these data should be made to [Xiao-min Gu, guxiaomin@lixin.edu.cn].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] I. Kim, "Managing Korea's system of technological innovation," *Interfaces*, vol. 23, no. 6, pp. 13–24, 1993.
- [2] G. Huber, "Synergies between organizational learning and creativity & innovation," *Creativity and Innovation Management*, vol. 7, no. 1, pp. 3–8, 1998.
- [3] S. Radosevic, "Defining systems of innovation: A methodological discussion," *Technology in Society*, vol. 20, no. 1, pp. 75–86, 1998.
- [4] L. Power, "The missing link: Technology, investment, and productivity," *Review of Economics and Statistics*, vol. 80, no. 2, pp. 300–313, 1998.
- [5] B. Morris, "High technology development: Applying a social network paradigm," *Journal of New Business Ideas and Trends*, vol. 4, no. 1, pp. 45–59, 2006.
- [6] I. De Noni, L. Orsi, and F. Belussi, "The role of collaborative networks in supporting the innovation performances of lagging-behind European regions," *Research Policy*, vol. 47, no. 1, pp. 1–13, 2018.
- [7] A. Marasco, M. De Martino, F. Magnotti, and A. Morvillo, "Collaborative innovation in tourism and hospitality: A systematic review of the literature," *International Journal of Contemporary Hospitality Management*, vol. 30, no. 6, pp. 2364–2395, 2018.
- [8] F. Caviggioli, "Technology fusion: Identification and analysis of the drivers of technology convergence using patent data," *Technovation*, vol. 55–56, no. 6, pp. 22–32, 2016.
- [9] S. Han, G. M. Yoo, and S. Kwak, "A comparative analysis of regional innovation characteristics using an innovation actor framework," *Science, Technology and Society*, vol. 23, no. 1, pp. 137–162, 2018.
- [10] C. Lindsay, P. Findlay, J. McQuarrie, M. Bennie, E. D. Corcoran, and R. Van Der Meer, "Collaborative innovation, new technologies, and work redesign," *Public Administration Review*, vol. 78, no. 2, pp. 251–260, 2018.
- [11] L. Zhen, "Tactical berth allocation under uncertainty," *European Journal of Operational Research*, vol. 247, no. 3, pp. 928–944, 2015.
- [12] D. C. Bello, L. P. Radulovich, R. G. Javalgi, R. F. Scherer, and J. Taylor, "Performance of professional service firms from emerging markets: role of innovative services and firm capabilities," *Journal of World Business*, vol. 51, no. 3, pp. 413–424, 2016.
- [13] O. F. Bustinza, E. Gomes, F. Vendrell-Herrero, and T. Baines, "Product-service innovation and performance: The role of collaborative partnerships and R & D intensity," *R & D Management*, vol. 49, no. 1, pp. 33–45, 2019.
- [14] L. Zhen, "Modeling of yard congestion and optimization of yard template in container ports," *Transportation Research Part B: Methodological*, vol. 90, pp. 83–104, 2016.
- [15] V. Jayaraman and Z. Liu, "Aligning governance mechanisms with task features to improve service capabilities---an empirical study of professional service outsourcing in India," *Operations Management Research*, vol. 12, no. 1–2, pp. 19–39, 2019.
- [16] H. Håkansson and I. Snehota, "Developing relationships in business networks," *Journal of Purchasing*, pp. 1–261, 1995.
- [17] R. P. J. Rajapathirana and Y. Hui, "Relationship between innovation capability, innovation type, and firm

- performance,” *Journal of Innovation & Knowledge*, vol. 3, no. 1, pp. 44–55, 2018.
- [18] C. L. Hwang and K. Yoon, “Multiple attribute decision making,” *Lecture Notes in Economics & Mathematical Systems*, vol. 404, no. 4, pp. 287–288, 1981.

Research Article

Optimization of Continuous Berth Scheduling by Taking into Account Double-Line Ship Mooring

Cheng Luo, Hongying Fei , Dana Sailike, Tingyi Xu, and Fuzhi Huang

School of Management, Shanghai University, Shanghai 200444, China

Correspondence should be addressed to Hongying Fei; feihy@shu.edu.cn

Received 2 July 2020; Revised 20 August 2020; Accepted 31 August 2020; Published 18 September 2020

Academic Editor: Tingsong Wang

Copyright © 2020 Cheng Luo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

“Double-Line Ship Mooring” (DLSM) mode has been applied as an initiative operation mode for solving berth allocation problems (BAP) in certain giant container terminals in China. In this study, a continuous berth scheduling problem with the DLSM model is illustrated and solved with exact and heuristic methods with an objective to minimize the total operation cost, including both the additional transportation cost for vessels not located at their minimum-cost berthing position and the penalties for vessels not being able to leave as planned. First of all, this problem is formulated as a mixed-integer programming model and solved by the CPLEX solver for small-size instances. Afterwards, a particle swarm optimization (PSO) algorithm is developed to obtain good quality solutions within reasonable execution time for large-scale problems. Experimental results show that DLSM mode can not only greatly reduce the total operation cost but also significantly improve the efficiency of berth scheduling in comparison with the widely used single-line ship mooring (SLSM) mode. The comparison made between the results obtained by the proposed PSO algorithm and that obtained by the CPLEX solver for both small-size and large-scale instances are also quite encouraging. To sum up, this study can not only validate the effectiveness of DLSM mode for heavy-loaded ports but also provide a powerful decision support tool for the port operators to make good quality berth schedules with the DLSM mode.

1. Introduction

With the deepening of economic globalization, the world container transportation volume has increased dramatically in recent years [1]. As one of the most important components of container transportation network, container terminals play an important role in world economy. Considering that the efficiency of berth allocation problem has great impact on the output of container terminals, a lot of studies have been dedicated to the berth scheduling problems.

In general, berth schedules are determined by specifying berthing time and position for the coming vessels by taking into account various constraints, such as the berth capacity, announced arrival time and departure time of container ships, and certain specific berthing requirements. In order to avoid collisions between vessels, the single-line ship mooring (SLSM) mode, which specifies that “no more than one vessel can be allocated to the same berth position at the same time,”

is normally applied in container terminals over the world, and this rule is regarded as a default in most of the studies on berth scheduling [2].

As one of the most important economic role in the world economy, China is continuously developing the economic innovation, resulting in huge container throughput of the international hubs in China. Yangshan deep-water port is located in Shanghai, a mega city in China. As the largest sea-island artificial deep-water port, Yangshan deep-water port is an important part of Shanghai International Shipping Center, and its annual throughput has increased constantly since it was built in the year of 2005.

Having been one of the world’s busiest container terminals, Yangshan deep-water port has applied a so-called “Double-Line Ship Mooring” (DLSM) mode to build berth schedules since the year of 2019. Different from the widely used SLSM mode, DLSM mode allows two container ships to be moored simultaneously at the same berth location, enabling more container ships to be moored at their ideal berth

and cranes to operate two container ships at the same time, a reasonable way to improve the efficiency of berth operation. Despite that, the real efficiency of the port with DLSM mode depends on how sophisticated the operators are according to the investigation made at Yangshan port, because berth scheduling with DLSM is much more complex than with the SLSM system. In consequence, it is essential to develop an effective decision support system to help berth operators improve the quality and efficiency of daily schedules with the DLSM mode.

This study aims at minimizing the additional operation costs for the vessels not placed at their ideal berthing position and penalties for the ships not being able to leave before their preplanned departure time, for a continuous berth allocation problems (BAP) model with the DLSM mode. The main contributions of this study are as follows: (1) as the first study on berth scheduling with the DLSM mode, it validates the contribution of DLSM model in comparison with the widely used SLSM mode; (2) a mixed integer programming model is constructed for the continuous BAP with DLSM mode, enabling a benchmark for the future studies on such topic; (3) a PSO algorithm is developed to obtain good quality DLSM berthing schedules within reasonable execution time, offering the berth operators at busy ports a powerful decision support tool to improve their work efficiency.

The reminder of this paper is structured as follows. Section 2 is dedicated to the literature review of related work. The problem description is given in Section 3. In Section 4, the targeted problem is formulated as a mixed integer programming model, which can be solved by CPLEX solver, and then, the details of the proposed PSO algorithm are given in Section 5. Section 6 is dedicated to discussions about the experimental results, and this paper ends up with conclusions and perspectives.

2. Related Work

According to the literature, berth allocation problem (BAP) is regarded as the most important issue faced by the management of container terminals. As the quality of berth schedule has a great influence on the improvement of port operation efficiency, a lot of researchers have been studying BAPs, and numerous results were published [2]. Since Imai et al. [3] firstly addressed a static berth allocation problem (SBAP) and further developed a dynamic BAP model (DBAP for public berth system [4]), numerous studies have been published on BAP aiming to optimize the operation efficiencies by taking into account the various constraints affecting BAP [2], such as tidal influence [5], vessel service priority [6], time-varying water depth [7], and channel flow control [8]. As to our knowledge, no study on BAPs with the DLSM mode is observed in the literature; thus, this study aims to fill the research gaps in this field by taking into account the DLSM mode when constructing berth schedules. Furthermore, this study is specifically dedicated to continuous BAP, which copes with the real situation of Yangshan deep-water port that we have investigated and can also be applied to many other huge container hubs.

With regard to the methodologies applied to BAPs, it can be observed in the literature that the BAPs are normally formulated as Mixed Integer Programming (MIP) models, which can be solved with commercial programming solvers for small-size problems [2–11]. Since BAPs are NP-hard, lots of researchers have also taken much effort on developing near-optimal solutions with efficient heuristics and meta-heuristics, such as genetic algorithm and its variants [12–15], subgradient optimization method [15], simulated annealing [9], evolutionary algorithms [16, 17], particle swarm optimization [10], and island-based meta-heuristic algorithm [18].

Particle swarm optimization (PSO) is an evolutionary computation technique firstly developed by Eberhart and Kennedy [19]. Since the particles used in PSO algorithm are only updated by internal velocity and fewer parameters should be tuned, the principles of PSO can be easily understood and widely adapted to various specific applications. According to the literatures, PSO algorithms have been proven quite efficient in dealing with berth allocation [20] yard allocation [21], quay crane scheduling [10, 22]; and many other planning and scheduling problems [23, 24]; in consequence, a PSO algorithm is also proposed in this study to develop efficient berth schedules by taking into account DLSM.

3. Problem Description

3.1. Assumptions. In general, the berth operators are in charge of arranging each vessel arriving at the port to a suitable berth position according to the availability of the wharf resources and respecting the preplanned arrival and departure of the vessel. According to the practices observed at the targeted container terminal, a set of assumptions are defined as follows:

- (1) Each vessel arrives at the port on the preplanned arrival time
- (2) If a vessel cannot leave the port before the preplanned departure time, the port has to incur a penalty
- (3) The coordinate corresponding to the leftmost end of the vessel is used to represent its berthing position, using the leftmost boundary of the wharf as the coordinate reference point
- (4) Each vessel has a predefined minimum-cost berthing position, which is determined according to the goods that will be loaded/unloaded at the port, and if a vessel cannot be berthed at its ideal berthing position, additional operation cost will occur
- (5) DLSM mode is applied, i.e., at most, two vessels can be simultaneously moored at the same berth position
- (6) When two vessels are moored in double-line, the length of the inner-side ship cannot be longer than the vessel moored outside
- (7) When two vessels are moored in double-line, the inner-side one must be berthed earlier and leave later than the outside one

3.2. *Notations.* For a better understanding, two types of notations are applied in this study: (1) Latin letters are used to denote parameters; (2) Greek letters are applied to represent decision variables.

3.2.1. Parameters

V : set of the vessels waiting to be allocated at the berth;
 L : length of the wharf at the port;
 T : planning horizon of this study;
 M : sufficiently large positive number;
 i, j, k : indices of vessels $i, j, k \in V$;
 c_{1i} : unit distance cost for transporting containers from/to vessel $i, i \in V$;
 c_{2i} : penalty cost per unit of time caused by the late departure of vessel i from the port, $i \in V$;
 p_i : predefined minimum-cost berthing position of vessel $i, i \in V$.
 e_i^{arr} : preplanned arrival time of vessel i ;
 h_i : total handling time required by the port to finish the necessary unloading/unloading operations of vessel $i, i \in V$;
 e_i^{dep} : preplanned departure time of vessel $i, i \in V$;
 l_i : length of vessel $i, i \in V$. In this study, the necessary gap that must be reserved to guarantee the safety is also integrated in this value for each vessel;

3.2.2. Decision variables

μ_i : integer variable representing the actual berthing position of vessel $i, i \in V$;
 θ_i : Integer variable representing the actual berthing time of vessel i , i.e., the moment that this vessel is moored at its berthing position where it can be operated by the cranes, $i \in V$;
 $\sigma_{i,j}$: binary variable representing the relative berthing position of two adjacent vessels, which equals 1 if vessel i is moored on the left to vessel j ($i, j \in V$) and 0 otherwise;
 $\pi_{i,j}$: binary variable indicating the sequential relationship between the berthing time of two vessels, which equals 1 if vessel i is berthed before vessel j ($i, j \in V$) and 0 otherwise;
 $\varepsilon_{i,j}$: binary variable representing the relative line position of two vessels that are moored in double-line, which equals to 1 if vessel i is berthed on the inner-side to vessel j ($i, j \in V$) and 0 otherwise.

3.3. Timeline Corresponding to the Berth Scheduling of a Vessel.

As shown in Figure 1, the timeline of the berth scheduling of vessel i starts from its arrival time at the port, denoted as e_i^{arr} , and ends up at the moment when all the necessary unloading/loading operations are completed, and the vessel is ready to leave the port for the next destination.

It is worth noting that (1) a gap between the preplanned arrival time e_i^{arr} and actual berthing time θ_i may be observed in the condition that no berth location is available on the arrival of the vessel which has to be waiting at the anchorage until in-wharf permission is delivered; (2) although setup operations (such as berthing, mooring ropes, and removing twist locks) are necessary before and after the cranes operating the vessels, the setup time of vessel i is integrated into its overall operation time h_i in this study to simplify the expression because it is observed in the targeted port that the setup time is generally not vessel-dependent.

3.4. *Berth Scheduling with DLSM Mode.* When the DLSM mode is applied at the port and two vessels will be scheduled to be moored in double-line at the same berthing position, it is necessary to determine the relative berthing line positions between these two vessels. Through the interview with the berth operators at the targeted port, a general rule is applied to ensure the berthing safety: Vessel i can be moored in the inner-side line to vessel j ($i, j \in V$) as long as the following conditions can be satisfied: (1) the length of vessel i is larger than that of vessel j ; (2) vessel i arrives no later than vessel j at the port; (3) the actual departure time of vessel i is not earlier than that of vessel j ; (4) the berthing position of the inner-side vessel should not be larger than that of the outside moored ship while the coordinate of the rightmost end of the former along the wharf must be at least as large as that of the latter.

Figure 2 is the three-dimensional schematic diagram of a DLSM example with three vessels, where vessels i and j are moored in double-line.

For a better understanding of this example, the corresponding side-wharf section and time-wharf one are detailed in Figures 3 and 4, respectively. It can be observed that vessel i is placed on the inner-side line to vessel j since the following conditions are satisfied (1) $l_i \geq l_j$ (as shown in Figure 3); (2) $\theta_i \leq \theta_j$ (as shown in Figure 4); (3) $\theta_i + h_i \geq \theta_j + h_j$ (as shown in Figure 4); (4) $p_i \leq p_j, p_i + l_i \geq p_j + l_j$ (as shown in Figure 4); in a berth schedule, a vessel can either be placed at its minimum-cost position (e.g., vessels j and k) or nonoptimal berthing position (e.g., vessel i).

As for vessel k , since it is moored on the right to the pair of double-lined ships, the coordinates of its berthing position must be larger than those of the rightmost of the inner-side moored ship to avoid overlaps, i.e., $\mu_k \geq \mu_i + l_i$.

4. Mathematical Formulation

4.1. *Construction of the Objective Function.* As mentioned in Section 1, the objective function of the targeted BAP problem consists of two parts:

- (1) Minimization of the additional operation cost for vessels not located at its minimum-cost berthing position

Considering that the ports normally predefine the minimum-cost berthing position for each vessel according the cargos that will be unloaded/loaded to maximize the operation efficiency, it is obvious that the larger is the deviation between the berthing

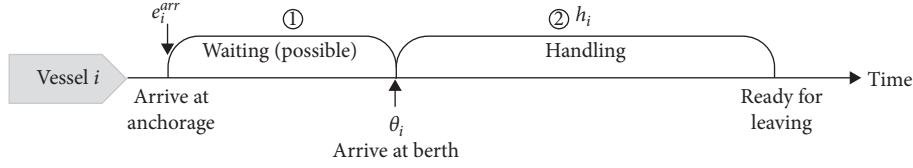
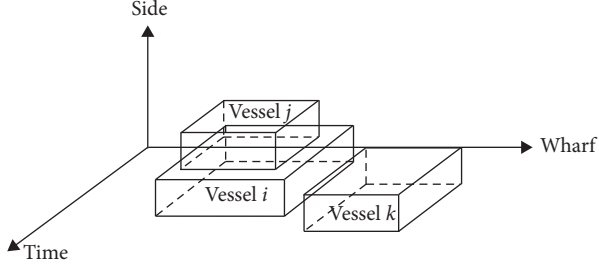
FIGURE 1: Timeline corresponding to the berth scheduling of vessel i .

FIGURE 2: Time-side-wharf schematic diagram of a DLISM example.

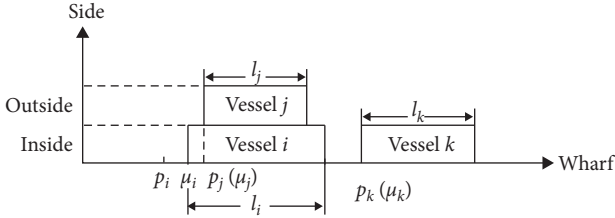


FIGURE 3: Side-wharf section of the DLISM example.

position and the predefined minimum-cost position of a vessel, the more is operation cost at the wharf; in this study, this part of the objective function is formulated as $\sum_{i \in V} c_{1i} |\mu_i - p_i|$.

- (2) Minimization of the penalty cost for vessels not leaving before their preplanned departure time

Let e_i^{dep} denote the preplanned departure time of vessel i . As is known, when the actual departure time of vessel i is later than e_i^{dep} , an additional cost will be incurred due to the influence of such delay on the rest of the voyage, and if the delay in the ship's departure is caused by the inefficient operation of the port, the port would have to pay certain fine for such delay. Therefore, this study aims also at minimizing the total penalties related to the delays of the vessels within the planning period, and this part of the objective function can be formulated as $\sum_{i \in V} c_{2i} (\theta_i + h_i - e_i^{dep})^+$, where $x^+ = \max\{0, x\}$.

To sum up, the objective function can be formulated as follows:

$$\text{【BAP】} \min \sum_{i \in V} \left\{ c_{1i} |\mu_i - p_i| + c_{2i} (\theta_i + h_i - e_i^{dep})^+ \right\}. \quad (1)$$

4.2. Mixed-Integer Programming Model. Considering that formula (1) is nonlinear, it should be linearized to meet the requirements of linear programming solver that will be applied in this study to optimally solve small-size instances.

Let $\alpha_i^+ = \mu_i - p_i$ when $\mu_i - p_i \geq 0$, $\alpha_i^- = p_i - \mu_i$ when $\mu_i - p_i < 0$, and β_i^+ and β_i^- denote the non-negative and negative value of $(\theta_i + h_i - e_i^{dep})$, respectively. Furthermore, let $\beta_i^+ = 0$ when $\theta_i + h_i - e_i^{dep} < 0$ to ensure that the smaller is β_i^+ , the better is the solution. Formula (1) can be linearized to formulas (2), (3), and (4) and then the mixed-integer programming model of the targeted problem can be constructed as follows:

$$\text{【BAP】} \min \sum_{i \in V} \{c_{1i} (\alpha_i^+ + \alpha_i^-) + c_{2i} \beta_i^+\}, \quad (2)$$

subject to

$$\mu_i - p_i = \alpha_i^+ - \alpha_i^-, \quad \forall i \in V, \quad (3)$$

$$\theta_i + h_i - e_i^{dep} = \beta_i^+ - \beta_i^-, \quad \forall i \in V, \quad (4)$$

$$\sum_{j \in V} \varepsilon_{i,j} \leq 1, \quad \forall i \in V, \quad (5)$$

$$\mu_i + l_i \leq L, \quad \forall i \in V, \quad (6)$$

$$\mu_i + l_i \leq \mu_j + M(1 - \sigma_{i,j}), \quad \forall i, j \in V, i \neq j, \quad (7)$$

$$\theta_i + h_i \leq \theta_j + M(1 - \pi_{i,j}), \quad \forall i, j \in V, i \neq j, \quad (8)$$

$$\theta_i \leq \theta_j + M(1 - \varepsilon_{i,j}), \quad \forall i, j \in V, i \neq j, \quad (9)$$

$$\theta_j + h_j \leq \theta_i + h_i + M(1 - \varepsilon_{i,j}), \quad \forall i, j \in V, i \neq j, \quad (10)$$

$$\mu_i \leq \mu_j + M(1 - \varepsilon_{i,j}), \quad \forall i, j \in V, i \neq j, \quad (11)$$

$$\mu_j + l_j \leq \mu_i + l_i + M(1 - \varepsilon_{i,j}), \quad \forall i, j \in V, i \neq j, \quad (12)$$

$$\sigma_{i,j} + \sigma_{j,i} + \pi_{i,j} + \pi_{j,i} + \varepsilon_{i,j} + \varepsilon_{j,i} \geq 1, \quad \forall i, j \in V, i \neq j, \quad (13)$$

$$\sigma_{i,j} + \varepsilon_{i,j} \leq 1, \quad \forall i, j \in V, i \neq j, \quad (14)$$

$$\pi_{i,j} + \varepsilon_{i,j} \leq 1, \quad \forall i, j \in V, i \neq j, \quad (15)$$

$$\theta_i \geq e_i^{arr}, \quad \forall i \in V, \quad (16)$$

$$\mu_i, \theta_i, \alpha_i^+, \alpha_i^-, \beta_i^+, \beta_i^- \geq 0, \quad \forall i \in V, \quad (17)$$

$$\sigma_{i,j}, \pi_{i,j}, \varepsilon_{i,j} \in \{0, 1\}, \quad \forall i, j \in V, i \neq j. \quad (18)$$

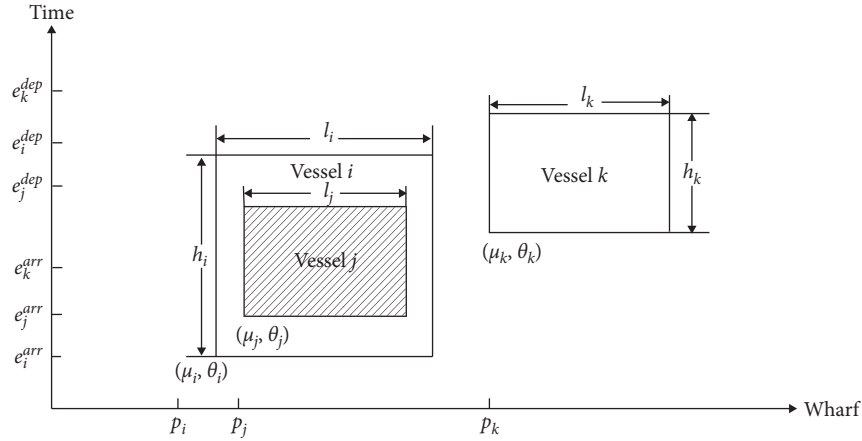


FIGURE 4: Time-wharf section of the DLSM example.

Objective function (2) and constraints (3) and (4) indicate that the objective of this study is to minimize both total additional operation cost corresponding to vessels not berthing at their minimum-cost berthing position vessels and the penalty cost incurred when vessels cannot leave before their preplanned departure time.

Constraint (5) ensures that no more than two vessels can be berthed at the same position simultaneously. Constraint (6) indicates that the rightmost end of each vessel must be limited by the length of the wharf. Constraint (7) ensures the position relationship of two adjacent vessels along the wharf. Constraint (8) indicates the sequential relationship between the berthing time of two vessels that will be berthed at the same position but not in double-line. Constraints (9)–(12) indicate the conditions to be respected when two vessels are berthed in double-line at the berth as detailed in Section 3.4. Constraints (13) ensure that at least one relationship between two vessels waiting to be berthed within the planning period, as shown in Figure 4, holds.

Constraints (14) and (15) ensure the SLSM mode and DLSM mode cannot be applied to the same pair of vessels simultaneously, i.e., if vessel i is berthed in double-line with vessel j , it can neither be berthed to the left of vessel j nor be berthed before the arrival or after the departure of vessel j . Constraint (16) ensures that vessels can only be berthed after their arrival. Constraints (17) and (18) define the range of decision variables.

5. PSO Algorithm for BAP with DLSM Mode

5.1. Introduction to PSO. In PSO algorithms, the particle swarm concept originated as a simulation of a simplified social system by introducing a number of simple entities—the particles—in the search space, where each particle represents a solution approach corresponding to a given position and velocity, which can be used to evaluate the objective function at its current location.

The movement of each particle is guided by their position, according to their own best position and a swarm's best position, which represents the quality of searching, and the velocity decides the direction in which the particle would

move in the next generation. These particles search for optimal solutions through updating generations. Formulas (19) and (20) represent how velocity and position update in the classical PSO algorithm, respectively:

$$v_{id}^k = v_{id}^{k-1} + c_1 r_1 (pbest_{id} - x_{id}^{k-1}) + c_2 r_2 (gbest_d - x_{id}^{k-1}), \quad (19)$$

$$x_{id}^k = x_{id}^{k-1} + v_{id}^k, \quad (20)$$

where v_{id}^k and v_{id}^{k-1} represent the current and previous flight velocity of particle i on dimension d in iteration k , respectively, x_{id}^k and x_{id}^{k-1} represent the current and previous position of particle i on dimension d in iteration k , ω^{k-1} is the inertial weight coefficient, which can adjust the search range of solution space, c_1 and c_2 are acceleration weights, which adjust the learning maximum step length, r_1 and r_2 are two random functions with a value range of $[0, 1]$, whose function is to increase the randomness of the search. $pbest_{id}$ denotes the best position of particle i on dimension d up to iteration k , while $gbest_d$ denotes the best position of the whole swarm on dimension d until iteration k .

Considering that the classical PSO algorithm mentioned above may lead the particles to grow unlimitedly, which influences the particles' convergence to the optimal solution, Shi and Eberhart [25] improved the updating mechanism, by introducing an inertia weight coefficient, which can be dynamically adjusted to balance the quality of solution and convergence velocity of the algorithm, as shown in the following formula:

$$v_{id}^k = \omega^{k-1} v_{id}^{k-1} + c_1 r_1 (pbest_{id} - x_{id}^{k-1}) + c_2 r_2 (gbest_d - x_{id}^{k-1}), \quad (21)$$

where $\omega^{k-1} = [(\gamma_{\max} - \gamma_{k-1}) / \gamma_{\max}] * (\omega_{\max} - \omega_{\min}) + \omega_{\min}$, which is the inertia weight coefficient, ω_{\max} and ω_{\min} denote the maximum and minimum values of the inertia weight coefficient, respectively, and γ_{\max} represents the maximum number of iterations.

The PSO algorithm proposed in this study is based on the updating mechanism proposed by Shi and Eberhart (1998).

5.2. *Encoding.* Assuming that n vessels are waiting to be scheduled within the planning period, n random numbers, denoted as τ_i , $i \in \{1, \dots, n\}$, are randomly generated in the range of 0 and 1.0, where each random number corresponds to the vessel with the same index.

Sort those generated random numbers in descending order and then allocate the corresponding vessels to berth positions one after another, i.e., the greater the random number τ_i is, the earlier vessel i is allocated to a berth position. Ties are broken by selecting the vessel with smallest index.

For a better understanding, here illustrated in Figure 5 is the encoding process with an example of 5 vessels, where a solution with the corresponding berthing order of the vessels as 5-3-1-4-2 is obtained.

5.3. *Decoding.* The decoding process, which is applied in this study to construct the berth schedule corresponding to a given solution obtained by the proposed PSO algorithm, consists of three steps as follows:

- (i) Step 1: initialization of the berthing schedule

The initial berth schedule can be generated by arranging each vessel one after another in the order defined by the solution to its minimum-cost berthing position. It is worth mentioning that although placing vessels to their pre-defined minimum-cost berthing positions can avoid additional operation costs, it is hardly possible for berth operators to arrange all the vessels to their minimum-cost berthing positions without overlapping any of them at a busy port. In consequence, there is a good chance that the berth schedule obtained at this step is infeasible due to the overlaps, and therefore, actions have to be taken to detect and resolve possible overlaps.

- (ii) Step 2: detection of overlaps

Considering that the overlap between two vessels takes place if and only if both berthing periods and spaces of these two vessels are partly overlapped; the overlap between two vessels j and k ($j, k \in V$) can be detected by verifying constraints (22)–(25); in Figure 6, an example of three vessels with overlap detected between two vessels j and k is shown:

$$\mu_k < \mu_j + l_j, \quad k, j \in V, \quad (22)$$

$$\theta_k < \theta_j + h_j, \quad k, j \in V, \quad (23)$$

$$\mu_j < \mu_k + l_k, \quad k, j \in V, \quad (24)$$

$$\theta_j < \theta_k + h_k, \quad k, j \in V. \quad (25)$$

- (iii) Step 3: overlaps resolving

Once overlaps are detected, the current berth schedule is not yet feasible, and thus actions must be taken to remove those overlaps. The procedure

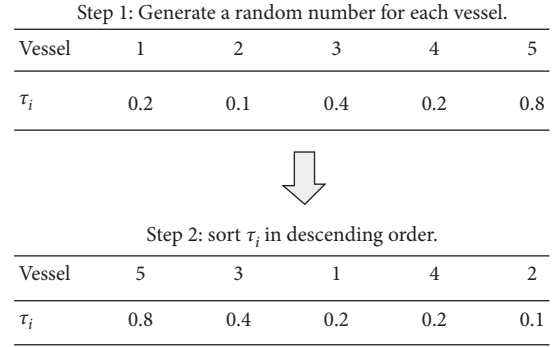


FIGURE 5: Schema of the encoding process.

resolving overlaps between two vessels j and k ($j, k \in V$) is as follows:

Step 3.1: removing overlap detected between two vessels

In this study, the overlap detected between two vessels is eliminated by fixing one vessel and moving the other one towards all possible directions until no overlap is observed between them. Here shown in Figure 7 is an example with two overlapped vessels, which are represented with solid line rectangles. Let vessel j be fixed and vessel k can be moved towards four possible directions to eliminate the overlap: (i) left (in condition that the leftmost end of vessel k does not exceed the left end of the wharf), (ii) right (in condition that the rightmost end of vessel k does not exceed the right end of the wharf), (iii) up (to delay its berthing time), and (iv) outside (in condition that the DLSM constraints are satisfied). The possible positions of vessel k after performing these movements are mentioned with dashed line rectangles, and the rectangle corresponding to the outside movement is shaded on this time-wharf section.

Upon further analysis of the four movements mentioned above, it can be observed that only movement (iii) can result in a feasible solution because (1) movement (i) is not available because there is not enough space on the left (dashed rectangle exceeds the left boundary of the wharf); (2) movement (ii) introduces an overlap between vessel k and vessel i ; (3) movement (iv) is not available as well because the constraints related to the DLSM mode, as described in Section 4, cannot be satisfied. Nevertheless, the feasibility of the solution obtained by (ii) can be improved by taking into account the relationship between the vessel being moved, i.e., vessel k , and the nearby vessels that may be overlapped by the newly placed vessel k , e.g., vessel i in the example shown in Figure 7.

Step 3.2: improving the feasibility of berthing schedule by taking into account the nearby vessels having overlaps with certain moved vessel

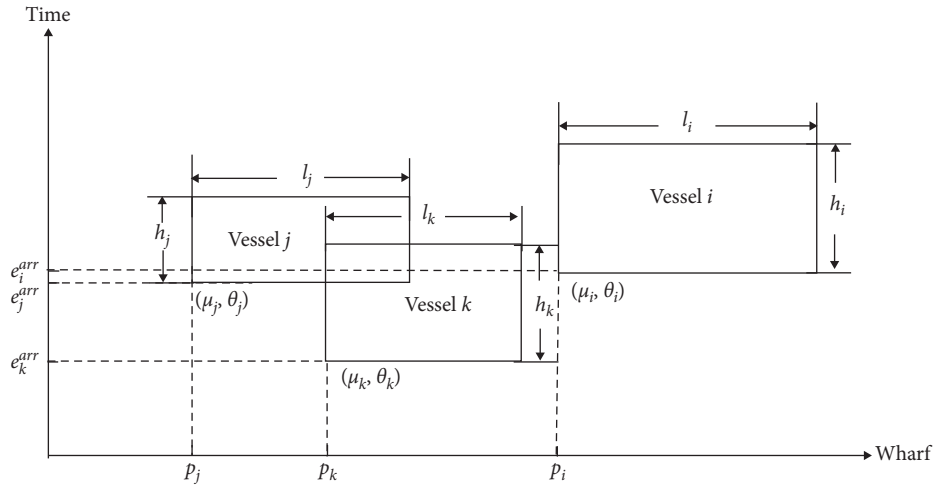


FIGURE 6: An example of three vessels obtained at step 1 with overlap detected.

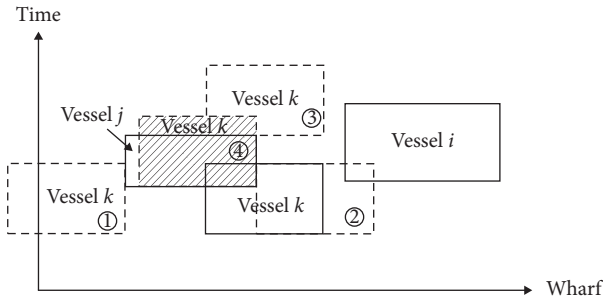


FIGURE 7: Illustration of possible movements made to remove the overlap between two vessels in time-wharf section.

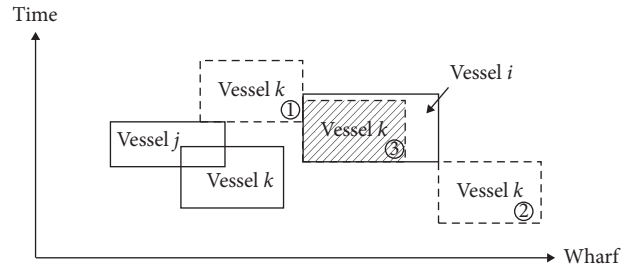


FIGURE 8: Possible movements of vessel k around the nearby vessel i .

Since it is possible to introduce new overlaps between the vessel being moved and some of the nearby ships, the relationship of all vessels that may have overlaps with the newly placed vessel must be considered to avoid introducing new overlaps.

For a better understanding, let us continue the illustration with the example mentioned in step 3.1. Since moving vessel k towards right may introduce an overlap between the vessels k and i , the movements of vessel k around vessel i are also considered to generate possible feasible berthing schedules. As shown in Figure 8, three new berthing schedules can be obtained by finding the optimal position of placing vessel k adjacent to vessel i in condition that it does not overlap with any other vessels. It is worth noting that the berthing schedule corresponding to the optimal position above vessel i is not shown in Figure 8 because that berthing schedule can be dominated by at least the one with vessel k on lower left, i.e., the berthing schedule corresponding to movement (i) shown in Figure 8.

Step 3.3: accepting the best feasible berthing schedule

Compare all of the possible feasible berthing schedules generated by the adjustments described

in steps 3.1 and 3.2 and accept the best one, i.e., the feasible berthing schedule with the smallest objective value as the one that corresponds to the given solution obtained by the proposed PSO algorithm.

5.4. General Procedure of the Proposed PSO Algorithm. The general procedure of the proposed PSO algorithm is as follows.

- (i) Step 1: set up the parameters of the PSO algorithm, such as the number of particles and the value of inertia weight coefficient.
- (ii) Step 2: initialize the position and velocity in allowable ranges for each particle and set iteration $k = 1$.
- (iii) Step 3: calculate the fitness value, which is equal to the objective value of the proposed model, for each particle.
- (iv) Step 4: set the local-best value and global-best value for each particle, where the former equals the particle's current position and the latter the position of the best particle.
- (v) Step 5: update the velocity and the position for each particle.
- (vi) Step 6: update the fitness value for each particle.

- (vii) Step 7: compare the current fitness value of each particle with the local-best one. If the current fitness value of a particle is better, update the local-best position of this particle; otherwise, it remains unchanged.
- (viii) Step 8: find out the particle with the best fitness function from the current swam. If the current best fitness value is better than that of the recorded global-best one, replace the global-best position with the position of the current best particle; otherwise, the global-best one remains unchanged.
- (xi) Step 9: if the number of iteration k attains the predefined threshold, the PSO algorithm terminates and reports the recorded global-best particle as the final solution; otherwise, set $k = k + 1$ and return to step 3.

6. Experimental Results

6.1. Experimental Settings. In this study, instances of different scales are randomly generated with the method introduced by Park and Kim [15]. The length of wharf is set as 1200 meters. The planning horizon T is set as 120 time units, where the time unit is one hour.

The cost coefficients c_{1i} and c_{2i} are set as 2 and 10, respectively, as proposed by Meisel and Bierwirth [26]. In order to ensure that most of the vessels can leave the port before their preplanned departure time, the value of the preplanned departure time of a vessel is determined by adding 1.0 to 2.0 times of the corresponding operation time to its preplanned arrival time, i.e., $e_i^{dep} = e_i^{arr} + h_i * q$ ($i \in V$) and q is a decimal randomly generated between 1.0 and 2.0, as proposed by Park and Kim [15]. The generation of the other parameters is detailed in Table 1.

The numerical experiments are programmed in C# (VS2017) on a PC with 2.3 GHz Intel Core i5 CPU and 4 GB RAM, and CPLEX 12.5 is applied as the programming solver for small-size instances. Both programming solver and the proposed PSO algorithm are set to terminate within 3 hours (10,800 s).

6.2. Comparison between Different Mooring Modes. First of all, experiments are conducted to compare between two different mooring modes, i.e., DLSM mode and SLSM mode by considering both objective values and execution time for small-size instances, i.e., the instances with up to 25 vessels.

As shown in Table 2, it can be observed that the optimal solutions for both modes can be obtained by CPLEX solver within 5 seconds for the instances with no more than 15 vessels. As for SLSM mode, the execution time used to solve instances with the SLSM mode by the CPLEX solver (as shown in column “CPU1”) increases dramatically when the number of vessels is beyond 20. When the DLSM mode is applied, solution for instances with up to 20 vessels can be obtained within 10 seconds and the instances with 25 vessels can still be obtained within 30 minutes (as shown in column “CPU2”). As shown in column “Diff_CPU,” the different rate of the execution time ($\text{Diff_CPU1} = (\text{CPU2} - \text{CPU1}) /$

TABLE 1: Parameters used in the experiments.

Parameter	Distribution type	Range
e_i^{arr}	Uniform distribution	$U(1, 96)$
h_i	Uniform distribution	$U(10, 24)$
p_i	Uniform distribution	$U(1, 1200)$
l_i	Uniform distribution	$U(150, 350)$

$\text{CPU1} * 100\%$) varies from -79.16% to -98.87% for the instances with 20 and 25 vessels, and it is reasonable to conclude that the application DLSM mode can greatly improve the work efficiency of port operators.

With regard to the objective values, it can be observed that the DLSM mode obviously dominates the SLSM mode because the objective values of solutions with the DLSM mode (shown in column “OBJ2”) are at least as good as those with the SLSM mode (shown in column “OBJ1”). According to the difference rates shown in column “Diff_Obj1” ($\text{Diff_Obj1} = (\text{OBJ2} - \text{OBJ1}) / \text{OBJ1} * 100\%$), operation costs can be reduced in average of 20.35% and the maximum reduction rate reaches 37.14%.

To sum up, it can be concluded that DLSM mode can help the port operators in not only improving their work efficiency but also reducing overall operation costs.

It should also be mentioned that the optimal solutions cannot be obtained by CPLEX solver within 3 hours for the instances with more than 25 vessels for neither of these two modes. Therefore, we can conclude that CPLEX solver is only effective for solving small-scale problems regardless of whether DLSM is applied, and thus it is necessary to develop efficient heuristics to obtain good quality solution within reasonable execution time for large-scale instances so as to cope with the real requirements of the huge terminal containers such as Yangshan port.

6.3. Comparison between Different Methodologies. As mentioned before, the CPLEX solver is just capable of solving BAP models for small-scale instances with both SLSM and DLSM modes, though much more vessels must be scheduled during even 120 hours. Thus, in this study, a PSO algorithm has been proposed to obtain good quality solutions within reasonable execution time for large-scale instances.

As shown in Table 3, when comparing the solutions obtained by CPLEX solver and the proposed PSO algorithm for instances with DLSM modes, we can observe that both CPLEX solver and the proposed PSO algorithm can get the final solution very quickly for the instances within 20 vessels. As for instances with more vessels, the CPLEX solver becomes more and more inefficient and cannot obtain optimal solutions within three hours for instances with beyond 30 vessels, though PSO can still get final solution within several minutes.

With regard to objective values, the proposed PSO can obtain optimal solutions for the instances with 8 vessels and most of the instances with 10 vessels and even one instance with 20 vessels; near-optimal solutions can be obtained for the rest of the instances with 10 vessels and most of the cases with 15 vessels and even most of the cases with 25 vessels

TABLE 2: Comparison between DLSM and SLSM modes for small-scale instances.

Instances	SLSM		DLSM		Diff_Obj1 (%)	Diff_CPU1 (%)
	OBJ1	CPU1 (s)	OBJ2	CPU2 (s)		
8-1	300	0.2	252	0.2	-16.00	0.00
8-2	324	0.3	304	0.2	-6.17	-33.33
8-3	488	0.3	488	0.4	0.00	33.33
10-1	430	0.3	430	0.5	0.00	66.67
10-2	656	0.4	656	0.3	0.00	-25.00
10-3	754	0.6	474	0.5	-37.14	-16.67
15-1	2642	4.4	2020	2.3	-23.54	-47.73
15-2	804	2.2	544	2.3	-32.34	4.55
15-3	940	0.5	940	1.1	0.00	120.00
20-1	2816	733.2	1842	8.3	-34.59	-98.87
20-2	3574	205.2	2604	7.7	-27.14	-96.25
20-3	1932	116.4	1232	3.6	-36.23	-96.91
25-1	4532	3630	3384	668.7	-25.33	-81.58
25-2	6886	7245.2	4616	1510.1	-32.97	-79.16
25-3	4440	7280.1	2936	1421.0	-33.87	-80.48
Average					-20.35	-28.76

TABLE 3: Comparison between the performance of CPLEX and PSO algorithm for solving problems with the DLSM mode.

Instances	CPLEX		PSO		Diff_Obj2 (%)
	OBJ2	CPU2 (s)	OBJ3	CPU3 (s)	
8-1	252	0.2	252	8.2	0.00
8-2	304	0.2	304	4.4	0.00
8-3	488	0.4	488	9.4	0.00
10-1	430	0.5	430	10.1	0.00
10-2	656	0.3	656	8.0	0.00
10-3	474	0.5	484	16.2	2.11
15-1	2020	2.3	2052	40.7	1.58
15-2	544	2.3	574	41.3	5.51
15-3	940	1.1	1066	23.6	13.40
20-1	1842	8.3	2102	73.7	14.12
20-2	2604	7.7	2974	73.9	14.21
20-3	1232	3.6	1232	73.6	0.00
25-1	3384	668.7	3904	103.3	15.37
25-2	4616	1510.1	4854	122.8	5.16
25-3	2936	1421.0	3022	111.3	2.93
30-1	Cannot	>3 h	7566	326.3	—
30-2	obtain	>3 h	6820	251.1	—
30-3	the	>3 h	8248	211.4	—
35-1	optimal		7898	272.2	—
35-2	solution		12006	273.1	—
35-3			12760	279.5	—
40-1			15800	354.1	—
40-2			15640	565.2	—
40-3			20558	474.2	—
45-1			26150	685.2	—
45-2			35374	744.6	—
45-3			31726	644.2	—

with quite small difference rate, which can be illustrated in column “Diff_Obj2” ($\text{Diff_Obj2} = (\text{OBJ3} - \text{OBJ2}) / \text{OBJ2} * 100\%$). It hints that the proposed PSO algorithm is also possible to get solutions of good quality for large-scale instances though further studies should be made to test the condition of such performance.

TABLE 4: Comparison between optimal solutions with SLSM mode and solutions obtained by PSO with the DLSM mode.

Instances	CPLEX-SLSM		PSO-DLSM		Diff_Obj3 (%)
	OBJ1	CPU1 (s)	OBJ3	CPU3 (s)	
8-1	300	0.2	252	8.2	-16.00
8-2	324	0.3	304	4.4	-6.17
8-3	488	0.3	488	9.4	0.00
10-1	430	0.3	430	10.1	0.00
10-2	656	0.4	656	8.0	0.00
10-3	754	0.6	484	16.2	-35.81
15-1	2642	4.4	2052	40.7	-22.33
15-2	804	2.2	574	41.3	-28.61
15-3	940	0.5	1066	23.6	13.40
20-1	2816	733.2	2102	73.7	-25.36
20-2	3574	205.2	2974	73.9	-16.79
20-3	1932	116.4	1232	73.6	-36.23
25-1	4532	3630	3904	103.3	-13.86
25-2	6886	7245.2	4854	122.8	-29.51
25-3	4440	7280.1	3022	111.3	-31.94
Average					-16.61

Since it is observed in Table 3 that the gap between solutions obtained by the CPLEX solver and the PSO algorithm with the DLSM mode is relatively significant for some of the instances, a further comparison is made between the results obtained by PSO with DLSM mode and the optimal solutions obtained by the CPLEX solver with the SLSM mode.

As shown in Table 4, solutions obtained by the PSO algorithm with DLSM mode are better than the optimal solutions obtained by the CPLEX solver with the SLSM mode, and the former can save up to 35.81% of the cost ($\text{Diff_Obj3} = (\text{OBJ3} - \text{OBJ1}) / \text{OBJ1} * 100\%$) among all the instances tested in this study.

Considering that hundreds of vessels should be operated every day at huge container terminals, the proposed PSO will be much more practical than CPLEX for supporting the decision-making of the port operators to not only improve their working efficiency but also reduce operation costs related to berth scheduling operations.

7. Conclusions and Perspectives

The study aims at minimizing the total operation cost of the continuous berth scheduling problem by taking into account the Double-Line Shipping Mooring (DLSM) mode, where both the additional operation cost for vessels not moored at their minimum-cost berthing position and penalty cost related to vessels not being able to leave before its pre-planned departure time are considered.

The problem is firstly formulated as a mixed integer programming model and solved by the CPLEX solver for small-scale instances. As for larger size instances that cannot be optimally solved by CPLEX solver, a PSO algorithm is proposed to obtain good quality solutions within reasonable execution time.

Numerical experiments are conducted to compare not only the efficiency between the traditional Single-Line Shipping Mooring (SLSM) mode and the innovative DLSM mode but also the performances between CPLEX solver and the proposed PSO algorithm. It can be concluded with the experimental results that (1) DLSM mode outperforms the SLSM mode in reducing not only total operation cost but also execution time. (2) The proposed PSO algorithm can generate optimal or near-optimal solution for small-scale instances. (3) The proposed PSO algorithm is much more efficient than the CPLEX solver for large-scale instances, which copes with the requirements of berthing management in Yangshan Deep-Water Port, one of the busiest container terminals in the world.

To sum up, as the first research dedicated to BAP with DLSM mode, this study can help not only in validating the advantages of DLSM mode but also offering an efficient decision support tool to berth operators in busy ports to improve their working efficiency.

Motivated by the results obtained in this study, it is interesting to keep improving the efficiency of the proposed algorithm and to apply such method in the targeted port.

Data Availability

All the experimental data can be generated with the rules described in the paper.

Conflicts of Interest

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

References

- [1] Q. Meng, S. Wang, H. Andersson, and K. Thun, "Container-ship routing and scheduling in liner shipping: overview and future research directions," *Transportation Science*, vol. 48, no. 2, pp. 265–280, 2014.
- [2] D. Kizilay and D. T. Eliyi, "A comprehensive review of quay crane scheduling, yard operations and integrations thereof in container terminals," *Flexible Services and Manufacturing Journal*, 2020.
- [3] A. Imai, K. I. Nagaiwa, and C. W. Tat, "Efficient planning of berth allocation for container terminals in Asia," *Journal of Advanced Transportation*, vol. 31, no. 1, pp. 75–94, 1997.
- [4] A. Imai, E. Nishimura, and S. Papadimitriou, "The dynamic berth allocation problem for a container port," *Transportation Research Part B: Methodological*, vol. 35, no. 4, pp. 401–417, 2001.
- [5] V. H. Barros, T. S. Costa, A. C. M. Oliveira, and L. A. N. Lorena, "Model and heuristic for berth allocation in tidal bulk ports with stock level constraints," *Computers & Industrial Engineering*, vol. 60, no. 4, pp. 606–613, 2011.
- [6] L. Dai and L. Tang, "Berth allocation with service priority for container terminal of hub port," in *Proceedings of the 2008 4th International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 1–4, Logs Engineering & Management, Dalian, China, October 2008.
- [7] T. Qin, Y. Du, and M. Sha, "Evaluating the solution performance of IP and CP for berth allocation with time-varying water depth," *Transportation Research Part E: Logistics and Transportation Review*, vol. 87, pp. 167–185, 2016.
- [8] L. Zhen, Z. Liang, D. Zhuge, L. H. Lee, and E. P. Chew, "Daily berth planning in a tidal port with channel flow control," *Transportation Research Part B: Methodological*, vol. 106, pp. 193–217, 2017.
- [9] K. H. Kim and K. C. Moon, "Berth scheduling by simulated annealing," *Transportation Research Part B: Methodological*, vol. 37, no. 6, pp. 541–560, 2003.
- [10] B. C. Jos, M. Harimanikandan, C. Rajendran, and H. Ziegler, "Minimum cost berth allocation problem in maritime logistics: new mixed integer programming models," *Indian Academy of Sciences/Sādhanā*, vol. 44, p. 149, 2019.
- [11] L. Zhen, H. Hu, W. Wang, X. Shi, and C. Ma, "Cranes scheduling in frame bridges based automated container terminals," *Transportation Research Part C: Emerging Technologies*, vol. 97, pp. 369–384, 2018.
- [12] E. Lalla-Ruiz, J. L. González-Velarde, B. Melián-Batista, and J. M. Moreno-Vega, "Biased random key genetic algorithm for the tactical berth allocation problem," *Applied Soft Computing*, vol. 22, pp. 60–76, 2014.
- [13] E. Nishimura, A. Imai, and S. Papadimitriou, "Berth allocation planning in the public berth system by genetic algorithms," *European Journal of Operational Research*, vol. 131, no. 2, pp. 282–292, 2001.
- [14] S. R. Seyedalizadeh Ganji, A. Babazadeh, and N. Arabshahi, "Analysis of the continuous berth allocation problem in container ports using a genetic algorithm," *Journal of Marine Science and Technology*, vol. 15, no. 4, pp. 408–416, 2010.
- [15] Y.-M. Park and K. H. Kim, "A scheduling method for berth and quay cranes," *OR Spectrum*, vol. 25, no. 1, pp. 1–23, 2003.
- [16] M. A. Dulebenets, "Application of evolutionary computation for berth scheduling at marine container terminals: parameter tuning versus parameter control," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 1, pp. 25–37, 2018.
- [17] M. A. Dulebenets, "An adaptive island evolutionary algorithm for the berth scheduling problem," *Memetic Computing*, vol. 12, no. 1, pp. 51–72, 2020.
- [18] M. Kavooosi, M. A. Dulebenets, O. Abioye et al., "Berth scheduling at marine container terminals: a universal island-based metaheuristic approach," *Maritime Business Review*, vol. 5, no. 1, pp. 30–66, 2020.
- [19] R. C. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proceeding of the 6th International Symposium on Micromachine and Human Science*, pp. 39–43, Nagoya, Japan, October 1995.
- [20] C.-J. Ting, K.-C. Wu, and H. Chou, "Particle swarm optimization algorithm for the berth allocation problem," *Expert Systems with Application*, vol. 41, no. 4, pp. 1543–1550, 2014.

- [21] L. Zhen, "Modeling of yard congestion and optimization of yard template in container ports," *Transportation Research Part B: Methodological*, vol. 90, pp. 83–104, 2016.
- [22] P. Guo, W. Cheng, and Y. Wang, "A modified generalized extremal optimization algorithm for the quay crane scheduling problem with interference constraints," *Engineering Optimization*, vol. 46, pp. 1411–1429, 2014.
- [23] H.-P. Hsu and C.-N. Wang, "Resources planning for container terminal in a maritime supply chain using multiple particle swarms optimization (MPSO)," *Mathematics*, vol. 8, no. 5, p. 764, 2020.
- [24] M. Zhong, Y. Yang, Y. Zhou, and O. Postolache, "Adaptive autotuning mathematical approaches for integrated optimization of automated container terminal," *Mathematical Problems in Engineering*, vol. 2019, Article ID 7641670, 14 pages, 2019.
- [25] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Proceedings of the IEEE world congress on Computational Intelligence*, pp. 69–73, Anchorage, AK, USA, 1998.
- [26] F. Meisel and C. Bierwirth, "Heuristics for the integration of crane productivity in the berth allocation problem," *Transportation Research Part E: Logistics and Transportation Review*, vol. 45, no. 1, pp. 196–209, 2009.

Research Article

A Comparative Performance Investigation of Swarm Optimisers on the Design of Hydrostatic Thrust Bearing

Ismail Sahin 

Gazi University, Technology Faculty, Department of Industrial Design Engineering, Teknikokullar-06500, Ankara, Turkey

Correspondence should be addressed to Ismail Sahin; isahin@gazi.edu.tr

Received 14 June 2020; Revised 8 July 2020; Accepted 21 August 2020; Published 1 September 2020

Academic Editor: Tingsong Wang

Copyright © 2020 Ismail Sahin. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the design of hydrostatic thrust bearings, power loss that occurs during operation is an important parameter that affects the design, and due to such features, it falls within the interest of design optimisation studies. The fact that the decimal places of the constraints and design variables used for minimum power loss optimisation of hydrostatic thrust bearings are highly effective on the result is a challenge for the design optimisation studies carried out on the problem and has yet made it rather attractive for the researchers. In this study, it is this feature of the problem that makes it the most important motivator in researching the performance of different metaheuristic optimisers in solving the minimum power loss problem. To this end, 7 different optimisers, four of them for the first time, were applied under equal conditions with various pop sizes and a number of iterations, and their performances were addressed under this challenging benchmark problem. The performances of these methods were compared to each other. In addition to the success of optimisers in reaching a solution, their performance in different populations and iterations is also discussed in the study. Considering the results, it is seen that MVO is the most effective optimiser in solving the problem and is followed by the WOA, PSO, and GWO. The application of WOA, MVO, CS, and SSA, for the first time, on the problem has exhibited that these methods could be used in optimisation of such delicate engineering problems.

1. Introduction

Plain bearings are the machine components of choice by virtue of their vibration, impact, and noise damping properties where high speed and load are required, and it is a fact that they have long operating lives. Plain bearings are classified as journal, linear, and thrust bearing, depending on the direction of the load they carry. Thrust bearings can be examined in three groups as hydrostatic, hydrodynamic, and hydrostatic-hydrodynamic bearings [1].

In hydrostatic thrust bearings, pressure is applied to balance the external force and separate the surfaces [2]. With the pressure generated by an external pump, oil is conveyed on the bearings and the bearing surfaces are separated by the oil film. As there is no contact between moving parts, it allows hydrostatic thrust bearings to have lower friction, wear, and vibration [3]. Due to their advantages such as higher accuracy, higher load capacity, incomparable smoothness, higher hardness, and lower friction values that

may be achieved in motion and positioning, hydrostatic thrust bearings are widely used in the industry [4–6]. The power loss that occurs during operation and the increase in oil temperature are considered performance measures in optimum hydrostatic thrust bearing design [7]. For this reason, minimisation of power loss during operation has been a core issue in the studies aimed at the optimisation of hydrostatic thrust bearings.

In the studies carried out for optimum design of hydrostatic thrust bearings, metaheuristic methods are heavily used [1, 2, 8–13]. Studies conducted on this subject are recently focused on swarm intelligence methods. Swarm intelligence algorithms mimic the intelligence of swarms, herds, or flocks of creatures in nature [14].

In this study, the power loss minimisation problem of hydrostatic thrust bearings defined by Siddall [12] was attempted to be solved through 7 different optimisation methods. The performances of the swarm intelligence approaches used in the study for the minimisation problem of

power loss during operation of hydrostatic thrust bearings were comprehensively discussed. The problem is interesting, compelling, and yet utterly attractive by reason of the fact that the decimal places of the active constraints and design variables are highly sensitive on the results. This has made it a rather good benchmark problem [11]. This structure of the problem makes it difficult to achieve feasible results. In the study, more delicate search was carried out in the search space to overcome this difficulty. The challenging structure of the problem was an important motivator in the research of the performance of different metaheuristic optimisers to produce optimum solutions. With this motivation, the problem was solved through popular swarm intelligence optimisers such as particle swarm optimisation (PSO), multiverse optimiser (MVO), grey wolf optimiser (GWO), cuckoo search (CS), whale optimisation algorithm (WOA), salp swarm algorithm (SSA), and artificial bee colony (ABC). Out of these, MVO, CS, SSA, and WOA methods have been applied to the problem for the first time. The performances of these methods were compared with each other. Another important goal of the study is to investigate the impact of population size and number of iterations on the solution. For this reason, the population size was selected as 100, 400, and 800 and the number of iterations was selected as 100, 1000, and 5000.

The remainder of this article is organised as follows: in Section 2, a literature search in which the studies on the subject are examined is presented; in Section 3, the problem is thrust bearing; in Section 4, optimisation methods used in experimental studies are introduced; in Section 5, comparisons of experimental results obtained with optimisers are explained; and Section 6 includes the conclusion and suggestions for future studies.

2. Related Works

The advantages of hydrostatic thrust bearings date back to 1940s. Automatic fluid pressure balancing system of Hoffer is considered the pioneer of hydrostatic thrust bearings [15]. Over the next half-century, a series of patent and research studies carried out covering topics such as the structure of hydrostatic thrust bearings, oil flow, balance problems, and infinite stiffness and have formed the basis of modern hydrostatic thrust bearing designs [16–21]. The hydrostatic system developed by Slochum et al. for use in high-pressure press benches was the first hydrostatic bearing [20]. In this new system, the support equipment used for the shaft's bearing had high strength and friction resistance.

In the subsequent studies, it was observed that oil film thickness, recess pressure, pressure distribution, and oil flow rate were effective on the performance of hydrostatic bearings [22]. These characteristics affect high wear resistance, which is the most important feature expected from thrust bearings [23]. Bearing materials are expected, in their selection, to have features such as lower friction coefficient, higher wear resistance, higher loading capacity, better erosion resistance, better thermal conductivity, and lower thermal expansion. However, oil viscosity, oil film thickness, oil flow rate, and pressure amount also have significant effects on the

performance of the system. In addition to experimental and analytical studies aimed at achieving the optimum values of these features, computer aided studies date back to early 1980s. Siddall's study, in which the problem of minimising power loss during operation of hydrostatic thrust bearings is described, is one of the pioneering studies in this respect [12]. Siddall reduced the power loss in existing bearings by 2,288.0 ft-lb/s (4.14 hp) through the Hooke and Jeeves (HJ) method.

Siddall's problem has become a benchmark problem used for testing intelligent optimisation approaches today. Among them, population-based approaches stand out. These studies used the minimisation of the power loss of hydrostatic thrust bearings extensively, inspired by Siddall's identified problem, in testing the metaheuristic methods they developed. Deb and Goyal's genetic adaptive search (GeneAS) method is the first of these [9]. Deb and Goyal demonstrated that the thrust bearing they designed could withstand more weight with smaller film thickness and would exhibit lesser power loss as per the algorithm of Siddall. Afterwards, Coello achieved faster and better results using a multipurpose optimisation technique in place of the penalty functions used in GA [8]. Another study focused on inconsistencies in previous studies with regard to unit and design criteria, using the particle swarm optimisation method [10]. Improved PSO algorithm of He et al. reduced power loss by around 30% compared to Siddall's study. Rao et al. achieved good results through the teaching-learning-based optimisation (TLBO) method they developed, compared to the studies in the literature [11]. Kentli and Sahbaz have solved the problem through sequential quadratic programming (SQP) method and compared the results with Siddall's study [1]. Sahin et al. [2] have developed a mathematical model of GWO, a highly popular population-based approach in recent years, and solved the problem through a new model (enhanced GWO) that increases the diversity of the valid solutions by keeping the search area wider [2]. In the study, all the studies in the literature were examined together, and a comprehensive comparison was made and very successful results were obtained through E-GWO. Talatahari and Azizi solved the problem through chaos game optimiser in an up-to-date study [13]. Best fitness results and design variable values obtained in studies that address the minimum power loss problem are presented in Table 1.

The optimisation of features such as maximum load-carrying capacity of bearings [24–26], stiffness [26], film thickness [7], and optimal recess shape [4, 27] is among other important research topics in hydrostatic thrust bearings.

3. Hydrostatic Thrust Bearing Design for Minimum Power Loss

The hydrostatic thrust bearing design problem discussed as part of this study was defined by Siddall [12]. The purpose function of the problem in which minimisation of power loss is targeted during the operation of a thrust bearing exposed to the axial load seen in Figure 1 is formulated in the following equation [12]:

TABLE 1: Algorithms and its optimal solutions for minimum power loss problem in the literature.

	Variables				Best fitness $f(x)$
	R	R_0	μ ($\times 10^{-6}$)	Q	
Siddall [12], HJ	7.1550805	6.688682	8.320765	9.168461	29221.321
Deb and Goyal [9], GeneAS	6.778	6.234	6.096	3.809	25937.058
Deb and Goyal [9], BGA	7.7077	6.549	6.619	4.849	27554.5428
Coello [8], GASO	6.271	12.901	5.605	2.938	23403.432
He et al. [10], PSO	5.956868685	5.389175	5.402133	2.301546	19586.5788
Rao et al. [11], TLBO	5.9557805	5.389013	5.358697	2.269655	19505.3131
Kentli and Sahbaz [1], SQP	5.95580	5.38904	8.63332	8.00001	26114.545
Şahin et al. [2], GWO	5.9576206	5.390681	0.000005	2.278085	19530.6552
Şahin et al. [2], E-GWO	5.9557805	5.389013	5.358697	2.269656	19505.3134
Talatahari and Azizi [13], CGO	5.963440023	5.395587861	5.36	2.264822	19454.9541

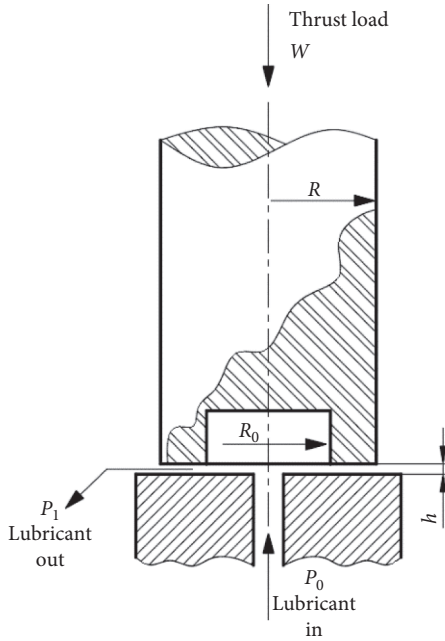


FIGURE 1: Hydrostatic thrust bearing.

$$\text{minimise: } f(x) = \frac{QP_0}{0.7} + E_f. \quad (1)$$

Four design variables were used in the problem [12]: flow rate (Q), recess radius (R_0), bearing step radius (R), and viscosity (μ) (equation (2)).

Design vector:

$$X = (Q, R_0, R, \mu). \quad (2)$$

The value ranges of the design variables of the problem are as follows:

$$\begin{aligned} 1,000 &\leq R \leq 16,000, \\ 1,000 &\leq R_0 \leq 16,000, \\ 1.0 \times 10^{-6} &\leq \mu \leq 16 \times 10^{-6}, \\ 1,000 &\leq Q \leq 16,000. \end{aligned} \quad (3)$$

Seven nonlinear constraints were identified in the problem (equations (4)–(10)). These are minimum load-carrying capacity, weight capacity, inlet oil pressure, oil

temperature rise, oil film thickness, step radius, exit loss, and contact pressure [2, 10, 12]. Out of 7 constraints, 6 are active constraints considering an accuracy of 3 decimal places, and all the design variables are highly sensitive in the optimisation problem subject to

$$g_1(x) = W - W_s \geq 0, \quad (4)$$

$$g_2(x) = P_{\max} - P_0 \geq 0, \quad (5)$$

$$g_3(x) = \Delta T_{\max} - T_0 \geq 0, \quad (6)$$

$$g_4(x) = h - h_{\min} \geq 0, \quad (7)$$

$$g_5(x) = R - R_0 \geq 0, \quad (8)$$

$$g_6(x) = 0.001 - \frac{\gamma}{gP_0} \left(\frac{Q}{2\pi Rh} \right) \geq 0, \quad (9)$$

$$g_7(x) = 5000 - \frac{W}{\pi(R^2 - R_0^2)} \geq 0, \quad (10)$$

where $g_1(x)$ is for weight capacity, which must be greater than weight of generator, $g_2(x)$ is for inlet oil pressure, $g_3(x)$ is for oil temperature rise, $g_4(x)$ is for oil film thickness, $g_5(x)$ is for step radius and must be greater than recess radius, $g_6(x)$ is for limits on significance off exit loss and must be greater than 0001, and $g_7(x)$ is the limit for contact pressure and must be greater than 5000.

The parameters given in the constraints are calculated by the following equations:

$$\begin{aligned} W &= \left(\frac{\pi P_0}{2} \right) \left[(R^2 - R_0^2) \left(\ln \frac{R}{R_0} \right) \right], \\ P_0 &= \frac{6\mu Q}{\pi h^3} \ln \frac{R}{R_0}, \\ E_f &= 9336Q\gamma C\Delta T, \end{aligned} \quad (11)$$

$$\Delta T = 2(10^P - 560),$$

$$P = \frac{\log_{10} \log_{10} (8.122e6\mu + 0.8) - C_1}{n},$$

where W is the load-carrying capacity, P_0 is the inlet pressure, E_f is the friction loss, ΔT is the temperature, γ is the weight density of oil (0.0307 lb/in^3), C is the specific heat of oil ($0.5 \text{ Btu/lb}^\circ\text{F}$), n and C_1 are the oil constants, and h is the film thickness. $C_1 = 10.04$ and $n = -3.55$ were chosen for SAE 20 grade oil (see Table 2).

Other specifications of the design problem are as follows: W_s (weight of generator) = 101000 lb (45804.99 kg); P_{\max} (maximum pressure available) = 1000 psi ($6.89655 \times 10^6 \text{ Pa}$); ΔT_{\max} (maximum temperature rise) = 50°F (10°C); h_{\min} (minimum oil thickness) = 0.001 in. (0.00254 cm); $g = 32.3 \times 12 = 386.4 \text{ in./seg}^2$ (981.465 cm/seg^2), and N (angular speed of shaft) = 750 RPM.

4. Optimisation Methods Applied to the Problem

In this study, seven population-based optimisation methods were applied to the problem. These are PSO, ABC, GWO, MVO, CS, WOA, and SSA. MVO, CS, WOA, and SSA optimisation methods were applied to the problem for the first time.

4.1. Salp Swarm Algorithm (SSA). The salp swarm algorithm (SSA) is an optimisation algorithm developed, inspired by the swarming mechanism of salps resembling jelly fish [14]. The inspiration for the method is the swarming behaviour that salp chains exhibit while searching and collecting in the ocean. In the SSA swarm model, the leading salp moving towards the food source is followed by the follower salps. If the food source is replaced with a global optimum, the salp chain automatically moves towards it. The results in mathematical functions exhibit that the SSA algorithm can effectively develop the initial random solutions and zoom closer to the optimum level [28].

4.2. Cuckoo Search (CS). CS is a metaintuitive search algorithm inspired by the reproductive strategy of cuckoo birds developed by Yang and Deb [29]. The method is based on some cuckoo species placing their eggs in the nests of other cuckoos or different species. In this method called the brood parasitism method, the cuckoo that desires to increase the chances of its own eggs may destroy other eggs, engage directly with other birds, or leave the nest [30]. Yang and Deb have set three rules to use this to solve optimisation problems [29]:

- (i) Each cuckoo bird leaves one egg at a time in a randomly selected nest.
- (ii) The best nests with high quality eggs are conveyed onto future generations.
- (iii) The number of host nests is constant. Likelihood of the landlord finding, a foreign egg is $pa \in [0, 1]$. In this case, the host bird can throw the egg or leave the nest to build a new nest.

4.3. Multiverse Optimiser (MVO). Developed by Mirjalili and his colleagues, MVO is inspired by the theory of multiverse and big bang theory in physics [31]. In the

TABLE 2: Values of n and C_1 for various grades of oil.

Oil	C_1	N
SAE 5	10.85	-3.91
SAE 10	10.45	-3.72
SAE 20	10.04	-3.55
SAE 30	9.88	-3.48
SAE 40	9.83	-3.46
SAE 50	9.82	-3.44

method in which the mathematical model of white hole, black hole, and worm hole concepts in cosmology is created, the search process is a two-step process, as in other population-based methods. But here, white hole and black hole concepts are used in the discovery of search spaces. Worm hole supports MVO in exploiting the search spaces. In MVO, each solution is assumed to resemble a universe, and every variable in the solution is an object in this universe. Another difference of MVO compared to other algorithms is the use of the concept of time, which is a common term in multiverse theory and cosmology [31].

4.4. Artificial Bee Colony (ABC). ABC is a swarm intelligence-based algorithm developed by Karaboga [32]. The algorithm was inspired by the behaviour of honey bees that collect nectars. Bees are divided into three groups: scout, onlooker, and employed bees. Worker bees seek food in nature with a special dance (the waggle dance) and communicate with each other. The onlookers watch the dance to make a choice and follow the bee to the food source [33]. Scooters discover abandoned food sources and replace them with new resources. In this method, the location of food refers to the possible solution to the optimisation problem and the amount of nectar in food refers to the suitability of the solution [34].

4.5. Particle Swarm Optimisation (PSO). PSO, one of the most popular approaches to swarm intelligence, was developed by Kennedy and Eberhart [35]. The algorithm is inspired by the foraging and navigation capabilities of bird flocks. The algorithm consists of particles that are placed in a search space and move around by combining their own previous location and the current global optimal solution. Potential solutions are encoded as randomly initiated particles and directed to move within the search space to find the most appropriate solution [36]. In PSO, the global optimum solution is searched taking into account the speed and location of each particle that makes up the flock [34].

4.6. Whale Optimisation Algorithm (WOA). WOA, which is based on the modelling of the hunting behaviour of humpback whales, was developed by Mirjalili and Lewis [33]. Humpback whales exhibit a three-stage behaviour, namely, search for prey, encircling prey, and bubble-net foraging, as they hunt. In WOA, this behaviour is mimicked.

The method uses updating whales (individuals) their position based on the position of the prey as base as in GWO.

4.7. Grey Wolf Optimiser (GWO). GWO was developed by Mirjalili and Lewis [37], inspired by the hunting method of grey wolves in nature and the social hierarchy within the pack. In GWO, each solution in the population corresponds to a wolf in the pack. There are four different hierarchical levels in the pack, which are called alpha, beta, delta, and omega. Mathematical model of GWO is constituted by this hierarchical structure and hunting method [38].

5. Results and Discussion

This study presented is aimed at minimising the power loss during operation of hydrostatic thrust bearings formulated by Siddall [12]. Siddall's ADRANS algorithm was based on the Hooke and Jeeves (HJ) pattern search method for the solution. In the study, the achievements of SSA, MVO, CS, GWO, ABC, WOA, and PSO methods were tested in solving the problem using the purpose function, design constraints, design vector, and parameters defined by Siddall.

In the first phase of experimental studies, statistical achievements of the algorithms were evaluated. Algorithms were run on a computer with an Intel® Core™ i7, 2.6 GHz CPU, and 8 GB RAM, running on the Windows 10 × 64 operating system using Python 3.7 software. In order to investigate the impact of population size and the number of iterations on the solution, the population was selected as 100, 400, and 800 in all algorithms and the number of iterations, defined as the stop criteria, was set to 100, 1000, and 5000. The algorithms were run 20 times for each pairwise of population size and number of iterations.

Siddall used the inches-pounds per second as unit of fitness value in his study. However, some studies were confirmed to use feet-pounds per second [9, 11, 13]. In this study, calculations were made in the units used by Siddall, sticking to the original state of the problem.

Each algorithm was run 20 times, and best, worst, average, and success percentage (SP) values of the results were calculated. SP takes the results obtained after running of the algorithm in different numbers and demonstrates the stability of the algorithm. In this context, the difference between the best value (F) achieved and the global optimum value achieved must be less than 0.1% of the global optimum (F^*) value (equation (12)), in order for the result to be considered successful:

$$\left| \frac{F^* - F}{F^*} \right| < 1E - 3. \quad (12)$$

5.1. Statistical Comparison of the Algorithms. The study aims to observe the impact of different population sizes and number of iterations on the achievement of the optimisers. Therefore, the performance of the algorithms is with 100, 400, and 800 population sizes and numbers of iterations is 100, 1000, and 5000 (Table 3). As seen in Table 3, the lowest objective function value in the study was achieved by the

MVO with a population of 800 and in 5000 iterations, as $f(x) = 19508,39014$. The results obtained by other algorithms, except for the ABC, are also quite successful. After MVO, the lowest objective function value was achieved by WOA, PSO, and GWO. The power loss value obtained with WOA is only 0.003% behind of MVO. SSA, another algorithm that is applied on the problem, has, however, achieved worse values than that of MVO by as much as 0.88%. This demonstrates that MVO, WOA, and SSA are rather successful in solving the power loss minimisation problem. Worst objective function values were materialised in ABC and CS as $f(x) = 22830,65571$ and $f(x) = 19950,96346$, respectively. With PSO, the third lowest objective function value was achieved following MVO and WOA. This result obtained is better than results of previous PSO study [10].

The performance of algorithms in solving the problem with different population sizes and in iterations differs. PSO, for example, performed poorly in low numbers of iteration. As can be seen from the table, performance of PSO in lower population sizes (100 and 400) and lower iteration numbers (100 and 1000) is worse compared to the population size and number of iterations where it achieved the best fitness. MVO has achieved the best objective function results in all populations and iterations except 100 populations and 5000 iterations. MVO has generated the best results at 5000 iterations for the whole population. This exhibits that MVO performs decisively. Algorithms other than ABC and CS exhibited closer performance at best fitness values at 400 and 800 populations and 1000 iterations. All algorithms except SSA and CS yielded the most successful fitness values during the experiments conducted with a population size of 800 and 5000 iterations.

As the number of iterations increased, success also increased in algorithms. It is not, however, possible to come up with the same conclusion for the results related to increase in population size. For instance, SSA, which acquired the lowest power loss in population size of 400, did not attain the same performance as the population size increased. The solution CS achieved with a population size of 800 is worse than what it achieved with a population size of 400. This may, consequently, indicate that the number of iterations is more effective on the solution than the population size.

In Table 4, the statistical results of algorithms were compared with each other. These values are the values obtained in a population size of 800 in 5000 iterations, where the best fitness value is achieved. In Table 4, the best optimal power loss value is seen to be achieved by MVO. In parallel with this, it can be seen that MVO is more successful, compared to other studies in terms of SP, mean, and worst values. MVO exhibited a steady achievement in terms of success percentage. MVO, with an SP rate of 1, is followed by PSO, GWO, and WOA with an achievement rate of 0.85, 0.40, and 0.15, respectively. However, WOA has the second-best fitness value, and the SP rate occurs to be 0.15. SP rates of SSA, ABC, and CS were 0.00.

In Table 5, the values of design variable used for solutions that occur in 800 populations and 5000 iterations, where the best objective function value is achieved, are given comparatively. When the table is analysed, it will be seen that the

TABLE 3: Best, worst, and mean fitness results of the algorithms at different populations and iteration sizes.

Algorithms	Population size # iterations	100			400			800		
		Worst	Mean	Best	Worst	Mean	Best	Worst	Mean	Best
ABC	100	97880,22377	59756,19037	35199,122	61438,2	38997,04537	23972,79965	45816,66109	35635,12678	25911,96344
	1000	42477,76908	33791,24852	25929,378	36734,93	30543,36807	25276,78226	37473,99432	30138,08625	23780,59516
	5000	33425,55472	28379,06693	24551,577	30956,61847	26934,92081	23447,8621	26339,46279	24666,75976	22830,65571
CS	100	86472,45356	52863,75942	37332,47	73568,12	46060,36489	22864,9523	49565,0751	37655,52983	29197,89212
	1000	33805,431	28370,40843	24409,907	29135,1	26901,81866	22247,30483	28196,3445	26395,48853	23922,48927
	5000	21419,78879	20663,07987	20165,94	20842,00943	20455,23573	19926,80827	20466,14966	20297,33609	19950,96346
GWO	100	96943,09892	49335,12227	25103,291	49501,18	29989,26663	21571,10275	46829,225	26831,94961	20747,87574
	1000	80124,80937	51458,42565	23085,525	28241,92	20625,69807	19629,68579	20070,02109	19742,61778	19562,81207
	5000	36928,31868	24291,30478	19586,232	19673,73006	19575,76898	19542,05678	19556,18396	19533,09898	19516,83771
MVO	100	36290,93605	28601,01781	24000,85	33412,84	26288,74334	20299,44124	31998,05736	24091,47227	19827,22283
	1000	28544,07865	21942,57516	19605,425	20334,92	19739,74148	19537,59644	20306,92897	19632,75373	19517,10515
	5000	25179,91496	19985,30306	19529,24	19937,75722	19595,30664	19514,8083	19517,78342	19511,16351	19508,39014
PSO	100	41159,18268	32404,95809	26166,869	34563,34	28483,1693	24047,96276	30706,99491	26603,35596	23125,08598
	1000	27831,00564	23986,12466	21747,726	25286,4	21347,28043	20117,52228	21623,69067	20614,88738	19948,8345
	5000	20707,32489	19861,48152	19588,604	19638,34448	19560,9795	19520,64307	19529,52256	19522,63534	19511,3608
SSA	100	40696,77805	30843,76731	27046,163	41522,23	31848,27576	23198,94943	40628,97117	31973,0032	26916,72938
	1000	63824,98431	31286,77464	21030,705	31282,91	25501,85922	20211,64085	47549,89143	28910,74458	21365,3382
	5000	28169,30759	23867,89307	19525,274	30032,02749	23406,51657	19572,13612	26034,91429	22685,63005	19671,90412
WOA	100	56481,05312	43441,94424	24296,775	51397,87	33718,33034	22550,37552	41141,30307	27294,16099	20671,39233
	1000	31215,27977	25190,74613	20333,709	41090,55	25311,34203	19764,81337	23666,92413	20918,93497	19522,75021
	5000	35441,19553	23217,44358	19595,925	23441,77209	20477,36342	19521,03023	21765,56684	20590,50532	19509,09474

TABLE 4: Statistical comparison of best fitness value of the algorithms.

Algorithm	Best ($f(X)$)	Worst ($f(X)$)	Mean ($f(X)$)	SP
ABC	22830,65571	26339,46279	24666,75976	0
CS	19950,96346	20466,14966	20297,33609	0
GWO	19516,83771	19556,18396	19533,09898	0.4
MVO	19508,39014	19517,78342	19511,16351	1
PSO	19511,3608	19529,52256	19522,63534	0.85
SSA	19671,90412	26034,91429	22685,63005	0
WOA	19509,09474	21765,56684	20590,50532	0.15

TABLE 5: Design variables for the best solutions.

Algorithms	Variables			
	R_0	R	$\mu (10^{-6})$	Q
ABC	6,42110644	5,87339441	6,03021956	3,27847435
CS	5,99669606	5,42618025	5,47140755	2,39732714
GWO	5,95622341	5,38916513	5,36189828	2,27293937
MVO	5,95590533	5,38909289	5,35983982	2,27070608
PSO	5,95635114	5,38953695	5,36015715	2,27129760
SSA	5,95578188	5,38901445	5,48694850	2,36847888
WOA	5,95584921	5,38908429	5,36133333	2,27169540

MVO with the lowest objective function value also has the optimum variable values. Looking at the variable values of MVO and WOA, it can be seen that there is little difference between them, which is compatible with best fitness results. When the results of design variables are considered, it can be said that PSO, MVO, and WOA are very competitive approaches in obtaining minimum power loss and are followed by GWO.

If a comment is to be made, in terms of the algorithms that are applied on the problem for the first time, it can be concluded that MVO and WOA produce high-quality solutions both in objective function values as well as in variable values.

5.2. *Convergence Performances of Optimisers.* In this section, the convergence graphs with population sizes of 100, 400, and 800 in 100, 1000, and 5000 iterations are presented in order to comparatively see the convergence ratios of the algorithms in all population sizes and number of iterations. A general interpretation of the convergence speeds and solution performance of algorithms in different population sizes and number of iterations is thus made.

In Figure 2, convergence rates of algorithms with population of 800 and 5000 iterations where the best result is obtained across all populations and iterations are presented. As can be clearly seen from the figure, although MVO had a slower convergence in comparison with PSO and WOA in the beginning, it reached the best fitness value by surpassing PSO and WOA in the last iterations. WOA reached the best fitness value after the 1764th and PSO after the 4670th iterations. GWO, which exhibited a similar convergence performance to MVO, reached the best fitness value in the last iterations. ABC and CS demonstrated the worst convergence performances, respectively. Considering the

change in the convergence curves, the performance exhibited by WOA and PSO towards solution is assessed to be more stable. There are gradual changes in the convergence curves of MVO, GWO, and SSA. Among these algorithms, MVO and GWO achieved the best fitness values in the last iterations. SSA reached its best fitness value after the 4120th iteration.

The fastest convergence and best quality solutions in a population size of 100 and 100 iterations are seen in WOA and MVO (Figure 3). They are followed by GWO, SSA, and PSO. Although SSA is seen to be converging faster than GWO after the 48th iteration, GWO exhibits a better convergence after the 85th iteration. WOA reaches its best values at iteration #81 and MVO at iteration #100. Although MVO has a slower convergence rate than that of WOA, it is the algorithm that reaches the best solution. CS and ABC are the algorithms with the worst performance in convergence.

As shown in Figure 4, the convergence of WOA, MVO, and SSA in population of 100 and 1000 iterations is better than that of GWO and PSO. CS and ABC exhibits the worst convergence performances. Although WOA reaches its best value (616th iteration) before MVO, MVO reaches the best solution by exhibiting the best convergence performance after the 890th iteration ($f_x = 19605.42531$). MVO and WOA are followed by SSA and PSO.

With a population of 100 and 5000 iterations, WOA, CS, and PSO are seen to converge faster initially and they are followed by SSA, MVO, and GWO (Figure 5). As can be seen in the convergence graph in Figure 5, SSA exhibits better convergence performance in the last 1500 iterations and reached the best solution. Exhibiting rather slow convergence in the beginning, MVO reached the 2nd best solution after demonstrating a faster convergence in the last 500 iterations. ABC has the worst convergence performance. It is therefore observed that there is analogy between the

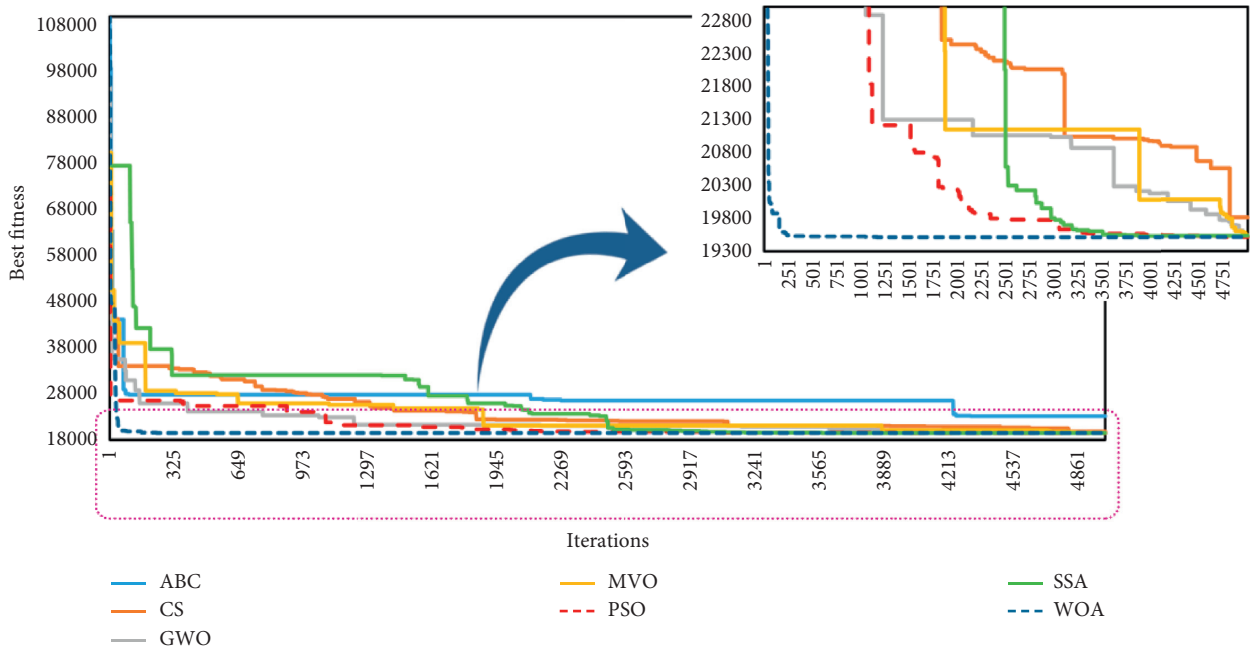


FIGURE 2: Convergence graph for best optimal solutions (in population size of 800 and number of iterations of 5000).

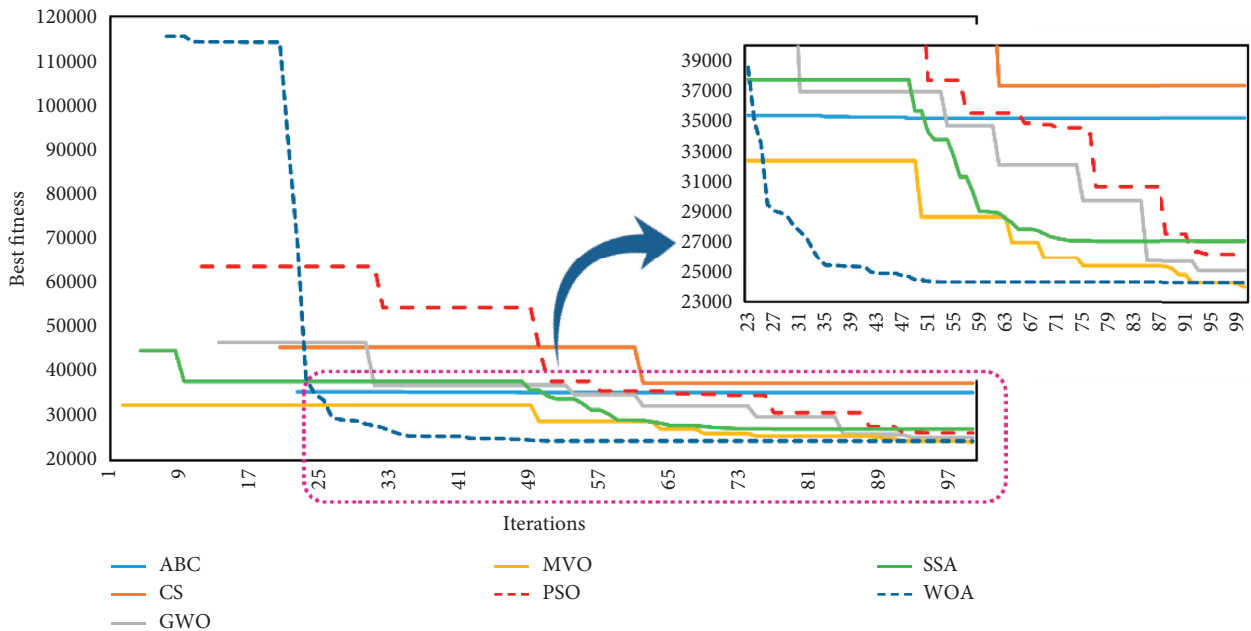


FIGURE 3: Convergence graph for a population size of 100 and number of iterations of 100.

optimum solutions achieved by the algorithms and the convergence speeds.

As seen in Figure 6, having a higher initial convergence speed with population of 400 and 100 iterations, CS is the algorithm with the highest initial convergence speed and it is followed by GWO, WOA, and SSA. While MVO, reaching the best solution ($f(x) = 20299,44124$), with a population of 400 and 100 iterations, has a slower initial convergence, it is seen that it exhibits a faster convergence

after the 30th iteration and reaches the solution. GWO exhibits a faster convergence after iteration #70 and reaches the second-best objective function value of $f(x) = 21571,10275$. PSO and ABC exhibit the worst convergence performances.

In the experiments conducted with a population of 400 and 1000 iterations, as number of iterations of ABC, GWO, PSO, and CS algorithms, excluding PSO, increased, the convergence speed got slower (Figure 7). As can be seen

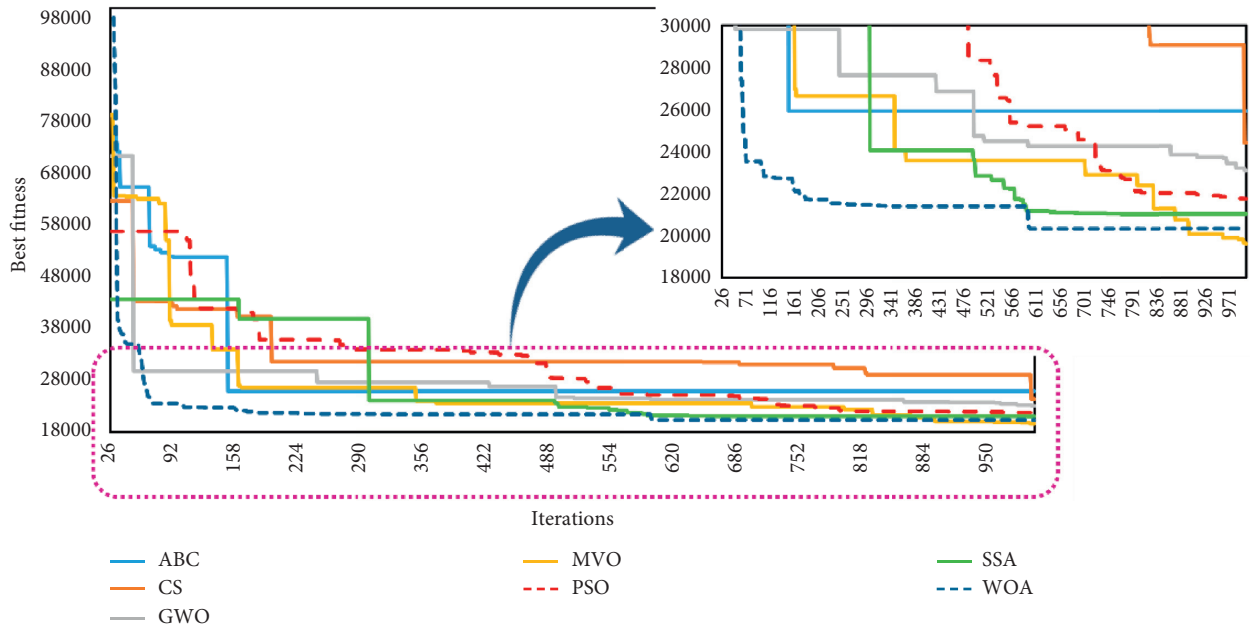


FIGURE 4: Convergence graph for a population size of 100 and number of iterations of 1000.

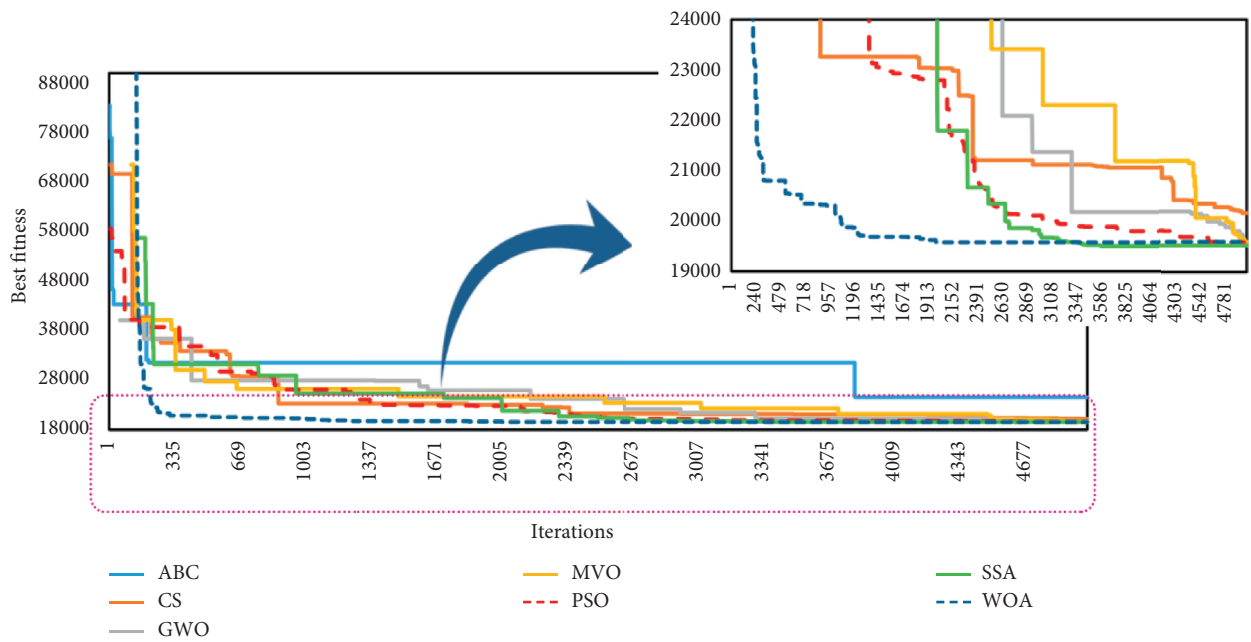


FIGURE 5: Convergence graph for population size of 100 and number of iterations of 5000.

from the convergence graph, having a slower initial convergence speed, WOA exhibits a faster convergence speed after the 80th iteration and SSA after the 570th iteration and they reach the optimal values. MVO, yielding the best solution ($f(x) = 19537,59644$), converges to its optimum value after the iteration #960, and GWO, yielding the second-best solution ($f(x) = 19629.68579$), after iteration #930.

It is seen in Figure 8 that initial convergence rates of WOA, PSO, and GWO are faster than that of ABC, CS, MVO, and SSA in the experiments conducted with a population size of 400

and 5000 iterations. MVO, with a slower initial convergence speed, exhibits a better convergence after the iteration #4550 and generates the best quality solution. PSO exhibits its performance in convergence speed and quality solution, here again, with iteration number of 5000, and generates the second-best objective function value of $f(x) = 19520,64307$. As can be seen from Figure 8, WOA, which converged to the best fitness value in early iterations by showing a rapid convergence in the beginning, was able to produce the third best solution. Having a good initial convergence performance, GWO generates the

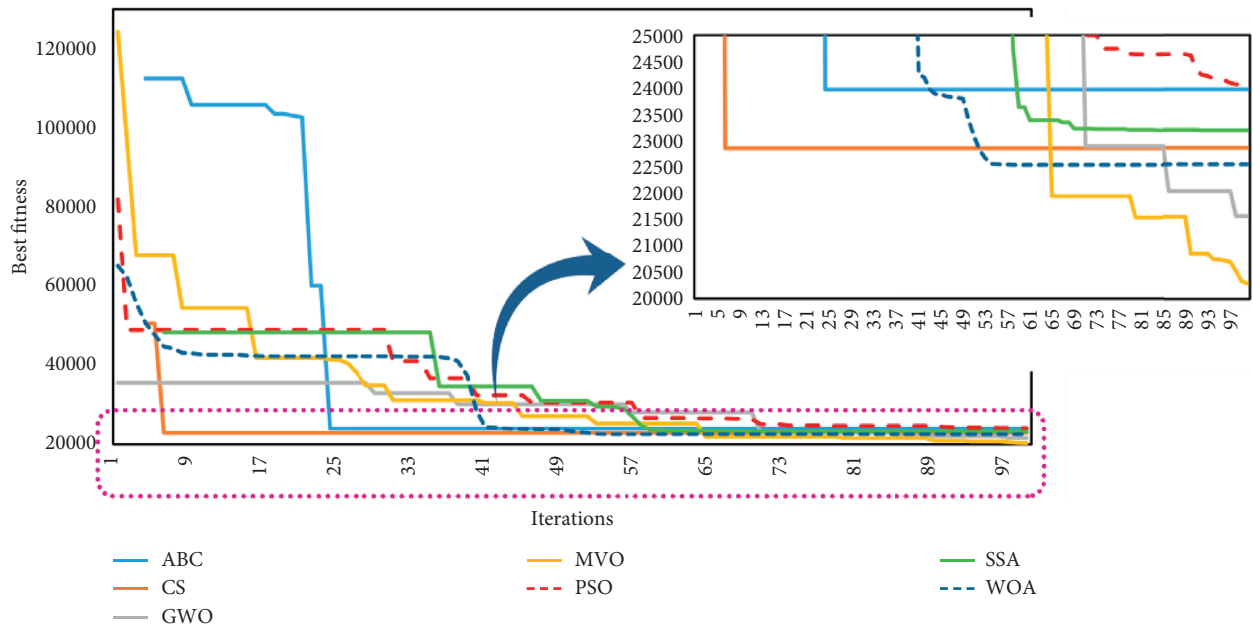


FIGURE 6: Convergence graph for a population size of 400 and number of iterations of 100.

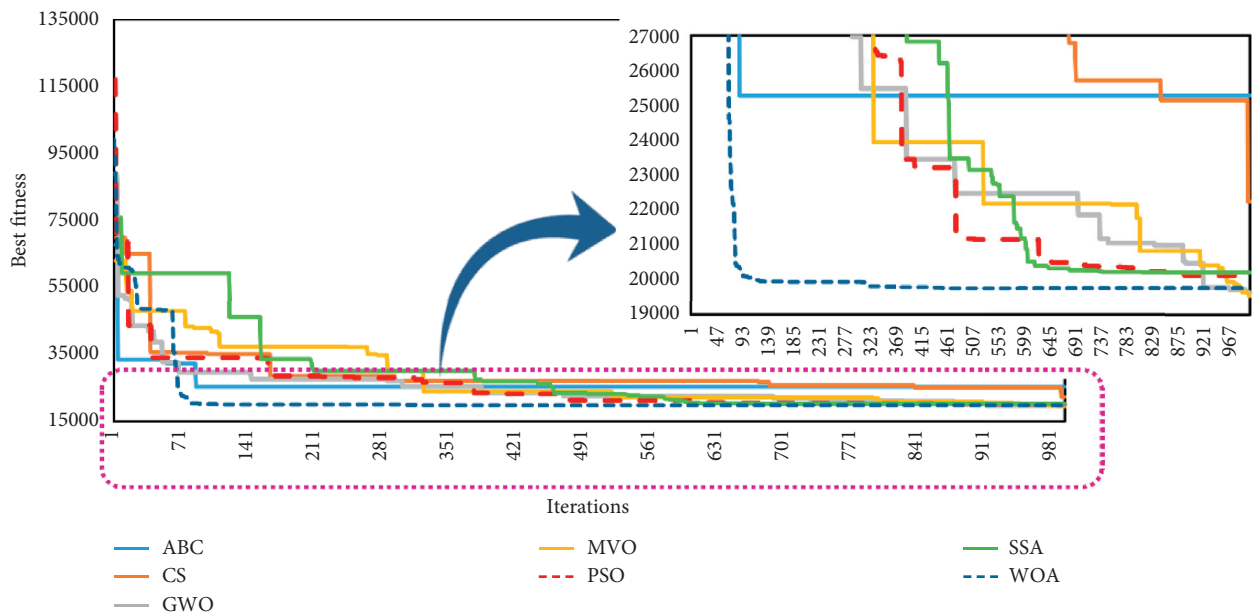


FIGURE 7: Convergence graph for a population size of 400 and number of iterations of 1000.

third-worst solution and ABC generates the worst solutions by exhibiting the worst convergence performance.

Figure 9 shows the convergence graphs of optimisers to their best values with a population size of 800 and 100 iterations. It can be seen here that the convergence rates of WOA, MVO, and GWO are better than that of ABC, PSO, SSA, and CS. While PSO exhibits a faster initial convergence, it exhibits a slower convergence in advancing iterations. In this series, MVO, WOA and GWO present quality solutions both in convergence speed and in best fitness values.

In the experiments carried out with a population size of 800 and 1000 iterations, it can be seen that the convergence speeds of WOA, GWO, and MVO are faster than PSO, SSA,

ABC, and CS (Figure 10). It is seen that MVO, WOA, and GWO generate quality solutions in both convergence speed and best fitness values with this population size and number of iterations. Although the initial convergence of the MVO is slower compared to that of GWO and WOA, it reaches the best objective function value of $f(x) = 19517,10515$.

5.3. Results of Execution Time Analysis. Execution time analysis is an important parameter used for evaluating the performance of optimisers. In the study, the speeds were compared with each other based on the time optimisers took for the solution of the problem where the best fitness value was

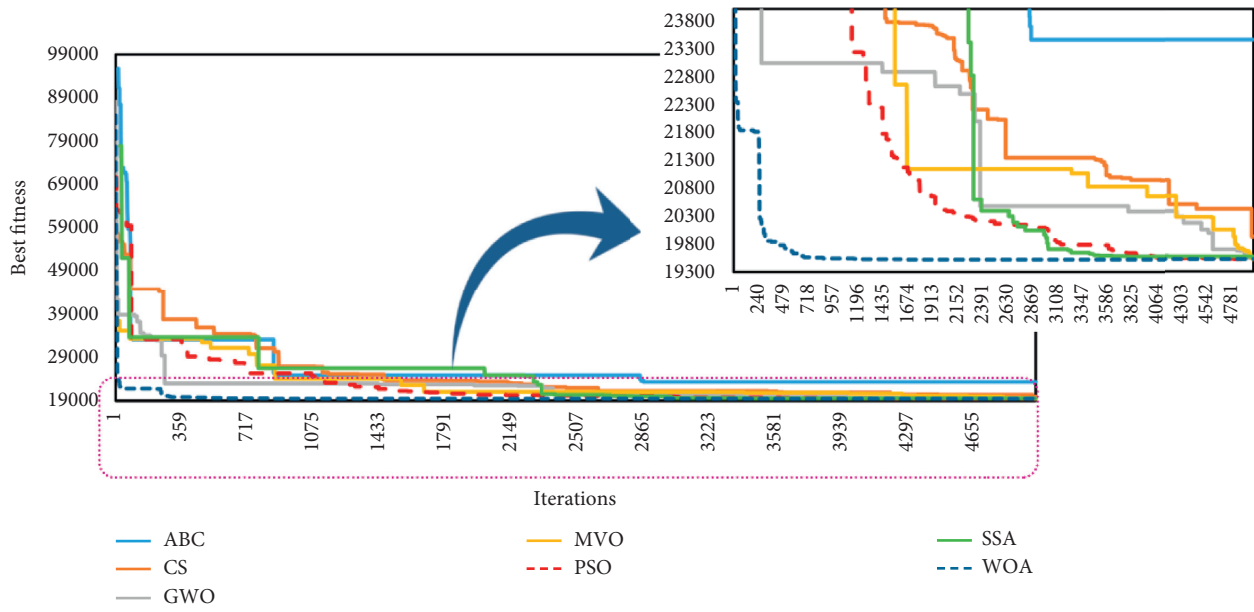


FIGURE 8: Convergence graph for a population size of 400 and number of iterations of 5000.

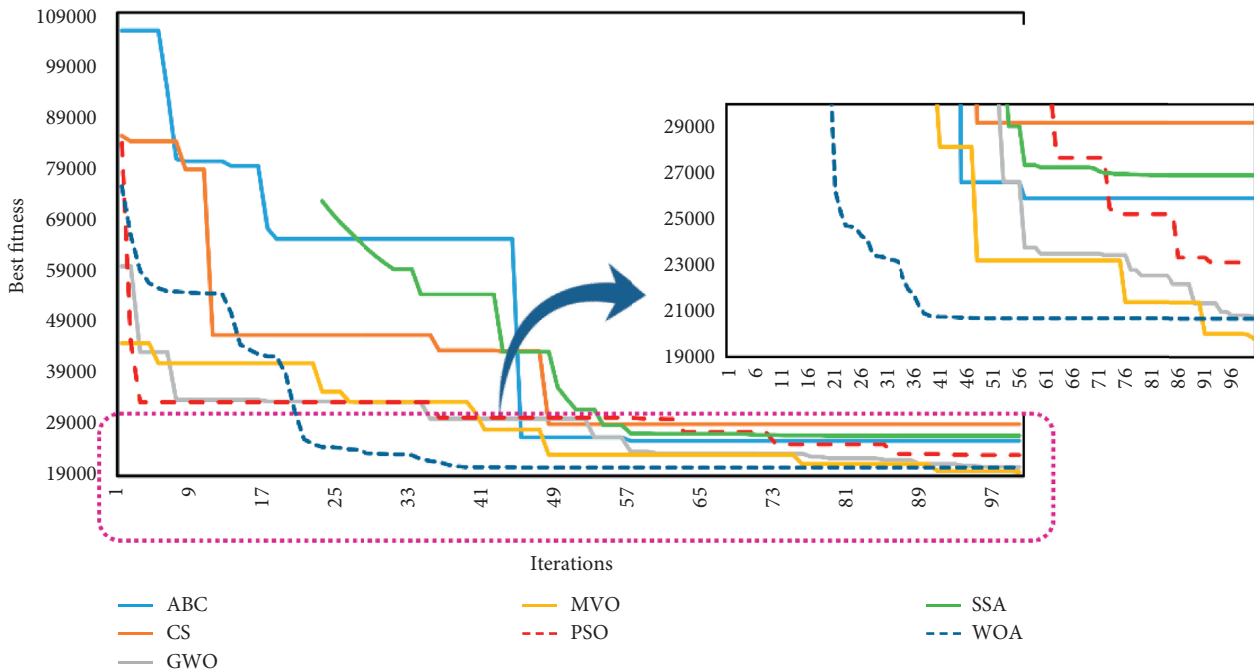


FIGURE 9: Convergence graph for a population size of 800 and number of iterations of 100.

reached with a population size of 800 and iteration number of 5000 and 20 studies (Figure 11). With this comparison, it is possible to construct the time complexity of algorithms. As seen in the graph shown in Figure 11, the longest calculation time for solving the problem is seen in ABC. It is followed by SSA and WOA. Having the best fitness value, PSO achieved this success by solving the problem at an average of 19.51 seconds. PSO, GWO, and MVO, which are most successful ones in solving the problem, are also more successful than other algorithms in terms of average calculation time. The

average running times of algorithms are close to each other. Although algorithms differ in terms of running time, it can be said that, with a general assessment, all algorithms solve the problem in a reasonable running time.

6. Conclusion

In this study, the hydrostatic thrust bearing design problem defined by Siddall [12] is discussed through 7 different swarm intelligence approaches. The purpose of the problem

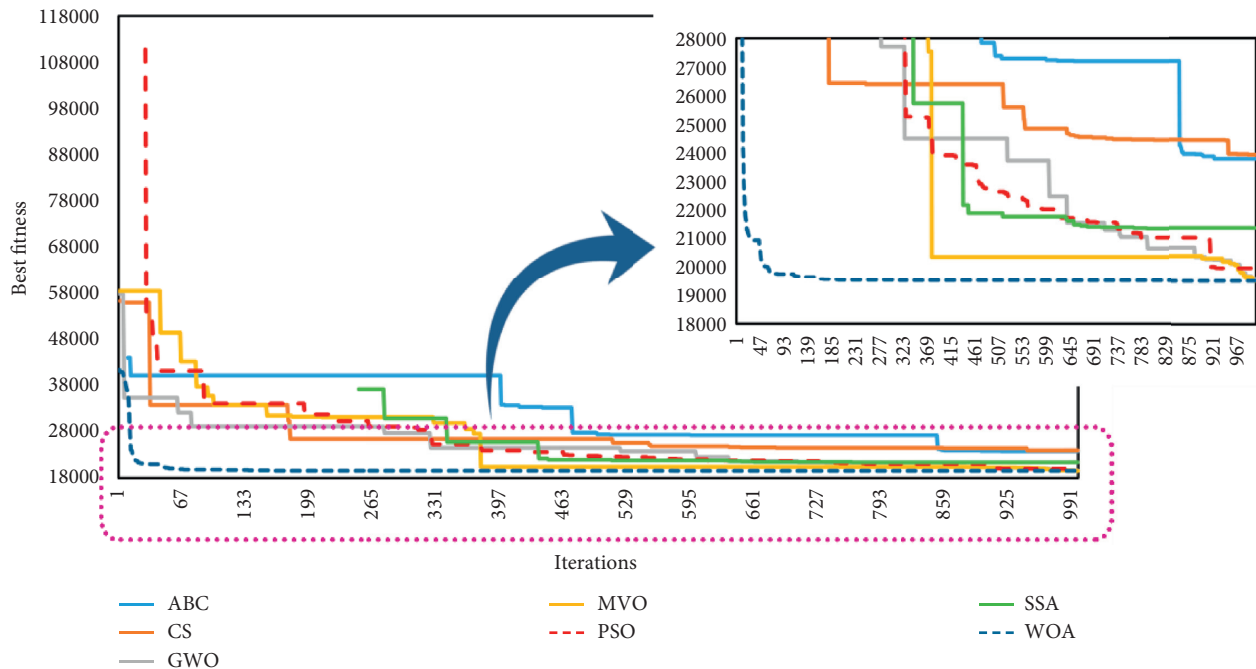


FIGURE 10: Convergence graph for a population size of 800 and number of iterations of 1000.

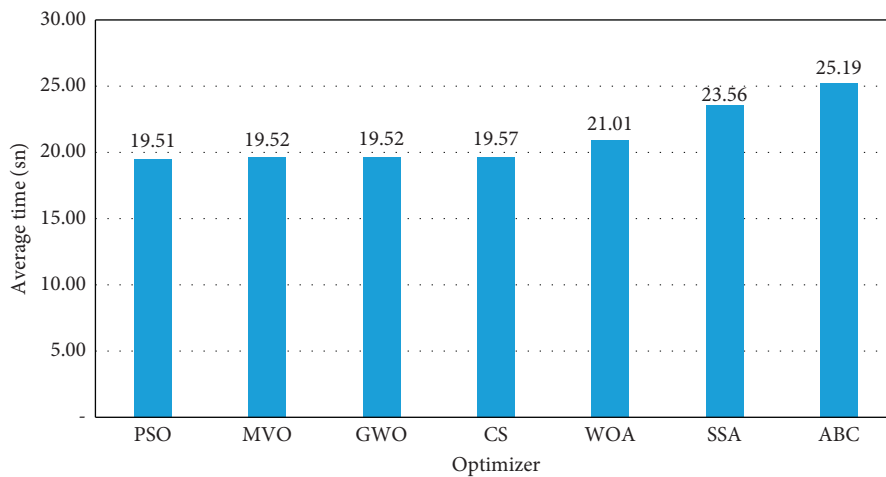


FIGURE 11: Average execution time of all optimisers for a population size of 800 and number of iterations of 5000.

is to minimize the power loss that occurs during operation of hydrostatic thrust bearings. In addition to the successes of 7 different optimisers in achieving an optimum solution, their performances in different population sizes and numbers of iterations are also discussed as the research subject. For this reason, optimisers were run with population sizes of 100, 400, and 800 in 100, 1000, and 5000 iterations. According to the results obtained, the lowest objective function value in the study was obtained by MVO with a population size of 800 and 5000 iterations. Following MVO, the best solution was reached through WOA, PSO, and GWO. The performances of WOA, MVO, CS, and SSA, which were applied for the first time to solve the minimum power loss problem, are particularly successful in higher iteration numbers.

The successful performance of MVO and WOA in the populations and iterations where the best fitness value is reached is remarkable. The results obtained with the other two algorithms (SSA and CS) applied for the first time to the solution of the problem are also very close to the best fitness value. This reveals the competitive aspects of algorithms that have not been applied to the problem earlier. The results of objective functions obtained with PSO [10] and GWO [2], which were earlier applied to the solution of the problem, are better than that of the studies in the literature.

When the performances of algorithms in different population sizes and iterations are analysed, it is seen that, as the number of iterations increased, the algorithms reached the solutions more easily. The increase in the number of iterations increased the precision of the solution.

Contribution of the increase in population size into quality solutions was not as much as that of the number of iterations. The outcomes of ABC and CS with a population size of 800 support this remark. Another remarkable outcome of the study is the speed of the algorithms. The best solution was reached in 19.51 seconds and the worst solution in 25.19 seconds with a population size of 800 and 5000 iteration numbers where the best fitness value was reached.

Quality solutions produced by swarm optimisers used for the first time in the study demonstrate that these methods can be used in delicate engineering design problems, such as the problem discussed in this article. However, the obligation for going up to higher numbers of iterations to reach the ideal solution is a significant issue here. In the upcoming studies, the focus will be on solving benchmark problems with hybrid models, such as the problem discussed in this study.

Data Availability

The data used to support the findings of this study are included within the article, and obtained other datasets can be requested from the corresponding author if needed.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

- [1] A. Kentli and M. Sahbaz, "Optimisation of hydrostatic thrust bearing using sequential quadratic programming," *Oxidation Communications*, vol. 37, no. 4, pp. 1144–1152, 2014.
- [2] İ. Şahin, M. Dörterler, and H. Gökçe, "Optimization of hydrostatic thrust bearing using enhanced grey wolf optimizer," *Mechanics*, vol. 25, no. 6, pp. 480–486, 2019.
- [3] W. B. Rowe, *Hydrostatic, Aerostatic and Hybrid Bearing Design*, Elsevier, Amsterdam, Netherlands, 2012.
- [4] O. J. Bakker and R. A. J. Van Ostayen, "Recess depth optimization for rotating, annular, and circular recess hydrostatic thrust bearings," *Journal of Tribology*, vol. 132, no. 1, Article ID 011103, 2010.
- [5] Z. Liu, Y. Wang, L. Cai, Y. Zhao, Q. Cheng, and X. Dong, "A review of hydrostatic bearing system: researches and applications," *Advances in Mechanical Engineering*, vol. 9, no. 10, Article ID 168781401773053, 2017.
- [6] W. B. Rowe, *Hydrostatic and hybrid bearing design*, Butterworth & Co. (Publishers) Ltd., Oxford, UK, 1983.
- [7] E. Solmaz, F. C. Babalik, and F. Öztürk, "Multicriteria optimization approach for hydrostatic bearing design," *Industrial Lubrication and Tribology*, vol. 54, no. 1, pp. 20–25, 2002.
- [8] C. A. C. Coello, "Treating constraints as objectives for single-objective evolutionary optimization," *Engineering Optimization*, vol. 32, no. 3, pp. 275–308, 2000.
- [9] K. Deb and M. Goyal, "Optimizing engineering designs using a combined genetic search," in *Proceedings of the Sixth International Conference on Genetic Algorithms*, pp. 521–528, San Francisco, CA, USA, July 1997.
- [10] S. He, E. Prempan, and Q. H. Wu, "An improved particle swarm optimizer for mechanical design optimization problems," *Engineering Optimization*, vol. 36, no. 5, pp. 585–605, 2004.
- [11] R. V. Rao, V. J. Savsani, and D. P. Vakharia, "Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems," *Computer-Aided Design*, vol. 43, no. 3, pp. 303–315, 2011.
- [12] J. N. Siddall, *Optimal Engineering Design*, Marcel Dekker, New York, NY, USA, 1982.
- [13] S. Talatahari and M. Azizi, "Optimization of constrained mathematical and engineering design problems using chaos game optimization," *Computers & Industrial Engineering*, vol. 145, Article ID 106560, 2020.
- [14] S. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Faris, and S. M. Mirjalili, "Salp Swarm algorithm: a bio-inspired optimizer for engineering design problems," *Advances in Engineering Software*, vol. 114, pp. 163–191, 2017.
- [15] F. W. Hoffer: "Automatic fluid pressure balancing system," U.S. Patent and Trademark Office, Washington, DC, USA, U.S. Patent No. 2,449,297, 1948.
- [16] F. G. Arneson: "Externally pressurized bearing structure," U.S. Patent and Trademark Office, Washington, DC, US, U.S. Patent No. 3,472,565, 1969.
- [17] H. E. Arneson: "Hydrostatic bearing structure," U.S. Patent and Trademark Office, Washington, DC, USA, U.S. Patent No. 3,305,282, 1967.
- [18] N. R. Kane and A. H. Slocum: "Modular hydrostatic bearing with carriage form-fit to PR," U.S. Patent and Trademark Office, Washington, DC, USA, U.S. Patent No. 5,971,614, 1999.
- [19] A. H. Slocum, U.S. Patent and Trademark Office, Washington, DC, USA, U.S. Patent No. 5,104,237, 1992.
- [20] A. H. Slocum, P. A. Scagnetti, N. R. Kane, and C. Brunner, "Design of self-compensated, water-hydrostatic bearings," *Precision Engineering*, vol. 17, no. 3, pp. 173–185, 1995.
- [21] N. Tully, "Static and dynamic performance of an infinite stiffness hydrostatic thrust bearing," *Journal of Lubrication Technology*, vol. 99, no. 1, pp. 106–112, 1977.
- [22] T. A. Osman, M. Dorid, Z. S. Safar, and M. O. A. Mokhtar, "Experimental assessment of hydrostatic thrust bearing performance," *Tribology International*, vol. 29, no. 3, pp. 233–239, 1996.
- [23] H. Shen, J. Zhou, Q. Peng, H. Li, and J. Li, "Multi-objective interplanetary trajectory optimization combining low-thrust propulsion and gravity-assist maneuvers," *Science China Technological Sciences*, vol. 55, no. 3, pp. 841–847, 2012.
- [24] M. Fesanghary and M. M. Khonsari, "Topological and shape optimization of thrust bearings for enhanced load-carrying capacity," *Tribology International*, vol. 53, pp. 12–21, 2012.
- [25] W. Wang, Y. He, J. Zhao, Y. Li, and J. Luo, "Numerical optimization of the groove texture bottom profile for thrust bearings," *Tribology International*, vol. 109, pp. 69–77, 2017.
- [26] C. Weißbacher, C. Schellnegger, A. John, T. Buchgraber, and W. Pscheidt, "Optimization of journal bearing profiles with respect to stiffness and load-carrying capacity," *Journal of Tribology*, vol. 136, no. 3, Article ID 031709, 2014.
- [27] S. K. Yadav and S. C. Sharma, "Performance of hydrostatic tilted thrust pad bearings of various recess shapes operating with non-Newtonian lubricant," *Finite Elements in Analysis and Design*, vol. 87, pp. 43–55, 2014.
- [28] A. E. Hegazy, M. A. Makhlof, and G. S. El-Tawel, "Improved salp swarm algorithm for feature selection," *Journal of King Saud University - Computer and Information Sciences*, vol. 32, no. 3, pp. 335–344, 2020.
- [29] X. S. Yang and S. Deb, "Engineering optimisation by cuckoo search," *International Journal of Mathematical Modelling and Numerical Optimisation*, vol. 1, no. 4, pp. 330–343, 2010.

- [30] M. Mareli and B. Twala, "An adaptive Cuckoo search algorithm for optimisation," *Applied Computing and Informatics*, vol. 14, no. 2, pp. 107–115, 2018.
- [31] S. Mirjalili, S. M. Mirjalili, and A. Hatamlou, "Multi-verse optimizer: a nature-inspired algorithm for global optimization," *Neural Computing and Applications*, vol. 27, no. 2, pp. 495–513, 2016.
- [32] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," vol. 200 Technical report-TR06, pp. 1–10, Erciyes University, Engineering Faculty, Computer Engineering Department, Kayseri, Turkey, 2005.
- [33] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016.
- [34] Ü. Atila, M. Dörterler, R. Durgut, and İ. Şahin, "A comprehensive investigation into the performance of optimization methods in spur gear design," *Engineering Optimization*, vol. 52, no. 6, pp. 1052–1067, 2020.
- [35] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, vol. IV, pp. 1942–1948, Perth, Australia, November 1995.
- [36] H. Shi, S. Liu, H. Wu et al., "Oscillatory particle swarm optimizer," *Applied Soft Computing*, vol. 73, pp. 316–327, 2018.
- [37] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.
- [38] M. Dörterler, İ. Şahin, and H. Gökçe, "A grey wolf optimizer approach for optimal weight design problem of the spur gear," *Engineering Optimization*, vol. 51, no. 6, pp. 1013–1027, 2019.

Research Article

Model and Hybrid Algorithm of Collaborative Distribution System with Multiple Drones and a Truck

Min Lin, Jun-Yan Lyu, Jia-Jing Gao, and Ling-Yu Li 

School of Management, Shanghai University, Shanghai 200444, China

Correspondence should be addressed to Ling-Yu Li; 2584687069@qq.com

Received 7 April 2020; Revised 12 May 2020; Accepted 19 August 2020; Published 1 September 2020

Academic Editor: Tingsong Wang

Copyright © 2020 Min Lin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper studies a coordinated system for multidrone single-truck distribution, where a truck delivers goods to a group of customers along a closed ground path with the help of a number of drones. For each delivery, the truck departs from the distribution centre with drones and all goods needed and returns back to the centre after fulfilling the delivery tasks. That is, the truck assigns these delivery tasks to several of its drones, each of which is responsible for sending goods to a different subgroup of customers in the empty air space. This study provides a new mixed-integer programming model of the routing problem with this distribution system based on urban road network. Meanwhile, a hybrid genetic algorithm and a hybrid particle swarm algorithm are designed. Experimental results show that the performance of the hybrid algorithms is better than that of the corresponding basic algorithms.

1. Introduction

Drones or UAVs (Unmanned Aerial Vehicle) could be used to transport goods such as packages, food, and medicine. Compared with traditional delivery with trucks, delivery with drones has several advantages. Firstly, a drone is able to be unmanned without well-paid drivers so that human and operational costs can be reduced. Secondly, a drone has the ability of flying at a high speed in the empty airspace to avoid ground road congestion and then improve distribution efficiency effectually. Thirdly, drones are especially suitable for the areas or situations in which land transport is not much convenient, such as remote mountains and islands and transport for emergency supplies.

However, there are also some obvious drawbacks for drone delivery. The low load capacity of a drone limits the weight or size of the goods it carries. In addition, since drones are powered by batteries, they have a very short flight radius compared to ordinary fuel trucks. Both of these disadvantages require the drone to travel frequently between central warehouse and customer demand point, which means that, after each delivery, the drone needs to return back to the warehouse to take the next goods, as shown in

Figure 1. However, although a normal fuel truck can travel longer distances and carry more goods, it may travel slowly and delivery efficiency may be affected easily due to the large weight itself and road traffic jams. In addition, truck drivers have to operate during the whole driving route so that high salary will be paid to drivers, which means high human cost. In contrast, drones are able to save this kind of cost by delivering automatically without being operated during the whole time. Besides, electric cost of drones is much cheaper than fuel cost of trucks, so truck expense is more in every mile, especially when customer points are very scattered.

A solution to the limited flight distance and load capacity of the drone is to allow the drone to collaborate with other types of delivery tools to complete the delivery tasks. Figure 2 shows a collaborative distribution system that uses a drone and a truck. The truck loads the goods for a group of customers at the distribution centre in advance, and it also needs to carry a drone. After leaving from the distribution centre, the truck is ought to move along a closed path on the ground. When the truck travels to a location near a customer demand point, it sends the drone carrying goods in the direction of this customer, while the truck continues to move on its intended path. The drone carries the goods along a

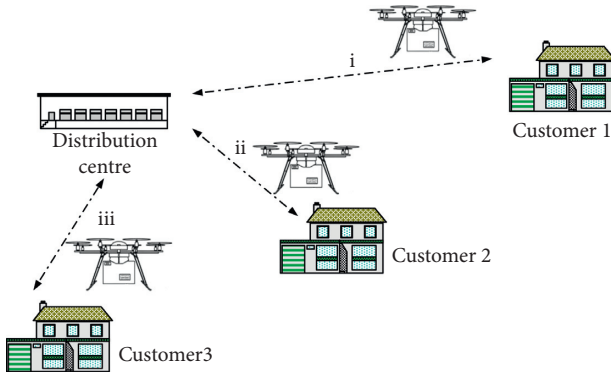


FIGURE 1: A distribution system that only uses a drone. (i) The drone delivers to Customer 1 from the distribution centre and returns to it. (ii) The drone delivers to Customer 2 from the distribution centre and then returns back. (iii) The drone delivers to another customer from the distribution centre and then returns back.

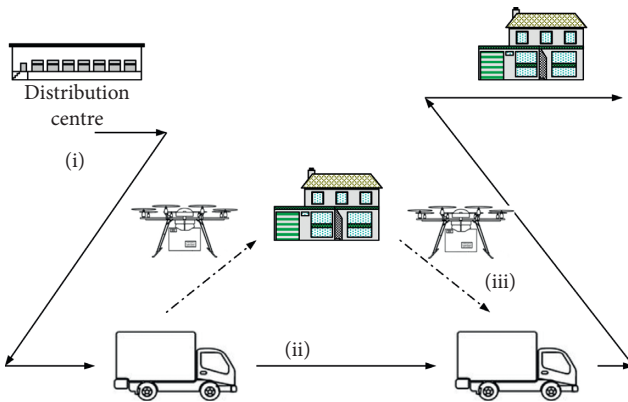


FIGURE 2: A collaborative distribution system that uses a drone and a truck. (i) The truck is loading goods from the distribution centre. (ii) When the truck approaches a customer, it sends a drone to deliver goods to the customer, and the truck continues to move. (iii) The drone returns to the truck and prepares the next customer's delivery.

straight line to the customer point, then returns back to the truck in another straight line after delivering, gets ready to pick-up new goods, and delivers them to the next customer. For every customer demand point, the drone performs the same actions as mentioned above so that the goods are able to be delivered to all the customers. At last, the truck returns to its original distribution centre.

At present, researchers have also begun to study the optimization of the collaborative distribution system where trucks and drones collaborate to complete delivery tasks, rather than simply relying on single delivery tools. Murray and Raj [1] studied a new extension of the “flying sidekick traveling salesman problem,” where a truck works with multiple various drones to deliver packages [1]. They described the problem as a mixed-integer linear programming and proposed a method by using three subproblem sequences. Zhen et al. [2] modelled and solved a nonlinear MIP model for different instance sizes with two efficient

algorithms, which are SOCP-based algorithm and dynamic linearization algorithm to handle nonconvex constraints and the model under different scales of route networks, respectively [2]. Bai et al. [3] studied the efficient routing problem for precedence-constrained package delivery for heterogeneous vehicles, where a truck and a microdrone are utilized to deliver packages to dispersed customers [3]. Trucks are restricted to the street network, while microdrones are available for the last-mile delivery with the restriction of their capacity and flight range. The goal of the problem is to minimize the delivery time in the priority constraint. The authors combined topological sorting technology and proposed several task distribution algorithms. Poikonen and Golden considered the k -multivisit drone routing problem, where every drone can take off from a truck and deliver one or more packages to customers [4]. Each drone can return to a truck to replace a battery or charge, take a new package, and then fly again to another customer demand point. de Freitas and Penna studied a flying sidekick traveling salesman problem, where the drone takes off from the truck, delivers the package to the customer, and returns to the truck in the third place [5]. When the drone flies, the truck delivers packages to other customers as long as the drone has enough batteries to hover in the air. Zhen imposed both a stochastic programming formulation that can cope with arbitrary probability distribution of operation time deviation and a robust formulation that is applicable to situation in which limited information about probability distributions is available [6]. Ha et al. studied the Traveling Salesman Problem with Drone (TSP-D) and proposed a hybrid genetic search with dynamic population management and adaptive diversity control based on a split algorithm to solve the completion time for the truck and drone [7]. Results show that the proposed algorithm outperforms existing methods in terms of solution quality and improves best known solutions found in the literature. Zhen et al. developed a column generation solution approach on a set partitioning-based reformulation of the original model [8]. Results show that this simple but practical solution approach can optimally solve the daily berth planning problem instances with reasonable and acceptable for the real-world applications. Liu presented a mixed-integer programming (MIP) model for on-demand meal delivery using drones and proposed an optimization-driven, progressive algorithm for online fleet dispatch operations [9]. Schermer et al. proposed an extension of the VRPD that is called the Vehicle Routing Problem with Drones and En Route Operations (VRPDERO), where drones may not only be launched and retrieved at vertices but also on some discrete points are located on each arc [10]. They also studied the Vehicle Routing Problem with Drones (VRPD), where given a fleet of trucks, each truck carries a given number of drones, and the objective consists in designing feasible routes and drone operations such that all customers are served and minimal make span is achieved [11]. Jeong et al. extended the previous vehicle routing models to the hybrid delivery systems by considering two important practical issues: the effect of parcel weight on drone energy consumption and restricted flying areas [12].

Zhen considered a combination of probabilistic and physics-based models for truck interruptions [13]. It enables to exactly evaluate link travel time, which then acts as the basis for proposing a mixed-integer programming model that minimizes the total expected travel time of moving containers. Wang et al. studied the traveling salesman problem with drone (TSP-D), which aims to find the coordinated routes of a drone and a truck to serve a list of customers. In practice, managers sometimes intend to attain a compromise between operational cost and completion time. Therefore, this article addresses a biobjective TSP-D considering both objectives [14]. Sacramento et al. published a paper where a mathematical model is formulated, defining a problem similar to the Flying Sidekick Traveling Salesman Problem, but for the capacitated multiple-truck case with time limit constraints and minimizing cost as objective function [15]. Karak and Abdelghany presented a mathematical formulation and efficient solution methodology for the hybrid vehicle-drone routing problem (HVDRP) for pick-up and delivery services. The problem is formulated as a mixed-integer program, which minimizes the vehicle and drone routing cost to serve all customers [16]. Javadi et al. extended the traveling Repairman Problem (TRP) by assuming a single truck which can stop at customer locations and launch drones multiple times for each stop location to serve customers. They also developed an efficient hybrid Tabu Search-Simulated Annealing algorithm to solve the problem [17]. Moshref-Javadi et al. presented a mathematical formulation and a heuristic solution approach for the optimal planning of delivery routes. They formulated the problem as a Mixed-Integer Linear Programming model and proposed an efficient Truck and Drone Routing Algorithm (TDRA) [18]. Gonzalez-R et al. proposed an iterated greedy heuristic based on the iterative process of destruction and reconstruction of solutions. This process is orchestrated by a global optimization scheme using a simulated annealing algorithm. The obtained results are quite promising even for large-size scenarios [19]. Chang and Lee focused on finding an effective delivery route for trucks carrying drones and proposed a new approach on a nonlinear programming model to find shift-weights that move the centres of clusters to make wider drone-delivery areas along shorter truck route after initial K-means clustering and TSP (Traveling Salesman Problem) modelling [20]. Carlsson and Song proved that the improvement in efficiency is related to the square root of the ratio of the speeds of the truck and the UAV by combining a theoretical analysis in the Euclidean plane with real-time numerical simulations on a road network [21]. Xia et al. studied how drones can be scheduled to monitor the sailing vessels in ECAs. They also modelled the dynamics of each sailing vessel using a real-time location function which allows to approximately represent the problem on a time-expanded network. Furthermore, they developed a Lagrangian relaxation-based method to obtain near-optimal solutions [22].

The contributions to this article are as follows. (1) This paper delved into the issue of collaborative distribution route planning for drones and trucks based on urban road networks. Although there is a lot of research on co-

distribution in recent years, research on co-distribution path planning based on urban road networks is not enough, and our paper filled this gap. (2) This problem involves many decision variables, including determining the routes of trucks and multiple drones and distributing tasks between multiple drones. A number of constraints are also involved in the issue, such as the closeness of the route of the truck, the cargo capacity of the truck, the cruising ability of a drone, and the sequential relationship between all launch sites and landing sites for one drone. (3) For megacities such as Shanghai, London, and Paris, the urban road network is particularly complex and large with a particularly large number of customers. It is apparent that, with the increasing size of the problem and the growing number of decision variables, performance of the algorithm gets worse. Thus, we designed a hybrid algorithm that considers the efficiency and quality of the solution and achieves better computational results at a faster rate.

2. Problem

In this section, we describe a coordinated system for multidrone single-truck distribution and then illustrate the routing problem of the coordinated system on urban road network.

2.1. Multidrone Single-Truck Distribution System. A coordinated system for multidrone single-truck distribution means that a truck works with multiple drones to deliver to a set of determined customers. That is, the truck assigns these delivery tasks to several of its drones, each of which is responsible for sending goods to a different subgroup of customers in the empty air space. In particular, a truck loads goods and a number of drones from the distribution centre at first. After leaving the distribution centre, the truck travels along a closed ground path. When approaching a customer demand point, the truck releases one of the remaining drones on it and continues to move forward on its intended path, while the drone carries goods to the customer point along a straight line and returns to the truck along another straight line after delivery. When the truck approaches the next customer demand point, the truck releases another drone from the remaining drones, the truck goes on to move forward, and the drone executes the same operation as mentioned above before returning to the truck. At this point, some of previously released drones may still be on delivery to customers and have not yet returned to the truck. If there are no extra drones on the truck, the truck has to wait for a drone to come back and then launch it to another customer point.

2.2. Collaborative Distribution System on Urban Road Network. A city road network refers to a network of all roads in the city. It can be expressed as $G = (V, E)$, where V is a collection of all road intersections and E is a collection of the whole roads. In V , the number of intersections is $0, 1, 2, \dots, n - 1$, and n , where the intersection of number 0

represents the distribution centre. The coordinate of Intersection u is

$$\mu_u = (\mu_u^1, \mu_u^2) \in \mathfrak{R}, \quad u = 0, 1, 2, 3, \dots, n. \quad (1)$$

A directed line from Intersection u to Intersection v is defined as follows:

$$(u, v), \quad u, v = 0, 1, 2, 3, \dots, n. \quad (2)$$

Figure 3 is an example of an urban road network. This network is composed of 32 intersections, where the distribution centre is Intersection 0, and the other intersections are numbered as 1, 2, ..., 30, and 31. The distribution centre 0 requires to deliver goods to six customers, $a, b, c, d, e,$ and f , respectively. A truck will start from the distribution centre to deliver those six customers in turn and then return to the distribution centre. There are many options for this closed ground route driven by the truck. Here are two options:

$$\begin{aligned} &0 \rightarrow 11 \rightarrow 10 \rightarrow 9 \rightarrow 16 \rightarrow 17 \rightarrow 22 \rightarrow 23 \rightarrow \\ &24 \rightarrow 18 \rightarrow 0, \\ &0 \rightarrow 18 \rightarrow 13 \rightarrow 12 \rightarrow 11 \rightarrow 10 \rightarrow 9 \rightarrow 16 \rightarrow \\ &17 \rightarrow 22 \rightarrow 23 \rightarrow 0. \end{aligned} \quad (3)$$

Figure 4 shows the route of the truck and the flight route of multiple drones, where the black thick arrow represents the route of the truck, and the red and green light arrows represent the flight routes of Drone 1 and Drone 2, respectively. In each delivery, Drone 1 moves in a red line to a customer and then returns back to the truck in another red line. The process is the same as Drone 2 while the colour of line is green. It is obvious to see that this collaborative distribution system has improved the delivery efficiency to a great extent.

The following is the procedure of collaborative distribution in Figure 4. The truck starts from the distribution centre and moves along a closed ground route indicated with the black thick arrow. First of all, it travels along the line $0 \rightarrow 11 \rightarrow 10 \rightarrow 9$. When the truck is gradually close to intersection 9, distance between the truck and Customer a decreases. The truck launches Drone 1 along with the red arrow to Customer a . After sending goods to Customer a , Drone 1 goes back along another red arrow to the truck which is moving forward and preparing the next delivery at the same time. The truck continues to drive along line $9 \rightarrow 16 \rightarrow 17$. When traveling to Intersection 16, the truck is getting closer to Customer c , and it launches Drone 1 to deliver to Customer c once again. Similarly, the drone moves along the red arrow, performs the same set of actions as mentioned above, and finally returns to the truck which is moving forward after Customer c 's delivery is completed. In the next stage, the truck continues to drive along the line $17 \rightarrow 22 \rightarrow 23$, and when approaching Intersection 22, the truck sends Drone 1 to Customer d along the red arrow. Noted that Drone 1 does not return to the truck when the truck is close to Intersection 23, so Drone 2 is sent to Customer f along the green arrow. Then, the truck continues

to drive along the line $23 \rightarrow 24 \rightarrow 18 \rightarrow 0$, and when it is near to Intersection 24, it sends Drone 2 along the green arrow to Customer e . Since Drone 2 has not returned to the truck when it is approaching to Intersection 18, the truck needs to send Drone 1 instead along the red arrow to Customer b . At this moment, all customers have been delivered and the truck could return to the distribution centre.

2.3. Notational Conventions. Before discussing the mathematical model, it is necessary to illustrate some mathematical symbols involved. Here are some of the main mathematical symbols.

$\| * \|$: we use Euclidean distance to denote travel distance according to Carlsson [21] and Agatz [23], for any two points $A(x_1, y_1)$ and $B(x_2, y_2)$:

$$\|A - B\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}. \quad (4)$$

\mathfrak{R} : the Euclidean plane.

h : the number of customers to visit in the Euclidean plane.

I : the collection of the customers that need to be served, $I = \{1, 2, \dots, h\}$.

i : customer number, $i \in I$.

p_i : the coordinates of the position of Customer i , $p_i = (p_i^1, p_i^2) \in \mathfrak{R}$.

g_i : the amount of demand for Customer i .

m : the total number of drones on a truck.

K : the collection of all the drones, $K = \{1, 2, \dots, m\}$.

k : drone number, $k \in K$.

q : load weight of the truck.

\wp : closed transport route of the truck.

G : urban road network $G = (V, E)$.

n : the total number of the road intersections.

V : a collection of all road intersections, $V = \{0, 1, 2, 3, \dots, n\}$.

(u, v) : a directed line from Intersection u to Intersection v , $u, v \in V$, called a directed road.

$\varepsilon_{u,v}$: a binary indicating whether a directed line segment (u, v) is present in the urban road network $G = (V, E)$. Note that there is no directed line segment from any intersection u to itself, that is to say, $(u, u) \notin E$:

$$\varepsilon_{u,v} = \begin{cases} 1, & \text{if } (u, v) \in E, u \neq v, \\ 0, & \text{if } (u, v) \notin E, u \neq v, \\ 0, & \text{if } u = v. \end{cases} \quad (5)$$

E : a collection of all the directed roads, $E = \{(u, v) \mid \varepsilon_{u,v} = 1, \forall u, v \in V\}$.

μ_u : The coordinates of the road intersections, $\mu_u = (\mu_u^1, \mu_u^2) \in \mathfrak{R}$, $u \in V$.

ϕ_0 : speed of the truck.

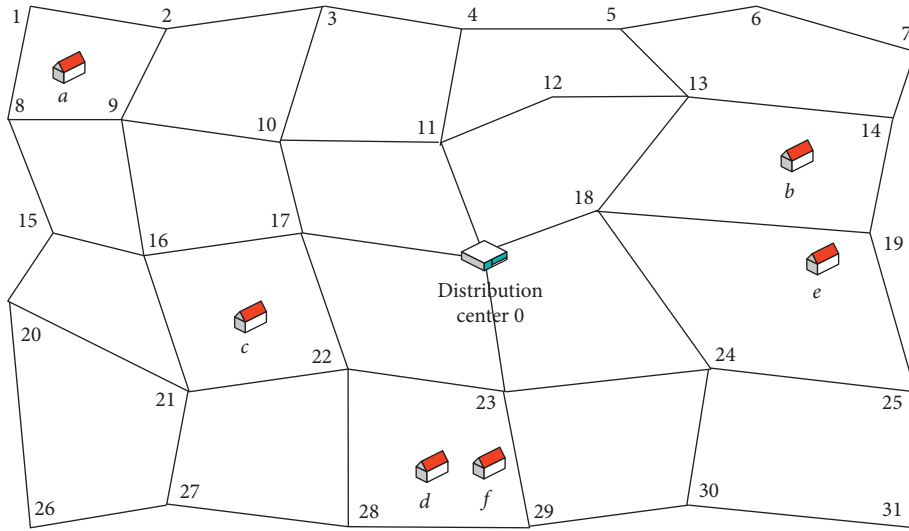


FIGURE 3: Example of an urban road network.

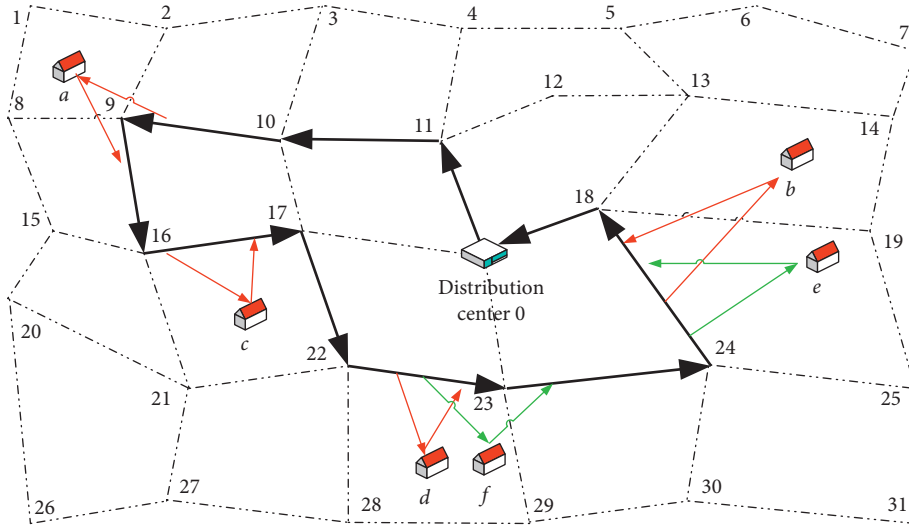


FIGURE 4: The movement route of a truck and drones.

φ_1 : speed of each drone, $\varphi_0 < \varphi_1$.

τ_k : the battery life of the drone k , determining its maximum endurance.

x_i “Launch site” of the drone released from the truck to Customer i .

y_i “Landing place” of the drone which returns onto the truck from Customer i .

2.4. *Decision Variables.* The decision variables for the routing problem of collaborative distribution system are listed as follows.

$\omega_{u,v}$ binary equals one if the directed line segment (u, v) belongs to a transport closed route \wp , while $u, v = 0, 1, \dots, n$; otherwise, equals zero:

$$\omega_{u,v} = \begin{cases} 1, & \text{if } (u, v) \in \wp, u \neq v, \\ 0, & \text{if } (u, v) \notin \wp, u \neq v, \\ 0, & \text{if } u = v. \end{cases} \quad (6)$$

$\theta_{i,k}$ binary equals one if Customer i is delivered by Drone k ; otherwise, equals zero.

Note that each Customer i corresponds to a Launching Point x_i and a Landing Point y_i for the drone. The drone starts from Launching Site x_i , sends goods to Customer i along a straight line, and then returns from Customer i along another straight line to Landing Site y_i on the truck. The details are shown in Figure 5, where the thick black line indicates the route of the truck, the blue dotted lines indicate the flight path of the first drone, the pink dotted lines indicate the flight path of the second drone, and the green one indicates the flight path of the third drone. x_{i-2} and y_{i-2} are, respectively, the launching point and landing point of the first drone for delivery of the customer point $i - 2$; x_{i-1} and y_{i-1} are that of customer point $i - 1$; x_i and y_i are the launching point and landing point of the second drone for

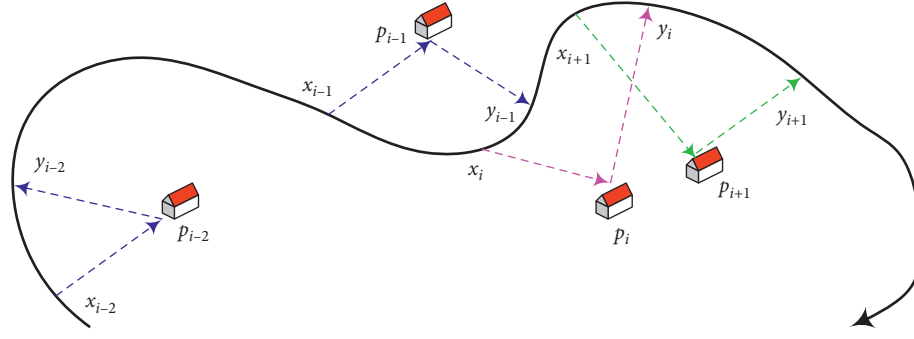


FIGURE 5: Launching sites and landing sites of drones.

delivery of the customer point i ; x_{i+1} and y_{i+1} are the launching point and landing point of the third drone for delivery of customer point $i + 1$.

2.5. Routing Problem of the Collaborative Distribution System. The routing problem of the coordinated system for multi-drone single-truck distribution can be defined as follows. Given a city road network $G = (V, E)$, the distribution centre is numbered as 0 while each customer is numbered as i ($i = 1, 2, \dots, h$); the location coordinates and demand quantity of all the customers are known. In addition, some assumptions are mentioned as follows to complete our problem:

- (i) Only one truck is provided, but is equipped with a certain number of drones
- (ii) A customer demand point can only be delivered by only one drone, while a drone is able to deliver to none or more customer demand points
- (iii) The weight of cargo for every customer is no more than the capacity of a drone
- (iv) All drones have the same speed
- (v) Each drone can only transport only one good each time
- (vi) The truck is just a platform for drones to launch and land, while itself does not visit customer demand points
- (vii) Drones could land on or fly off the truck no matter if the truck is motionless or locomotor
- (viii) The number of drones on each truck is fixed and the load weight of the truck excludes the total weight of drones on it, which means the weight of drones is negligible

Our problem aims to find out two sorts of routes in the shortest distribution time: one is the closed transport route \wp in which starting and terminal points are both distribution centre in the whole delivery of the truck, and the other is flight path of each drone on the truck, which means that launching and landing points of all drones are ought to be decided in this problem.

3. Mathematical Model

Some of relevant intermediate variables will be illustrated before recommending the mathematical model.

3.1. Intermediate Variables Related to the Closed Route of the Truck. We define a symbol δ_u that indicates the mileage of the truck from the distribution centre to Intersection u . Obviously, $\delta_0 = 0$.

If the intersection u is not on the transport closed route \wp , then there is no intersection v that makes $\omega_{v,u} = 1$. Set $\delta_u = -1$ when $u \notin \wp$.

If the intersection u is on the transport closed route \wp , then there is only one intersection v that makes $\omega_{v,u} = 1$; in other words, $\exists v, (v, u) \in \wp$. Because the length of the directed line segment (u, v) is

$$\|\mu_u - \mu_v\| = \sqrt{(\mu_u^1 - \mu_v^1)^2 + (\mu_u^2 - \mu_v^2)^2}, \quad (7)$$

so

$$\delta_u = \sum_{v=0}^n \omega_{v,u} \times \left(\delta_v + \sqrt{(\mu_u^1 - \mu_v^1)^2 + (\mu_u^2 - \mu_v^2)^2} \right). \quad (8)$$

In a word,

$$\delta_u = \begin{cases} 0, & \text{if } u = 0, \\ \sum_{v=0}^n \omega_{v,u} \times \left(\delta_v + \sqrt{(\mu_u^1 - \mu_v^1)^2 + (\mu_u^2 - \mu_v^2)^2} \right), & \text{if } u \neq 0, \forall u \in \wp, \\ -1, & \text{if } u \neq 0, \forall u \notin \wp. \end{cases} \quad (9)$$

3.2. Intermediate Variables Related to Launching and Landing Sites of Drones. In order to simplify the mathematical model, the following decision variables are utilized directly to substitute launching site x_i and landing site y_i .

$\beta_{i,u,v}$: a binary indicating whether the launching point x_i is located on the directed line segment (u, v) :

$$\beta_{i,u,v} = \begin{cases} 1, & \text{if } x_i \in (u, v), \\ 0, & \text{if } x_i \notin (u, v). \end{cases} \quad (10)$$

α_i : a decimal indicating dividing the distance between x_i and its start point u by the length of line segment (u, v) , so $\alpha_i \in [0, 1)$:

$$\alpha_i = \frac{\|x_i - \mu_u\|}{\|\mu_v - \mu_u\|}. \quad (11)$$

$\beta_{i,u,v}^*$: an integer indicating the line segment of \wp in which landing point y_i belongs to

$$\beta_{i,u,v}^* = \begin{cases} 1, & \text{if } y_i \in (u, v), \\ 0, & \text{if } y_i \notin (u, v). \end{cases} \quad (12)$$

α_i^* : a decimal indicating dividing the distance between y_i and its start point u by the length of line segment (u, v) , so $\alpha_i^* \in [0, 1)$:

$$\alpha_i^* = \frac{\|y_i - \mu_u\|}{\|\mu_v - \mu_u\|}. \quad (13)$$

Depending on the decision variables above, coordinates of the launching point x_i and landing point y_i could be calculated as follows:

χ_i^1 : the x -coordinate of the launching point:
 $\chi_i^1 = \sum_{u=1}^n \sum_{v=1}^n \beta_{i,u,v} \times (\mu_u^1 + \alpha_i \times (\mu_v^1 - \mu_u^1))$

χ_i^2 : the y -coordinate of the launching point:
 $\chi_i^2 = \sum_{u=1}^n \sum_{v=1}^n \beta_{i,u,v} \times (\mu_u^2 + \alpha_i \times (\mu_v^2 - \mu_u^2))$

γ_i^1 : the x -coordinate of the landing point:
 $\gamma_i^1 = \sum_{u=1}^n \sum_{v=1}^n \beta_{i,u,v}^* \times (\mu_u^1 + \alpha_i^* \times (\mu_v^1 - \mu_u^1))$

γ_i^2 : the y -coordinate of the landing point:
 $\gamma_i^2 = \sum_{u=1}^n \sum_{v=1}^n \beta_{i,u,v}^* \times (\mu_u^2 + \alpha_i^* \times (\mu_v^2 - \mu_u^2))$

Define a symbol ξ_i that indicates the mileage of the truck from the distribution centre to launching point χ_i :

$$\xi_i = \sum_{u=1}^n \sum_{v=1}^n \beta_{i,u,v} \times \left(\delta_u + \sqrt{(\chi_i^1 - \mu_u^1)^2 + (\chi_i^2 - \mu_u^2)^2} \right). \quad (14)$$

That is,

$$\xi_i = \sum_{u=0}^n \sum_{v=0}^n \beta_{i,u,v} \times (\delta_u + \alpha_i \times \|\mu_v - \mu_u\|). \quad (15)$$

Define a symbol ψ_i that indicates the mileage of the truck from the distribution centre to landing point γ_i :

$$\psi_i = \sum_{u=1}^n \sum_{v=1}^n \beta_{i,u,v}^* \times \left(\delta_u + \sqrt{(y_i^1 - \mu_u^1)^2 + (y_i^2 - \mu_u^2)^2} \right). \quad (16)$$

That is,

$$\psi_i = \sum_{u=0}^n \sum_{v=0}^n \beta_{i,u,v}^* \times (\delta_u + \alpha_i^* \times \|\mu_v - \mu_u\|). \quad (17)$$

The flight mileage of the drone when delivering to Customer i can also be calculated as follows:

$$\|\chi_i - p_i\| + \|p_i - \gamma_i\| = \sqrt{(\chi_i^1 - p_i^1)^2 + (\chi_i^2 - p_i^2)^2} + \sqrt{(p_i^1 - \gamma_i^1)^2 + (p_i^2 - \gamma_i^2)^2}. \quad (18)$$

3.3. Mathematical Model and Its Description.

$$\begin{aligned} \text{MinimizerR} = & \sum_{i=1}^h \left| \frac{1}{\phi_0} (\psi_i - \xi_i) - \frac{1}{\phi_1} (\|\chi_i - p_i\| + \|p_i - \gamma_i\|) \right| \\ & + \frac{1}{\phi_0} \sum_{u=0}^n \sum_{v=0}^n \omega_{u,v} \times \|\mu_u - \mu_v\|, \end{aligned} \quad (19)$$

subject to

$$\omega_{u,u} = 0, \quad \forall u \in V, \quad (20)$$

$$\omega_{u,v} \leq \varepsilon_{u,v}, \quad \forall u, v \in V, \quad (21)$$

$$\sum_{u=0}^n \omega_{u,v} \leq 1, \quad \forall v \in V, \quad (22)$$

$$\sum_{v=0}^n \omega_{u,v} \leq 1, \quad \forall u \in V \quad (23)$$

$$\sum_{v=0}^n \omega_{u,v} = \sum_{t=0}^n \omega_{t,u}, \quad \forall u \in V, \quad (24)$$

$$\delta_0 = 0, \quad (25)$$

$$\delta_u = \sum_{v=0}^n \omega_{v,u} \times (\delta_v + \|\mu_u - \mu_v\|), \quad \forall u \in V, u \neq 0, \quad (26)$$

$$\beta_{i,u,v} \leq \omega_{u,v}, \quad \forall u, v \in V, \forall i \in I, \quad (27)$$

$$\beta_{i,u,v}^* \leq \omega_{u,v}, \quad \forall u, v \in V, \forall i \in I, \quad (28)$$

$$\sum_{u=1}^n \sum_{v=1}^n \beta_{i,u,v} = 1, \quad \forall i \in I, \quad (29)$$

$$\sum_{u=1}^n \sum_{v=1}^n \beta_{i,u,v}^* = 1, \quad \forall i \in I, \quad (30)$$

$$\chi_i^1 = \sum_{u=1}^n \sum_{v=1}^n \beta_{i,u,v} \times (\mu_u^1 + \alpha_i \times (\mu_v^1 - \mu_u^1)), \quad \forall i \in I, \quad (31)$$

$$\chi_i^2 = \sum_{u=1}^n \sum_{v=1}^n \beta_{i,u,v} \times (\mu_u^2 + \alpha_i \times (\mu_v^2 - \mu_u^2)), \quad \forall i \in I, \quad (32)$$

$$y_i^1 = \sum_{u=1}^n \sum_{v=1}^n \beta_{i,u,v} \times (\mu_u^1 + \alpha_i^* \times (\mu_v^1 - \mu_u^1)), \quad \forall i \in I, \quad (33)$$

$$y_i^2 = \sum_{u=1}^n \sum_{v=1}^n \beta_{i,u,v}^* \times (\mu_u^2 + \alpha_i^* \times (\mu_v^2 - \mu_u^2)), \quad \forall i \in I, \quad (34)$$

$$\xi_i = \sum_{u=0}^n \sum_{v=0}^n \beta_{i,u,v} \times (\delta_u + \alpha_i \times \|\mu_v - \mu_u\|), \quad \forall i \in I, \quad (35)$$

$$\psi_i = \sum_{u=0}^n \sum_{v=0}^n \beta_{i,u,v}^* \times (\delta_u + \alpha_i^* \times \|\mu_v - \mu_u\|), \quad \forall i \in I, \quad (36)$$

$$\xi_i \leq \psi_i, \quad \forall i \in I, \quad (37)$$

$$\sum_{k=1}^m \theta_{i,k} = 1, \quad \forall i \in I, \quad (38)$$

$$\sum_{i=1}^h \theta_{i,k} \times \frac{1}{\phi_1} (\|\chi_i - p_i\| + \|p_i - \gamma_i\|) \leq \tau_k, \quad \forall k \in K, \quad (39)$$

$$(\xi_i - \xi_j) \times (\xi_i - \psi_j) \times \theta_{i,k} \times \theta_{j,k} \geq 0, \quad \forall i, j \in I, \forall k \in K, \quad (40)$$

$$(\psi_i - \xi_j) \times (\psi_i - \psi_j) \times \theta_{i,k} \times \theta_{j,k} \geq 0, \quad \forall i, j \in I, \forall k \in K, \quad (41)$$

$$\sum_{i=1}^h g_i \leq q. \quad (42)$$

Here, $\omega_{u,v}$, α_i , α_i^* , $\beta_{i,u,v}$, $\beta_{i,u,v}^*$, and $\theta_{i,k}$ are decision variables with $0 \leq \alpha_i < 1$ and $0 \leq \alpha_i^* < 1$. In particular, $\omega_{u,v}$, $\beta_{i,u,v}$, $\beta_{i,u,v}^*$, or $\theta_{i,k}$ is an integer for 0 or 1, while α_i or α_i^* is a decimal between 0 and 1, that is, $\alpha_i \in [0, 1)$ and $\alpha_i^* \in [0, 1)$.

The objective function (19) minimizes the total delivery time. We do not simply add the truck time and drone time to obtain the whole completion time during delivery tasks. This is because these two vehicles need to be synchronized at the drone's landing point, which means that the truck have to stop and wait for the drone if it arrives to the landing point early, while the drone is also ought to wait until the truck comes if the drone arrives to the landing point early.

In function (19), the first term $\sum_{i=1}^h |(1/\phi_0)(\psi_i - \xi_i) - (1/\phi_1)(\|\chi_i - p_i\| + \|p_i - \gamma_i\|)|$ represents the total time which the truck spends to wait for the drone or the drone spends to wait for the truck in all pairs of launching point χ_j and landing point γ_j . In this term, $(1/\phi_0)(\psi_i - \xi_i)$ shows the time which the truck takes from a launch point to the corresponding landing point, and $(1/\phi_1)(\|\chi_i - p_i\| + \|p_i - \gamma_i\|)$ indicates the total time that the drone takes to leave the truck from the launching point, reach the customer demand point, and then return to the truck at the landing point. The second term $(1/\phi_0) \sum_{u=0}^n \sum_{v=0}^n \omega_{u,v} \times \|\mu_u - \mu_v\|$ represents the total time of the truck on the closed loop when the drone and the truck are fully synchronized for each delivery. In other words, flight time of the drone is completely equal to the time of the truck for each pair of take-off points and landing points.

Constraint (20) indicates that there is no directed line segment from the intersection to itself. Constraint (21) says if $\varepsilon_{u,v} = 0$, that is to say $(u, v) \notin E$, then $\omega_{u,v} = 0$. Otherwise, if $\varepsilon_{u,v} = 1$, then $\omega_{u,v} = 0$ or $\omega_{u,v} = 1$. Constraint (22) shows that the number of input lines of any Intersection v cannot exceed 1. Constraint (23) ensures that the quantity of input lines of any Intersection u is 1 at most. These two constraints are used to guarantee that any line segment is passed by the truck only once. Constraint (24) means that an output or input line segment of an intersection is allocated to at most one input or output line segment or not. These constraints enable a closed loop of the truck, and there is no need for it to travel through all intersections in the road network. Instead, the truck just requires to go through part of the intersections, starting from and returning back to the distribution centre. Constraint (25) indicates that the total distance of the truck driving from the distribution centre to the distribution centre is 0. Constraint (26) represents the total distance of the truck driving from the distribution centre to the intersection u , $\forall u \in V, u \neq 0$. Constraint (27) says if $\omega_{u,v} = 0$, that is to say $(u, v) \notin \emptyset$, then $\beta_{i,u,v} = 0$. Otherwise, if $\omega_{u,v} = 1$, then $\beta_{i,u,v} = 0$ or $\beta_{i,u,v} = 1$. Constraint (28) says if $\omega_{u,v} = 0$, that is to say $(u, v) \notin \emptyset$, then $\beta_{i,u,v}^* = 0$. Otherwise, if $\omega_{u,v} = 1$, then $\beta_{i,u,v}^* = 0$ or $\beta_{i,u,v}^* = 1$. Constraint (29) indicates that a launching point can only be on one directed line segment. Constraint (30) indicates that a landing point can only be on one directed line segment. Constraints (31) and (32) represent the x -coordinate and y -coordinate of the launching point, respectively, while Constraints (33) and (34) are the x -coordinate and y -coordinate of the landing point, respectively. Constraints (35) and (36) show the mileage of the truck from distribution centre to launching point and to landing point separately. Constraint (37) imposes that, for the drone of each customer, the launching point corresponded on the truck trajectory is earlier than the landing point. Constraint (38) guarantees that a customer demand point is allocated to only one drone, while a drone could deliver to none or some customer demand points. Constraint (39) states the relationship between endurance mileage of a drone and its battery life. Constraints (40) and (41) ensure that for any Customer i and Customer j which are delivered by the same drone, both launching point χ_i and landing point γ_i , are either earlier than launching point χ_j or

later than landing point γ_j , which means that χ_i and γ_i are not able to locate between χ_j and γ_j on the truck trajectory. Constraint (42) imposes that the gross weight of goods loading on the truck is no more than the maximum carrying weight q .

Since our problem is a new extension of the Vehicle Routing Problem (VRP) which has been shown to be NP-hard, the problem in this paper is also NP-hard and the model is nonlinear. In addition, CPLEX is always used to solve small-scale vehicle routing problem, but not to large-scale vehicle routing problem due to the fact of low solution efficiency. Therefore, we utilize heuristic algorithm in this paper for large-scale problem.

4. Algorithm

With the increasing number of decision variables, problem scale, and search space of genetic algorithm (GA) or particle swarm optimization (PSO) for a complex problem, performances of algorithm is impacted negatively due to the slow convergence rate of algorithm and large amount of iterations.

In order to improve the convergence speed, this paper designs a hybrid genetic algorithm (h-GA) and a hybrid particle swarm optimization (h-PSO). The basic idea of these two hybrid algorithms is to decompose a problem into a master problem and a child problem and just encode decision variables in the master problem as a chromosome or a particle instead of encoding them all. Decision variables of the child problem are solved by a nonlinear programming algorithm or a heuristic method. This way greatly reduces the search space of genetic algorithm or particle swarm algorithm and then enhances the convergence rate to a great extent.

4.1. Hybrid Genetic Algorithm. Hybrid genetic algorithm breaks down this problem into a master problem and a child problem. The master problem is to device the best closed transport route, which the truck along will start from and return to the distribution centre after delivery tasks are completed.

According to the closed route solved by the master problem, the child problem is utilized for every customer to plan a best flight route, which the drone flies along from the truck to the customer and returns back to the moving truck after present delivering. As the demand and coordinate location of each customer is known, this child problem is equivalent to finding out the best launching point and landing point of a drone for each customer on the known closed transport route of the truck.

4.1.1. Genetic Algorithm of the Master Problem. In this section, the master problem is solved by genetic algorithm. Figure 6 shows the flowchart of this algorithm.

The main steps of the genetic algorithm are encoding and decoding of chromosomes and the fitness function. Please refer to relevant documents for details of other steps.

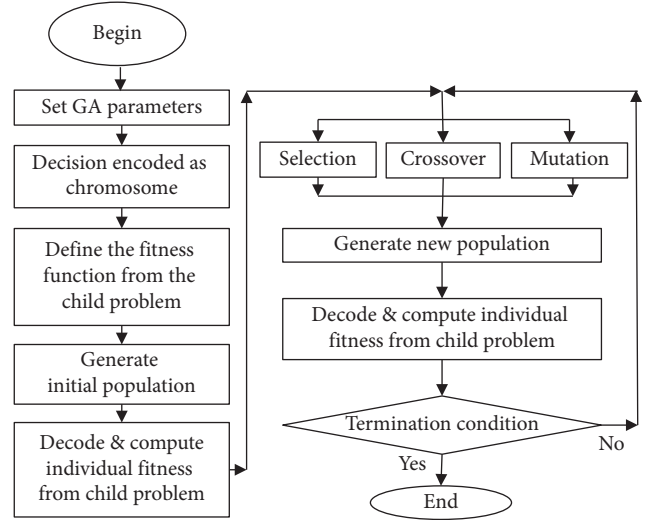


FIGURE 6: Genetic algorithm of the master problem.

(1) *Encoding and Decoding of Chromosomes of the Master Problem.* Because the master problem is simply solving a best closed transport line of the truck, the only decision variable needs to be encoded into chromosomes is $\omega_{u,v}$, where $u, v \in \{0, \dots, n\}$. Here, we adopt natural number encoding for $\omega_{u,v}$. On the contrary, other decision variables such as α_i , $\beta_{i,u,v}$, α_i^* , $\beta_{i,u,v}^*$, and $\theta_{i,k}$ do not need to be encoded into chromosomes, while will be solved by the child problem. As shown in Figure 7, genetic chromosomes just contain $\omega_{u,v}$, which is either 0 or 1. In this way, search space of the genetic algorithm is drastically reduced, and decision variables in the master problem can be solved in short time. For instance, if the closed loop of the truck is as

$$0 \rightarrow 11 \rightarrow 10 \rightarrow 9 \rightarrow 16 \rightarrow 17 \rightarrow 22 \rightarrow 23 \rightarrow 24 \rightarrow 18 \rightarrow 0, \quad (43)$$

then, in the chromosomes, the values of genes $\omega_{0,11}$, $\omega_{11,10}$, $\omega_{10,9}$, $\omega_{9,16}$, $\omega_{16,17}$, $\omega_{17,22}$, $\omega_{22,23}$, $\omega_{23,24}$, $\omega_{24,18}$, and $\omega_{18,0}$ are 1, and the value of the rest genes are 0.

For a chromosome, if the line segments of its related genes with value 1 can form a closed loop from distribution centre and back to it, the chromosome corresponds to a feasible solution of the master problem.

(2) *The Fitness Function of the Master Problem.* The fitness function shows whether the solution to the master problem is good or not, and the degree is depended on the following factors. The first factor is to see if this solution to the master problem satisfies the constraints; the second one is to see if the solution to its related child problem meets the constraints; and the third one is to see whether the target value of its child problem is good or not. Specifically, the solution to the master problem requires to be a closure of the path of the truck.

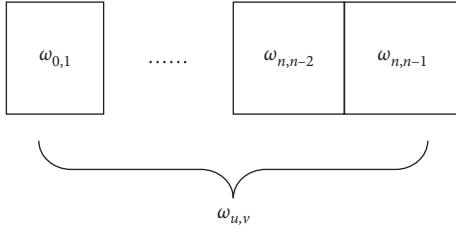


FIGURE 7: Chromosome of the master problem.

Then, the best flight route of the drone for each customer and target value of child problem are figured out based on the solution of the master problem.

The solution in child problem needs to consider the following conditions: (1) sequence of departure points; (2) landing points of the drone; (3) maximum endurance of the drone; (4) total weight of demand goods, which is at most the maximum load of the truck.

(3) *Random Generation of Initial Populations.* Each gene is randomly generated within its range of values so as to get a random chromosome. Repeat the above operation until a certain number of chromosomes are produced.

(4) *Replication of Chromosomes and Selection of the Elite.* Copy a certain number of chromosomes into the next generation with a roulette method. And for each generation of chromosomes, choose the best fitness individuals depending on a certain selection rate to replace the lowest one for the next generation.

(5) *Crossover Operation.* In the last generation, select two individuals randomly as fathers, randomly identify an intersection, and then perform the crossover operation.

(6) *Mutation Operation.* Randomly select mutation points in the chromosome, and randomly generate a value for the genes at these mutation points.

(7) *Construct New Population.* Production of individuals in a new population comes in three approaches: (1) produce new individuals by crossover operation; (2) generate them by mutation operation; (3) select 10% of the optimal fitness individuals of the original population.

(8) *Termination Condition.* If the number of iterations in the algorithm reaches the specified maximum, terminate the algorithm and output the optimal solutions. Otherwise, return to Step 4 and continue to perform the genetic operations above.

4.1.2. Heuristic Method of the Child Problem. This child problem enables to figure out the best take-off point and landing point for each customer demand point on a known closed transport line of the truck. Since the child problem is nonlinear programming, it can be solved by a heuristic method. Figure 8 shows the details of the heuristic.

(1) Determine the order of assigned customer points.

First, the child problem needs to determine the order of allocated customer points which the drone deliveries. If the total number of customers is n , the amount of all possible delivery orders is $n!$. The calculated amount will be enormous if all possible delivery orders are solved with the child problem. Thus, a construction method is used here to get a better delivery order in a shorter time.

Given the closed transport route φ of the truck, solution of this construction method is as follows. For each customer demand point p_i , calculate point \perp_i which is the closest to p_i on Route φ . Then, according to the sequence of these points \perp on Route φ , determine the order of assigned customer points. Specifically, if \perp_i is earlier than \perp_j on Route φ , then corresponding demand point p_i in the delivery order is arranged before p_j . On the contrary, if \perp_i is later than \perp_j , p_i is arranged after p_j .

(2) Given the truck route and delivery order as above, plan the drone route

When the truck closed route φ and customer delivery order are solved, this child problem is transformed into a convex optimization problem for the take-off point and landing point of the drone and can also be solved by a structured method as follows.

When solving customer's delivery order following the above step, point \perp_i which is the closest to each point p_i on Route φ is calculated. The corresponding decision variables for the delivery of a drone to the customer p_i are x_i and y_i . The truck and a drone are ought to avoid waiting for each other at a landing point in order to guarantee time effect. In other words, the best case is a drone and the truck arrives at the landing point at the same time. Thus, the total delivery time of the drone is equal to that of the truck:

$$\frac{1}{\phi_0} (\psi_i - \xi_i) = \frac{1}{\phi_1} (\|\chi_i - p_i\| + \|p_i - \gamma_i\|). \quad (44)$$

Assuming that take-off point x_i and landing point y_i are symmetric about the line segment from customer point p_i to point \perp_i , the flying distance of the drone from x_i to p_i is equal to that from p_i to y_i , and the driving distance of the truck from x_i to \perp_i is equal to that from \perp_i to y_i . So, an approximation can be derived:

$$\frac{1}{\phi_0} \|\chi_i - \perp_i\| \approx \frac{1}{\phi_1} \|\chi_i - p_i\|. \quad (45)$$

Assume that the line segment from customer point p_i to point \perp_i is perpendicular to the line segment from x_i to y_i . The approximate solution of the take-off point x_i and landing point y_i can be figured out according to the Pythagorean theorem, respectively.

(3) Solve decision variable $\theta_{i,k}$ according to the order of delivery and the drone route.

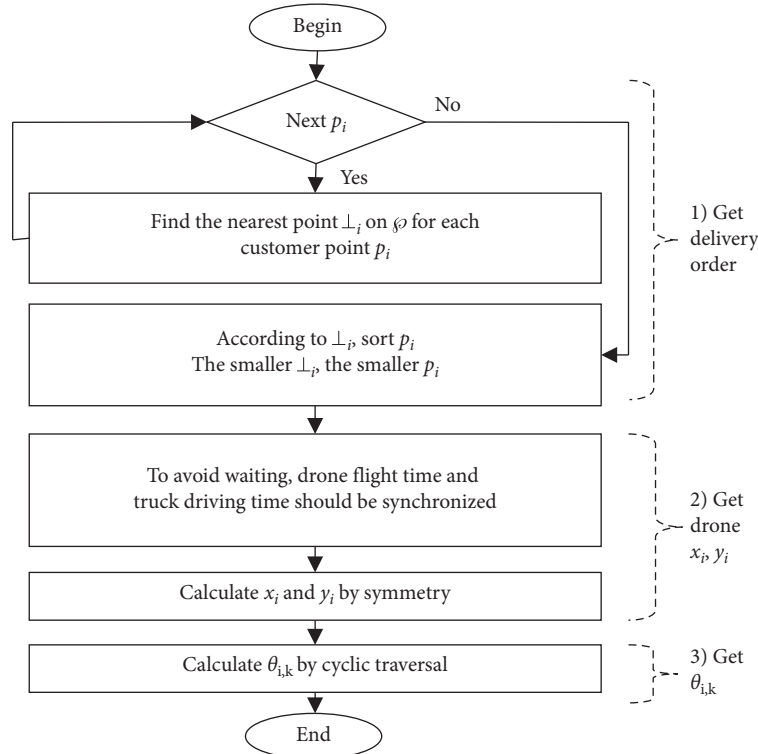


FIGURE 8: Heuristic algorithm of the child problem.

After the delivery order and drone route are solved by the above steps 1 and 2, decision variables $\theta_{i,k}$ can be solved in order by the sequence loop traversal method as follows. For each drone k , choose each unvisited customer demand point in turn according to the solved delivery order. If a customer point is the i th one for Drone k to visit successfully, then set $\theta_{i,k} = 1$. Otherwise, if launching point of the i th customer point is before the landing point of the $(i - 1)$ th customer point, set $\theta_{i,k} = 0$ and turn to consider the next customer point until all points are delivered.

4.2. Hybrid Particle Swarm Optimization. Similar to the hybrid genetic algorithm, the hybrid particle swarm algorithm here also divides the routing problem into a master problem and a child problem.

4.2.1. Particle Warm Optimization of the Master Problem. In this section, the master problem is solved by particle swarm optimization, the flowchart of which is shown in Figure 9.

The major steps of particle swarm optimization are the expression of particles and the evaluation function. For detailed information on its other steps, please refer to relevant documents.

- (1) Expression of the particle of the master problem.

The key to implement the particle swarm optimization is to find a suitable expression of the particle. Here, the

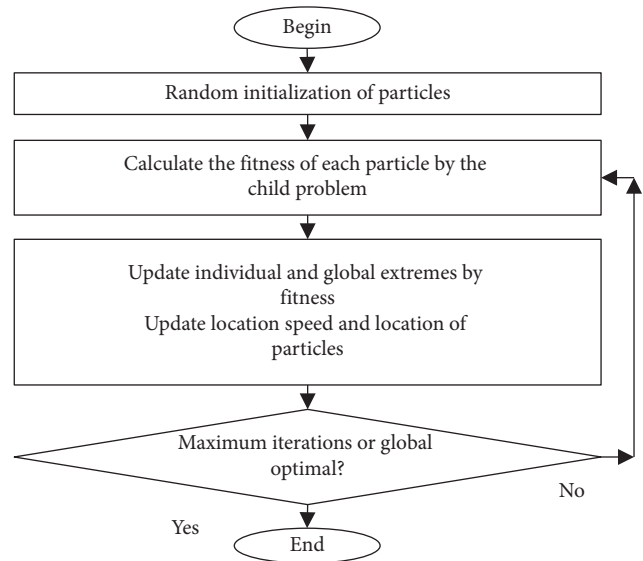


FIGURE 9: Particle swarm optimization of the master problem.

expression only contains decision variable $\omega_{u,v}$, where $u, v \in \{0, \dots, n\}$, and the value of $\omega_{u,v}$ is either 0 or 1.

- (2) Evaluation function of the master problem.

The evaluation function indicates whether a solution to the master problem is good or not. Criteria depend on the following factors: (1) this solution to the master problem satisfies the constraints; (2) this solution to its child problem satisfies corresponding constraints; (3) the target value of its corresponding child problem is good.

The implementation steps of the particle swarm algorithm are as follows:

- (I) Initialize the particle swarm.
 - (1) Divide the particle group into numbers of overlapping subgroups
 - (2) Randomly generate each particle X , where $\omega_{u,v}$ is set either 0 or 1 randomly
 - (3) Randomly generate each velocity vector V , each dimension of which is set with a real number between $-Vmax_d$ and $Vmax_d$ randomly
 - (4) Evaluate all particles with evaluation functions
 - (5) Set the initial evaluation value as optimal solution P_i for individual history and then search optimal solution P_l and P_g in the subgroups and total group, respectively
- (II) Repeat the following steps until satisfying the termination condition or reaching the maximum number of iterations.
 - (1) For each particle X , the latest velocity vector V is calculated according to the following speed update formula:

$$v_{id}(t+1) = w \times v_{id}(t) + c_1 \times \text{rand}(\cdot) \times [p_{id}(t) - x_{id}(t)] + c_2 \times \text{rand}(\cdot) \times [p_{gd}(t) - x_{id}(t)]. \quad (46)$$

Calculate the latest position vector according to the following position update formula:

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1). \quad (47)$$

If velocity vector V and position vector X exceed its range, then the boundary values are evaluated.

- (2) Evaluate all particles with evaluation functions.
- (3) If current evaluation of a certain particle is better than the past best evaluation, it will be set as the optimal evaluation value and the current location will be set as the optimal position of the last particle.
- (4) Find the optimal solution in the current subgroup and total group. If it is better than the historical best one, update P_l and P_g . If all the individuals in a subgroup are unsolvable, or multiple individuals in the subgroup have the same best values, then randomly take one of them as the current optimal solution in the subgroup.

4.2.2. Heuristic Method of the Child Problem. This child problem is equivalent to figuring out the best take-off point and the best landing point for each customer demand point on a known closed transport line of the truck. This is a nonlinear programming problem, which can be solved by using the heuristic method, as shown in Figure 8.

5. Experiment

In this section, we conduct numerical experiments using a PC (Intel i5, 2.90 GHz, 8 GB of Memory) in Windows 10 to validate the effectiveness of our models and efficiency of our solution methods. The algorithms mentioned above are implemented by C#.

5.1. Experimental Configuration. Figure 10 is the outcome of a program which is used in this experiment to generate a road network randomly, where small red dots represent customer points. Apart from producing any number of customer demand points at any position coordinates, this program enables to generate any size of road network randomly, which is denoted by numbers of line segments. Take some customer points, for example, Figure 11 shows the distribution route calculated by the algorithm, where blue lines and green lines present the flying routes of the first drone and the second drone, respectively, black lines mean the driving route truck, while small red dots represent customer demand points and the blue pentagon in the centre means distribution centre.

In order to test universality of results, we use same input parameters in this experiment with any number of demand points in a road network which is in the size of large, medium, and small, respectively. Besides the distribution centre which is noted as serial number 0, position coordinate and quantity of goods for each customer are generated in a random way.

5.2. Results and Discussion. Figure 12 shows the comparison of four algorithms in different scales, where the horizontal coordinate represents the number of iterations and the vertical coordinate represents current best fitness. Higher fitness means being closer to the optimal solution. In each graph, blue line shows the result of basic genetic algorithm (GA), orange line is that of hybrid genetic algorithm (h-GA), while grey line and yellow line present the result of basic particle swarm algorithm (PSO) and the hybrid particle swarm algorithm (h-PSO), respectively.

It is obvious to see from Figure 12 that, in experiments (a), (b), (e), and (f), the current optimal fitness of h-PSO exceeds all other three algorithms after about 70 iterations. In (c), (g), and (h), the current optimal fitness of h-PSO exceeds all other three algorithms after about 200 iterations. In (d), the current best fitness of h-PSO performs better after approximately 800 iterations.

However, from experiment (a), (b), and (d), we can observe that the current optimal fitness of h-GA is basically a little better than GA and much worse than PSO, while it is better than GA and PSO after about 500 iterations. The situation that current optimal fitness of h-GA is better than GA and PSO in experiments (c) and (g) happens after around 800 iterations, and the same case is shown in experiments (e) and (f) after about 200 iterations and 400 iterations, respectively. In (h), h-GA is basically a little better than GA, and a little worse than PSO.

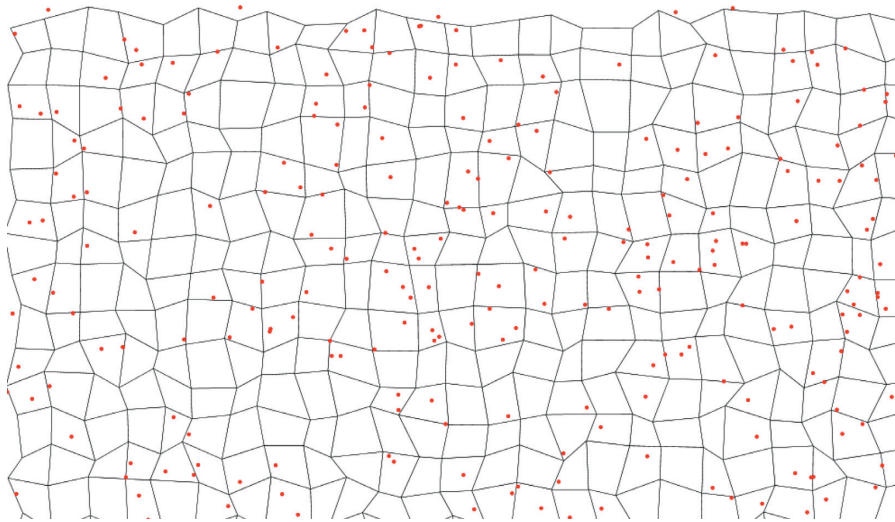


FIGURE 10: Road network in the algorithm experiment.

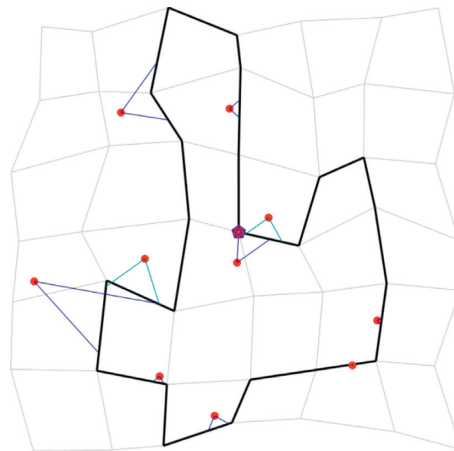


FIGURE 11: Distribution route calculated by the algorithm.

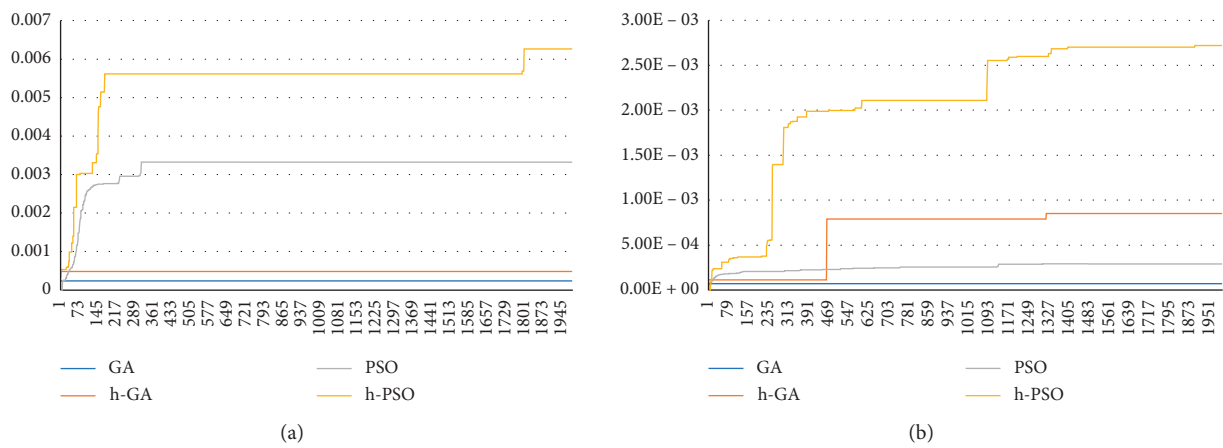


FIGURE 12: Continued.

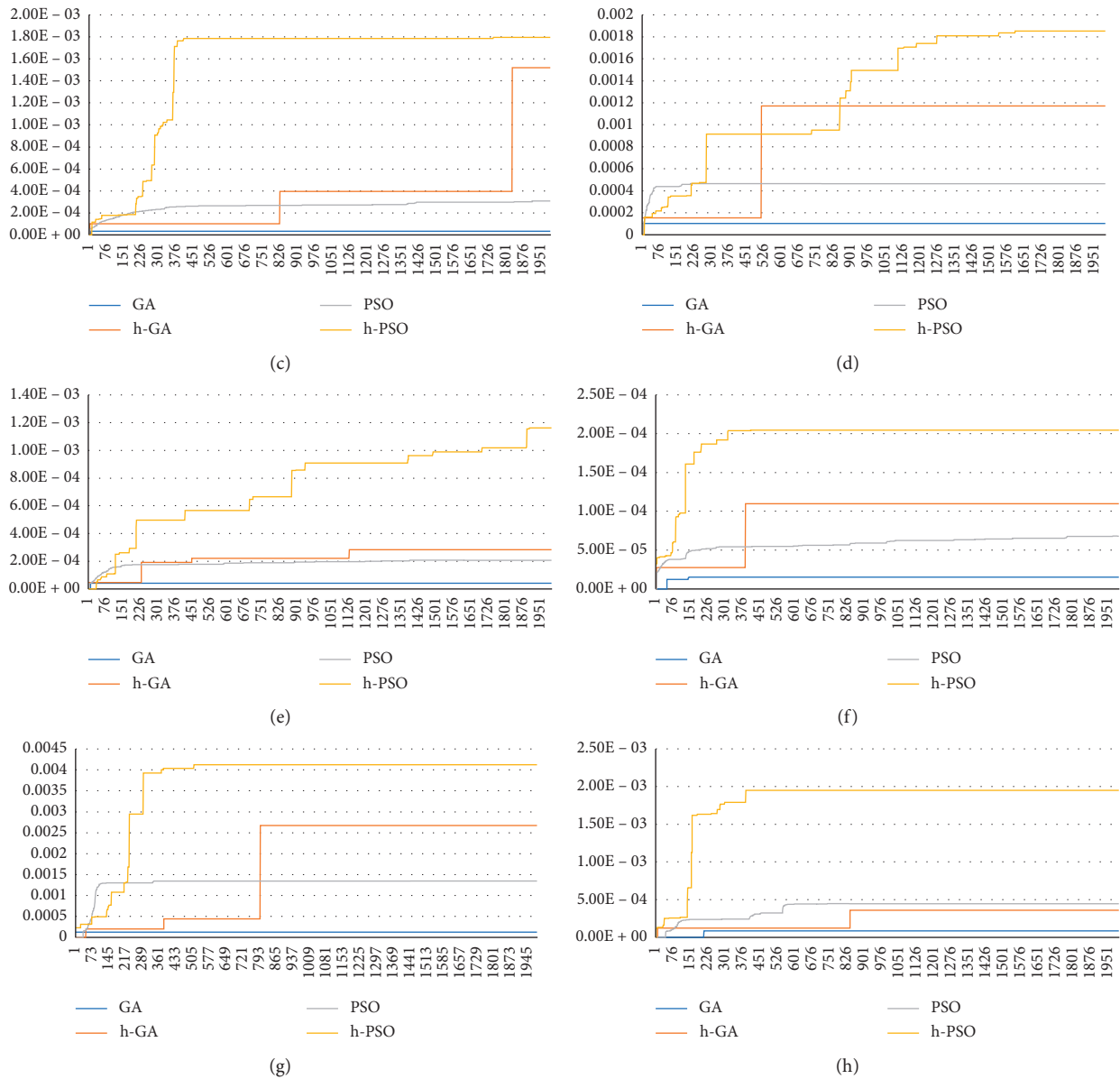


FIGURE 12: Algorithm comparison experiment. (a) Network size is 100 and customer number is 20. (b) Network size is 100 and customer number is 60. (c) Network size is 100 and customer number is 100. (d) Network size is 400 and customer number is 20. (e) Network size is 400 and customer number is 60. (f) Network size is 400 and customer number is 100. (g) Network size is 900 and customer number is 10. (h) Network size is 900 and customer number is 20.

Figures 13(a) and 13(b) are the comparison results of the average convergence fitness and average convergence iteration number of the four algorithms, respectively, which are the results of taking the average value after each algorithm is run for 50 times, respectively. In Figure 13, experiments (a), (b), and (c) are realized with the scale of 100 networks, while the number of customers is 20, 60, and 100, respectively. In experiments (d), (e), and (f), network size was replaced with 400, and customer number is also 20, 60, and 100, respectively. Network size in experiments (g) and (h) is 900. The number of customers in experiment (g) is 10 and that in experiment (h) is 20.

It can be seen from Figure 13(a) that, from experiments (a) to (h), the average convergence fitness of h-PSO is better

than the other three algorithms. In (a) and (h), the average convergence fitness of h-GA is better than GA and worse than PSO, while in (b), (c), (d), (f), and (g), the average convergence fitness of h-GA is better than both GA and PSO.

From Figure 13(b), we can see that, in experiment (a), when average convergence iteration number is about 500, the average convergence fitness of $h=$ PSO is around 70 while that of h-GA and PSO are no more than 20 and 40, respectively. Besides, GA only achieves an average convergence fitness which is less than 2 when the average convergence iteration number is about 250. Therefore, we can conclude that the solving efficiency of h-PSO is the best, and that of PSO is the second, which is better than h-GA and GA.

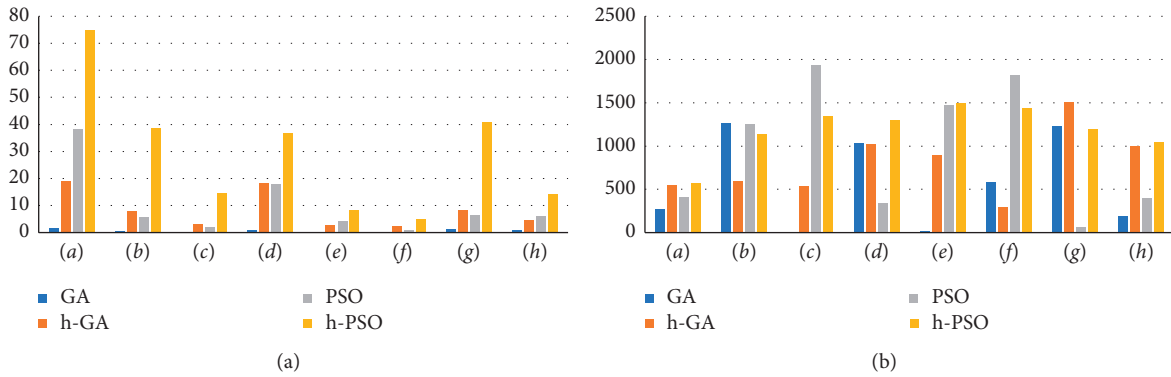


FIGURE 13: (a) Comparison of average convergence fitness between hybrid algorithms and basic algorithms. (b) Comparison of average convergence iterations between hybrid algorithms and basic algorithms.

In experiment (b), the solution efficiency of h-PSO is the best (the fitness is 40 after about 1000 iterations), that of h-GA is the second (the fitness is below 10 after about 500 iterations), that of PSO is the third (the fitness is below 10 after about 1000 iterations), and the efficiency of GA is the worst (the fitness is below 2 after about 1000 iterations). Experimental results are similar in (c)–(h). In short, the solving efficiency of these four algorithms could be sorted from the best to the worst as follows: h-PSO, PSO, h-GA, and GA.

Moreover, compared with basic algorithms, the corresponding hybrid algorithms are able to achieve better fitness with fewer iterations and thus achieve a better balance between solution efficiency and solution quality. In addition, the hybrid particle swarm optimization (h-PSO) is the best among all algorithms in terms of solution efficiency and quality.

6. Conclusions

This paper studies a coordinated system for multidrone single-truck distribution, where a truck carrying goods and a number of drones, starting from the distribution centre, will deliver goods to a group of customers along a closed ground path with the help of multiple drones and return to the distribution centre after completing the delivery tasks. In short, the truck assigns delivery tasks to drones, each of which is responsible for different subgroups to send goods to customers. Our study presents a new mixed-integer programming model for the routing problem of this distribution system based on an urban road network. After that, the hybrid genetic algorithm and hybrid particle swarm algorithm are designed. Experimental results show that effectiveness of the hybrid algorithms is better than that of corresponding basic algorithms for the routing problem. Our collaborative distribution model is suitable for the road networks of various cities because trucks are allowed to pass on this kind of network. Besides, in some areas such as remote mountains and islands, road for trucks is usually far away from customer points, which leads to the distance between trucks and customer points being very long for drones to delivery under the restrictions of electric quantity and

flight radius of drones. To solve this problem, better battery with bigger battery capacity could be changed, and part of the parameter in our model could be recomposed for different circumstances.

Some future research directions are suggested as follows. First, consider other more additional constraints that exist in reality, such as goods, must be delivered within the time specified by the customer. Second, add uncertainties in collaborative distribution route planning, including urban road conditions with a high degree of randomness such as vehicle breakdowns and traffic accidents, occasional drone malfunction, and uncertain customer demand. Third, consider better types of algorithms. Fourth, combine the simulation technique and the optimization algorithm to build a simulation optimization model to solve the synergistic distribution path problem.

Data Availability

No data were used to support this study.

Conflicts of Interest

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

References

- [1] C. C. Murray and R. Raj, "The multiple flying sidekicks traveling salesman problem: parcel delivery with multiple drones," *Transportation Research Part C: Emerging Technologies*, vol. 110, pp. 368–398, 2020.
- [2] L. Zhen, Y. Hu, S. Wang, G. Laporte, and Y. Wu, "Fleet deployment and demand fulfillment for container shipping liners," *Transportation Research Part B: Methodological*, vol. 120, pp. 15–32, 2019.
- [3] X. Bai, M. Cao, W. Yan, and S. S. Ge, "Efficient routing for precedence-constrained package delivery for heterogeneous vehicles," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 1, pp. 248–260, 2020.
- [4] S. Poikonen and B. Golden, "Multi-visit drone routing problem," *Computers & Operations Research*, vol. 113, Article ID 104802, 2020.

- [5] J. C. de Freitas and P. H. V. Penna, "A variable neighborhood search for flying sidekick traveling salesman problem," *International Transactions in Operational Research*, vol. 27, no. 1, pp. 267–290, 2020.
- [6] L. Zhen, "Tactical berth allocation under uncertainty," *European Journal of Operational Research*, vol. 247, no. 3, pp. 928–944, 2015.
- [7] Q. M. Ha, Y. Deville, Q. D. Pham, and M. H. Hà, "A hybrid genetic algorithm for the traveling salesman problem with drone," *Journal of Heuristics*, vol. 26, no. 2, pp. 219–247, 2019.
- [8] L. Zhen, Z. Liang, D. Zhuge, L. H. Lee, and E. P. Chew, "Daily berth planning in a tidal port with channel flow control," *Transportation Research Part B: Methodological*, vol. 106, pp. 193–217, 2017.
- [9] Y. Liu, "An optimization-driven dynamic vehicle routing algorithm for on-demand meal delivery using drones," *Computers & Operations Research*, vol. 111, pp. 1–20, 2019.
- [10] D. Schermer, M. Moeini, and O. Wendt, "A hybrid VNS/Tabu search algorithm for solving the vehicle routing problem with drones and en route operations," *Computers & Operations Research*, vol. 109, pp. 134–158, 2019.
- [11] D. Schermer, M. Moeini, and O. Wendt, "A matheuristic for the vehicle routing problem with drones and its variants," *Transportation Research Part C: Emerging Technologies*, vol. 106, pp. 166–204, 2019.
- [12] H. Y. Jeong, B. D. Song, and S. Lee, "Truck-drone hybrid delivery routing: payload-energy dependency and No-Fly zones," *International Journal of Production Economics*, vol. 214, pp. 220–233, 2019.
- [13] L. Zhen, "Modeling of yard congestion and optimization of yard template in container ports," *Transportation Research Part B*, vol. 90, pp. 83–104, 2016.
- [14] K. Wang, B. Yuan, M. Zhao, and Y. Lu, "Cooperative route planning for the drone and truck in delivery services: a bi-objective optimisation approach," *Journal of the Operational Research Society*, 2019.
- [15] D. Sacramento, D. Pisinger, and S. Ropke, "An adaptive large neighborhood search metaheuristic for the vehicle routing problem with drones," *Transportation Research Part C: Emerging Technologies*, vol. 102, pp. 289–315, 2019.
- [16] A. Karak and K. Abdelghany, "The hybrid vehicle-drone routing problem for pick-up and delivery services," *Transportation Research Part C: Emerging Technologies*, vol. 102, pp. 427–449, 2019.
- [17] M. Moshref-Javadi, S. Lee, and M. Winkenbach, "Design and evaluation of a multi-trip delivery model with truck and drones," *Transportation Research Part E: Logistics and Transportation Review*, vol. 136, 2020.
- [18] 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.
- [19] P. L. Gonzalez-R, D. Canca, J. L. Andrade-Pineda, M. Calle, and J. M. Leon-Blanco, "Truck-drone team logistics: a heuristic approach to multi-drop route planning," *Transportation Research Part C: Emerging Technologies*, vol. 114, pp. 657–680, 2020.
- [20] Y. S. Chang and H. J. Lee, "Optimal delivery routing with wider drone-delivery areas along a shorter truck-route," *Expert Systems with Applications*, vol. 104, pp. 307–317, 2018.
- [21] J. G. Carlsson and S. Song, "Coordinated logistics with a truck and a drone," *Management Science*, vol. 64, no. 9, pp. 4052–4069, 2018.
- [22] J. Xia, K. Wang, and S. Wang, "Drone scheduling to monitor vessels in emission control areas," *Transportation Research Part B: Methodological*, vol. 119, pp. 174–196, 2019.
- [23] N. Agatz, P. Bouman, and M. Schmidt, "Optimization approaches for the traveling salesman problem with drone," *Transportation Science*, vol. 52, no. 4, pp. 965–981, 2018.

Research Article

Multidepot Heterogeneous Vehicle Routing Problem for a Variety of Hazardous Materials with Risk Analysis

Bochen Wang,¹ Qiyuan Qian ,² Zheyi Tan,² Peng Zhang,¹ Aizhi Wu,¹ and Yi Zhou¹

¹Beijing Academy of Safety Science and Technology, Yunhe East Street 57, Beijing 101101, China

²School of Management, Shanghai University, Shang Da Road 99, Shanghai 200444, China

Correspondence should be addressed to Qiyuan Qian; qy_qian@shu.edu.cn

Received 2 July 2020; Revised 29 July 2020; Accepted 4 August 2020; Published 28 August 2020

Academic Editor: Xiaobo Qu

Copyright © 2020 Bochen Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study investigates a multidepot heterogeneous vehicle routing problem for a variety of hazardous materials with risk analysis, which is a practical problem in the actual industrial field. The objective of the problem is to design a series of routes that minimize the total cost composed of transportation cost, risk cost, and overtime work cost. Comprehensive consideration of factors such as transportation costs, multiple depots, heterogeneous vehicles, risks, and multiple accident scenarios is involved in our study. The problem is defined as a mixed integer programming model. A bidirectional tuning heuristic algorithm and particle swarm optimization algorithm are developed to solve the problem of different scales of instances. Computational results are competitive such that our algorithm can obtain effective results in small-scale instances and show great efficiency in large-scale instances with 70 customers, 30 vehicles, and 3 types of hazardous materials.

1. Introduction

In the past few years, the rapid development of China's chemical industry has led to the expansion of the road transportation market for hazardous materials. Thus, the transportation demand of hazardous materials presents a high-speed rising trend. Road transportation is still the main mode of transportation for hazardous materials, with the advantage of its strong mobility, flexibility, continuous transportation capacity, and less restrictive requirements. In contrast, the disadvantages of road transportation are the greatest potential risk, high unit transportation costs, and variability in operating conditions. Due to the characteristics of being flammable, explosive, corrosive, toxic, and radioactive, dangerous goods often cause more serious secondary hazards in traffic accidents, not only the serious loss of personnel and property, but also a huge negative impact on society. Enterprises need to consider the cost of transportation and, more importantly, to ensure the safety of the transportation process, which has brought great challenges. In order to reduce transportation costs, companies will choose shorter transportation routes, and these

transportation routes will pass through areas with higher population density, leading to increased potential risks in the transportation process, while choosing a farther transportation route will undoubtedly increase total costs. As reported in this research, the problem of hazardous materials transportation includes two basic objective, minimum cost and risk control [1]. The key to solving such problems is to make a balance between the two points.

According to the regulations of the Ministry of Transport of the People's Republic of China, different types of hazardous materials transportation vehicles have different restrictions during actual operation. For some flammable and explosive products, according to the regulations, transport on specific roads is not allowed without permission. However, limited quantities of dangerous goods with a total mass not exceeding 8000 kg can be transported as ordinary goods. In addition, the hazardous materials manufacturers have the characteristics of limited production and unsuitable long-term storage in the actual situation. During the transportation process, a single visit to a customer cannot take advantage of the vehicle transportation, so as to control the potential risks. It is reasonable for the carrier to complete

the transportation task by traversing multiple manufacturers of similar products. This is also the dilemma currently faced by hazardous materials transportation companies. In this paper, we propose a multidepot heterogeneous vehicle routing model in order to find the optimal scheme to complete the transportation task among different manufacturers.

The traditional vehicle routing problem (VRP) aims generally to make the transportation cost the lowest under certain restrictions. It is based on general cargo transportation and does not consider the vehicle type, which means that all the vehicles can carry different types of goods. However, in heterogeneous vehicles routing problem, different types of goods must be transported by the specified vehicle type. In hazardous materials transportation, gasoline requires tanker trucks while solid hazardous materials should be carried by special lorry, and they are not interchangeable. In the real business, transportation always takes place between multiple depots and customer points. In addition, risk needs to be considered in the road transportation of hazardous materials. The treatment of risk in this article aims to convert the hazard after the accident into a risk coefficient and introduce uncertain scenarios to predict the probability of risk. The product of the probability of risk and the risk coefficient is the risk cost, which is used to add weighted transportation costs to obtain the total objective. The risks include road accidents, leakage of dangerous goods, and pedestrians on the way. Comprehensive consideration of factors such as transportation costs, multiple depots, heterogeneous vehicles, risks, and multiple accident scenarios can make the study closer to the actual situation. However, in contrast, the more factors we considered, the more difficult it is to solve the problem. Most of the research involves only two or three factors, and few involve more than three factors, which also makes our research valuable.

We propose a more complicated mixed integer programming (MIP) model which is close to the actual situation. Existing solvers, such as CPLEX, can only solve small-scale calculation examples. As the scale of the problem becomes larger, the efficiency of CPLEX's solution continues to decrease. Therefore, we design a bidirectional tuning heuristic and a particle swarm optimization algorithm (PSO) to solve problems at different scales. The results show that the algorithms we designed can effectively solve larger-scale problems and can also find near-accurate solutions to small-scale problems.

The reminder of this paper is organized as follows. Section 2 reviews related literature. Section 3 formulates a mathematical model. In Section 4, two different algorithms are proposed. Section 5 reports the numerical experiments. Conclusions are summarized in the last section.

2. Related Literature

The difference between vehicle routing problem for hazardous materials (VRPHM) and traditional VRP is that VRPHM needs to consider the risk and it is a key issue in research which cannot be ignored. Bula et al. [2] pointed out that the research on VRPHM mainly focuses on the shortest path problem and the routing problem. The former research

focuses on controlling risks, and such research has already achieved many results. A population exposure model was proposed by [3], which defines the range of the affected population on each side, and the risk value on each side is summed up to obtain the total risk value on the entire path. Tarantilis and Kiranoudis [4] had also done a similar topic research; the difference was that this research is based on vehicle fleet to meet customer needs, rather than focusing on the shortest path, which is more realistic. Bula et al. [5] defined the risk value of hazardous materials during transportation. The risk value is composed of the probability of accident, the extent of the accident, and the population density of the route. Based on this, a MIP model with the minimum risk value as the objective was established. Cuneo et al. [6] focused on the transportation of fuel oil; they proposed an innovative risk index in the function as far as possible to reduce the risk of accidents during the transportation of fuel oil. The method of most related work is to calculate fixed risk values few papers take stochastic into consideration. However, in other fields, the stochastic is widely used to access uncertainty. Rabbani et al. [7] studied hazardous waste management and controlled the processing costs of industrial wastes by introducing scenarios to defined risks. The risk defined in this paper is much like the method used in [8].

Another direction of research is to solve the VRPHM from the perspective of traditional VRP. Pradhananga et al. [9] considered the multiobjective problem of transportation time and risk. Risk in their paper was obtained by multiplying the incidence rate of each arc accident by the population density of the region. Du et al. [10] constructed fuzzy bilevel programming, which is a multidepot problem, and proposed a fuzzy simulation-based heuristic algorithm to solve the problem.

To the best of our knowledge, VRPHM with heterogeneous vehicle and uncertainty risks is rarely studied. Homogeneous models are easy to study, but the transportation of hazardous materials usually involves the problem of heterogeneous vehicle. Golden et al. [11] considered the problem of distribution of mixed vehicles with unlimited number, but the article only focused on the capacity of the vehicles. In the general VRP problem, there are many studies on heterogeneous vehicles. A number of effective algorithms were proposed to solve this problem; [12–14] used exact algorithms while [15, 16] designed heuristic algorithms. Due to the fact that VRPHM needs to consider more factors than general VRP, few scholars have studied heterogeneous vehicles.

In sum, the contributions of this paper are as follows. First, we propose a complex model to solve multidepot VRPHM, considering risk, heterogeneous vehicles, transportation cost, and time limit. Second, we describe the risk as uncertainty in different scenarios. Third, two effective algorithms are developed.

3. Problem Definition and Formulation

In this section, we describe the multidepot VRPHM, define the parameters, and then formulate a MIP model.

3.1. Problem Definition. The objective of this problem is to minimize the transportation costs, the risk costs resulting from the transformation of transportation risks, and the penalty cost of additional working time. We also consider various sudden scenes during transportation and consider the distribution of hazardous materials in multiple depots and heterogeneous vehicles. Different vehicles have different capacities and distribution categories. The vehicle cost in this paper consists of fixed cost and variable cost. Fixed cost refers to the cost unrelated to transportation which includes depreciation cost and insurance cost, while variable cost is related to the transportation distance and chosen route, which includes transportation cost and risk cost. The fixed costs of starting a vehicle and the variable costs per unit of travel distance are also different. The transportation network of hazardous materials can be defined on a directed network $G = (O, A)$, where the set O represents the union of the set D composed of depots and the set N composed of customer points, and $A = \{(i, j) \mid i, j \in O\}$ is an edge set. The network is depicted clearly in Figure 1.

Let the set K represent the vehicles of fixed cost g_k . The maximum load of each vehicle is q_k . The vehicle determines whether it can visit the customer based on the type of hazardous materials, described using parameter $c_{i,k}$. Each vehicle must depart from its own depot 0 to the customer $i \in N$. After process time t_i , the vehicle can travel from customer $i \in N$ to customer $j \in N$. We introduce a virtual node $N + 1$ to represent the depot that the vehicle returns to after the transportation. This virtual node is the same node as 0. The transportation time is different according to the chosen route which can be expressed as $d_{i,j,r,k}$, where the

route set is $R_{i,j}$. Besides, the travel time is different from 0 to each customer if the vehicle belongs to different depots. Each vehicle k has the latest working time y_k , and an additional overtime costs c_{time} is needed if the total transportation time is beyond the latest working time. Accidents are very likely to occur in the transportation of hazardous materials. This paper introduces the concept of scenarios with different probability to eliminate such uncertainties. The set of scenarios is represented by S . The risk under different scenarios is related to the route chosen by the vehicle. We use $w_{i,j,r,s}$ to describe the probability of accident that may occur from the customer $i \in N$ to the customer $j \in N$.

Before formulating the model, some assumptions are listed as follows:

- (1) All vehicles must return to the depots where they came from after transportation
- (2) Each vehicle can only transport one type of hazardous materials
- (3) Each customer can be accessed multiple times by different vehicles
- (4) Each vehicle serves one customer no more than once
- (5) Each vehicle must wait for the preceding vehicle to complete the operation before starting service

3.2. Notations. In this section, we define all the parameters and decision variables (Table1).

3.3. Mathematical Model.

$$\text{Minimize } \sum_{k \in K} g_k \eta_k + \frac{1}{|S|} \left[\sum_{s \in S} \sum_{k \in K} \sum_{j \in N \cup \{e(N)\}} \sum_{i \in N \cup \{0\}} \sum_{r \in R_{i,j}} c_{\text{tran}} d_{i,j,r,k} \gamma_{i,j,k,r,s} + \sum_{s \in S} \sum_{k \in K} \sum_{j \in N \cup \{e(N)\}} \sum_{i \in N \cup \{0\}} \sum_{r \in R_{i,j}} c_{\text{risk}} w_{i,j,r,s} \gamma_{i,j,k,r,s} \right. \\ \left. + \sum_{s \in S} \sum_{k \in K} c_{\text{time}} (\beta_{e,k,s} - y_k)^+ \right], \quad (1)$$

$$\text{subject to } \sum_{k \in K} \alpha_{i,k} \geq 1, \quad \forall i \in N, \quad (2)$$

$$\alpha_{i,k} \leq c_{i,k}, \quad \forall i \in N, k \in K, \quad (3)$$

$$\alpha_{i,k} \leq \eta_k, \quad \forall i \in N, k \in K, \quad (4)$$

$$\sum_{i \in N \cup \{0\}} \sum_{r \in R_{i,j}} \gamma_{i,j,k,r,s} = \sum_{i \in N \cup \{e(N)\}} \sum_{r \in R_{i,j}} \gamma_{j,i,k,r,s} = \alpha_{j,k}, \quad \forall j \in N, k \in K, s \in S, \quad (5)$$

$$\sum_{j \in N \cup \{e(N)\}} \sum_{r \in R_{0,j}} \gamma_{0,j,k,r,s} = \sum_{j \in N \cup \{0\}} \sum_{r \in R_{0,j}} \gamma_{j,e,k,r,s} = 1, \quad \forall k \in K, s \in S, \quad (6)$$

$$\sum_{i \in N} \sum_{j \in N} \sum_{r \in R_{i,j}} \gamma_{i,j,k,r,s} \leq |N_1| - 1, \quad N_1 \subseteq N, 2 \leq |N_1| \leq |N|, \forall k \in K, r \in R_{i,j}, s \in S, \quad (7)$$

$$\gamma_{i,j,k,r,s} \leq b_{i,j,r,s}, \quad \forall i \in N \cup \{0\}, j \in N \cup \{e(N)\}, k \in K, r \in R_{i,j}, s \in S, \quad (8)$$

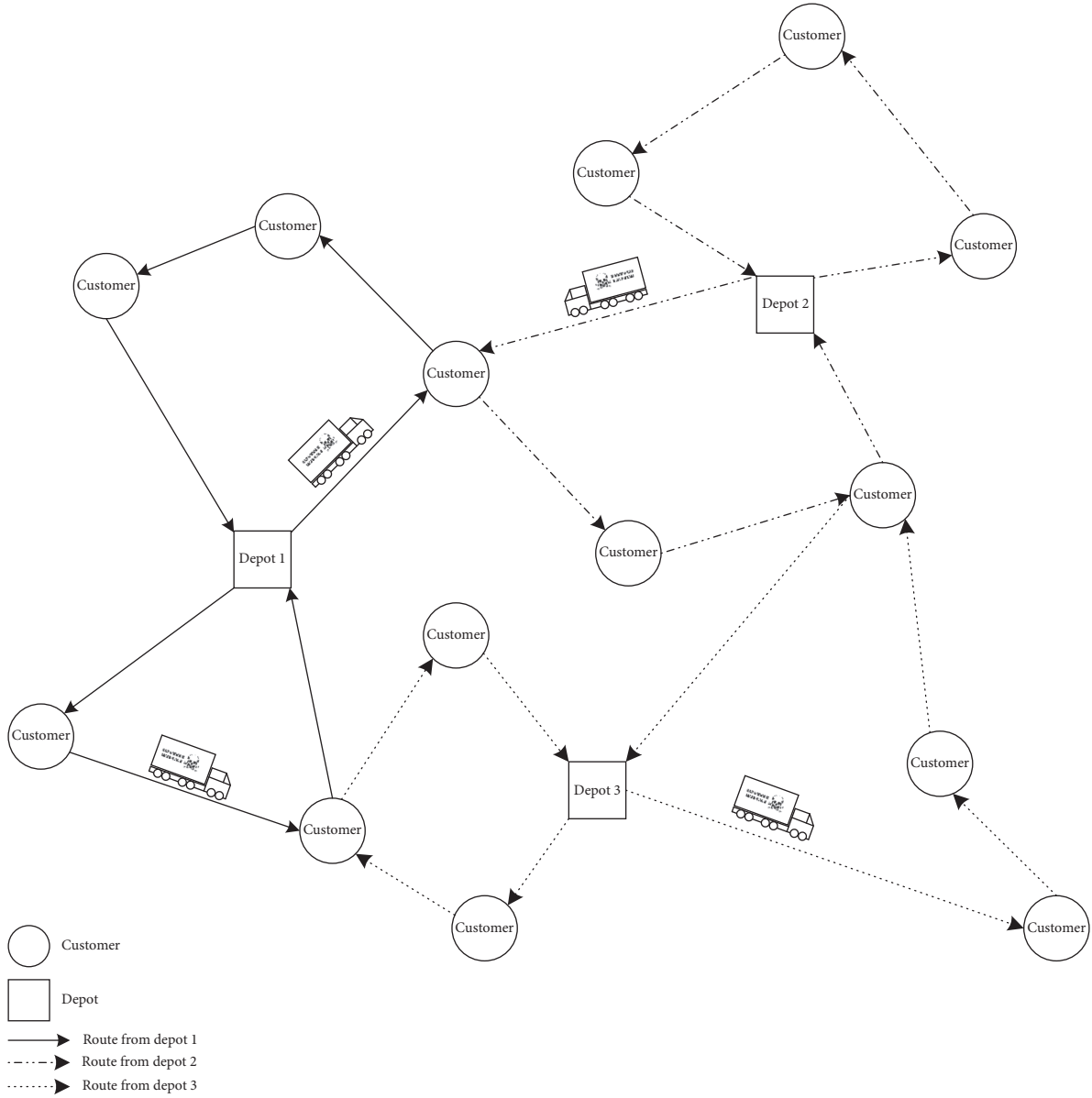


FIGURE 1: A vehicle routing network with three depots and several customers.

$$\sum_{i \in N} \zeta_{i,k,s,v} \leq q_k, \quad \forall k \in K, s \in S, \quad (9)$$

$$\sum_{k \in K} \zeta_{i,k,s,v} = f_{i,v}, \quad \forall i \in N, s \in S, v \in V, \quad (10)$$

$$\zeta_{i,k,s,v} \leq f_{i,v} \mu_{k,v} \alpha_{i,k}, \quad \forall i \in N, k \in K, v \in V, s \in S, \quad (11)$$

$$\sum_{v \in V} \mu_{k,v} = \eta_k, \quad \forall k \in K, \quad (12)$$

$$\delta_{i,k,s} + t_i + d_{i,j,r,k} \leq \beta_{j,k,s} + M(1 - \gamma_{i,j,k,r,s}), \quad \forall i \in N \cup \{0\}, j \in N \cup \{e(N)\}, k \in K, r \in R_{i,j}, s \in S, \quad (13)$$

$$\delta_{i,k,s} \geq \beta_{i,k,s}, \quad \forall i \in N, k \in K, s \in S, \quad (14)$$

TABLE 1: Notations.

Indices and sets	
i, j	Index of customer.
N	Set of customers.
k	Index of vehicle.
K	Set of vehicles.
s	Index of scenario.
S	Set of scenarios.
r	Index of route.
$R_{i,j}$	Set of routes from customer i to customer j .
v	Index of type of hazardous materials.
V	Set of type of hazardous materials.
Parameters	
$c_{i,k}$	Equals 1 if vehicle k can visit customer i .
$d_{i,j,r,k}$	Travel time from customer i to customer j for vehicle k on route r .
t_i	Service time at customer i .
$f_{i,v}$	Transportation demand of hazardous materials v at customer i .
q_k	Capacity of vehicle k .
y_k	Latest working time of vehicle k .
g_k	Fixed cost of vehicle k .
$w_{i,j,r,s}$	Risk probability from customer i to customer j on route r .
c_{tran}	Coefficient of transportation cost.
c_{risk}	Coefficient of risk cost.
c_{time}	Coefficient of work overtime cost.
Decision variables	
$\alpha_{i,k}$	Binary, equals 1 if vehicle k serves i .
$\gamma_{i,j,k,r,s}$	Binary, equals 1 if vehicle k travels from customer i to customer j by route r on scenario s .
$\varepsilon_{i,k,k',s}$	Binary, equals 1 if vehicle k' serves customer i earlier than vehicle k .
η_k	Binary, equals 1 if vehicle k is started.
$\mu_{k,v}$	Binary, equals 1 if vehicle k transports hazardous material v .
$\beta_{i,k,s}$	Float, time point when vehicle k arrives at customer i on scenario s .
$\delta_{i,k,s}$	Float, time point when vehicle k starts serving customer i on scenario s .
$\zeta_{i,k,s,v}$	Integer, quantity of hazardous material v transported by vehicle k at customer i on scenario s .

$$\delta_{i,k,s} + t_i \leq \delta_{i,k',s} + M(1 - \varepsilon_{i,k,k',s}), \quad \forall i \in N, k \in K, s \in S, \quad (15)$$

$$\sum_{k \in K \cup \{0(K)\}} \varepsilon_{i,k,e(K),s} = \sum_{k \in K \cup e(K)} \varepsilon_{i,0(K),k,s} = 1, \quad \forall i \in N, k \in K, k \in K, \quad (16)$$

$$\sum_{k \in K \cup \{0(K)\}} \varepsilon_{i,k,k',s} = \sum_{k \in K \cup e(K)} \varepsilon_{i,k',k,s} = 1, \quad \forall i \in N, k' \in K, s \in S, \quad (17)$$

$$\sum_{k \in K} \sum_{k' \in K} \varepsilon_{i,k,k',s} \leq |K_1| - 1, \quad K_1 \subseteq K, 2 \leq |K_1| \leq |K|, \forall i \in N, s \in S, \quad (18)$$

$$\alpha_{i,k}, \gamma_{i,j,k,r,s}, \varepsilon_{i,k,k',s}, \eta_k \in \{0, 1\}, \quad \forall i \in N \cup \{0\}, j \in N \cup \{e\}, k \in K, r \in R_{i,j}, s \in S, \quad (19)$$

$$\beta_{i,k,s}, \varepsilon_{i,k,k',s}, \zeta_{i,k,s} \geq 0, \quad \forall i, j \in N, k, k' \in K, s \in S. \quad (20)$$

The objective function (1) minimizes the total costs of vehicle startup, additional overtime costs, and risk conversion costs. Constraints (2) ensure that a customer is visited at least once. Constraints (3) limit whether vehicle k can serve customer i . Constraints (4) show that if vehicle k serves customer i , the vehicle must be started. Constraints (5) make sure that if a customer is visited by a vehicle, other nodes must be visited before and after this node. Constraints

(6) mean that, in any scenario, each vehicle must travel from the original depot and return to the same depot. Constraints (7) eliminate subpaths. Constraints (8) impose the condition that whether the vehicles are allowed to choose a route r from customer i to customer j under scenario s is determined by whether the route is passable. For example, some routes may be restricted during the morning and evening rush hours, specific time period, and temporary event.

Constraints (9) guarantee that the total loads of a vehicle will not exceed the maximum vehicle capacity. Constraints (10) and (11) ensure that all the demands of hazardous materials must be satisfied. Constraints (12) put the limit that each vehicle can only transport one type of hazardous material. Constraints (13) and (14) describe the time relationship between the time point of arrival and time point of service. Constraints (15) determine the service time of different vehicles at the same customer according to their sequence. Constraints (16) and (17) address the sequence of vehicle access for every customer. Constraints (18) ensure that there are no subpaths in vehicle loop. Constraints (19) and (20) limit the range of decision variables.

3.4. Linearizing the Products of Two Variables in the Constraint. In constraint (11), the load quantity contains the form of products of variables $\mu_{k,v}$ and $\alpha_{i,k}$. The nonlinear part is not conducive for solver to calculate. To linearize the constraint, some new variables are added as follows. Constraints (11) are then transformed into constraints (21)–(24).

A newly defined variable is as follows:

$\theta_{i,k,v}$: binary, equals 1 if $\alpha_{i,k} = 1$ and $\mu_{k,v} = 1$.

Newly defined constraints are as follows:

$$\theta_{i,k,v} \leq \mu_{k,v}, \quad \forall i \in N, k \in K, v \in V, \quad (21)$$

$$\theta_{i,k,v} \leq \alpha_{i,k}, \quad \forall i \in N, k \in K, v \in V, \quad (22)$$

$$\theta_{i,k,v} \geq \alpha_{i,k} + \mu_{k,v} - 1, \quad \forall i \in N, k \in K, v \in V, \quad (23)$$

$$\zeta_{i,k,s,v} \leq f_{i,v} \theta_{i,k,v}, \quad \forall i \in N, k \in K, v \in V, s \in S. \quad (24)$$

4. Algorithm Strategies

The model presented in Section 3 can be solved by solvers (e.g., CPLEX) directly in small-scale examples. To deal with large-scale problems, we develop two different heuristics, bidirectional tuning heuristic and particle swarm optimization algorithm (PSO), to solve our proposed model. In Sections 4.1 and 4.2, we describe the framework of the two heuristics, respectively.

4.1. Bidirectional Tuning Heuristic. The main idea of bidirectional tuning heuristic is to transform the model into several interrelated subproblems. Solving the subproblems using some fixed decision variables, we can obtain the remaining decision variables and then use them to gain other decision variables. In the iterative process, if the value of the objective is better than the optimal value, the optimal solution needs to be updated. The exit mechanism of the algorithm can be determined by the number of iterations or by the rules to judge whether the objective value can get a better solution.

Before we start solving, the decision variables need to be classified. The decision variables in this paper can be divided into three different types according to their definitions. The first type is $\alpha_{i,k}$, the variables used to decide whether a vehicle

visits a customer. The second type is $\mu_{k,v}$ and η_k , used to determine the type of hazardous material each vehicle transports. The other decision variables contain the dimensions of random scenarios and are limited by the first two kinds of decision variables, so they are always taken as the variables to be solved. We determine the customer that each vehicle visits and then optimize the transportation types with fixed $\alpha_{i,k}$. Then, we move to determining the transportation types and optimize the customers with fixed $\mu_{k,v}$ and η_k . The solving is repeated until the iteration reaches upper limit or no improvement can be obtained. The basic procedure can be defined as follows and the detailed process is shown in Algorithm 1.

Step 1: solve the model to find a feasible solution, which can be taken as the initial solution. Record the initial decision variables and the objective value.

Step 2: based on $\alpha_{i,k}$ from initialization, determine decision variable η_k and $\mu_{k,v}$. Compare the obtained objective value with the historical optimal value, and if the value is better than the historical optimal value, update the optimal value.

Step 3: solve the model with fixed η_k and $\mu_{k,v}$ to obtain $\alpha_{i,k}$. Update the objective value.

Step 4: judge the exit condition; if it is satisfied, the optimal solution is obtained and the algorithm stops; otherwise, go to step 2.

4.2. Particle Swarm Optimization (PSO). Particle swarm optimization (PSO) has been studied extensively since it was proposed [17]. As one of the swarm intelligence algorithms, PSO has been widely used in vehicle routing problem in the past few years, for example, vehicle routing problem with multiple pickup and delivery [18] and multidepot multitrip vehicle routing problem [19].

4.2.1. Initialization and Velocity Updating Strategy. The particle swarm optimization algorithm not only considers the personal optimal value, but also records the global optimal value of the entire group in the process of searching for solutions. It has excellent performance in solving optimization problems. The movement of particles depends on the update of their position and velocity; each particle contains the current position and the personal optimal position. The current position of the particles is adjusted by updating the velocity formula, the solution is judged by the specific fitness value, and the optimal position of the entire group is updated in the global range.

In the model we proposed, the decision variables $\alpha_{i,k}$ and $\mu_{k,v}$ are obviously related to other decision variables. When $\alpha_{i,k}$ is fixed, η_k can be determined by constraints (4) and $\gamma_{i,j,k,r,s}$ can also be determined by constraints (5)–(7). Furthermore, the decision variables $\beta_{i,k,s}$, related to time, are limited by $\gamma_{i,j,k,r,s}$ according to constraints (13). Then the variable $\delta_{i,k,s}$ can be derived from constraints (13). At this point, we can judge whether the service time $\delta_{i,k,s}$ of each customer is reasonable. If it is not reasonable, we update the

- (1) Initialization
- (2) Solve the model, and determine $\alpha_{i,k}$
- (3) Update the optimal solution
- (4) While ($n < \text{MaxIter}$)
- (5) Solve model with fixed $\alpha_{i,k}$, determine $\mu_{k,v}$ and η_k
- (6) Update the optimal solution
- (7) Solve model with fixed $\mu_{k,v}$ and η_k , determine $\alpha_{i,k}$
- (8) Update the optimal solution
- (9) End while

ALGORITHM 1: Bidirectional tuning heuristic.

time according to the time sequence of starting time and the processing time. It is also necessary to trace back to update variable $\beta_{i,k,s}$, which is the subsequent vehicle arrival time of the route. When all decision variables related to time are determined, we can easily confirm the decision variable $\varepsilon_{i,k,k',s}$. When we know both $\alpha_{i,k}$ and $\mu_{k,v}$, the load variable $\zeta_{i,k,s,v}$ can also be determined. We now have clarified the relationship between all the decision variables.

Based on the description above, we set $\mu_{k,v}$ as the outer variable of the algorithm. And the particles that we set up contain information to find the vehicles assigned to customer and to generate visit sequence of each vehicle, which is similar to the approach taken in this work [20]. We define the particle as P_m^n , and its velocity can be defined as Vel_m^n . The velocity update formula can be expressed by formula (25) and the position update formula is (26). We use $P\text{Best}_m^n$ to represent the best personal optimal position of the particle and $G\text{Best}_m^n$ to represent the global optimal position, where m means the number of particles (we take 300) and n means the current iteration number:

$$\text{Vel}_m^{n+1} = w^n \text{Vel}_m^n + c_1 r_1 (P\text{Best}_m^n - P_m^n) + c_2 r_2 (G\text{Best}_m^n - P_m^n), \quad (25)$$

$$P_m^{n+1} = P_m^n + \text{Vel}_m^n. \quad (26)$$

In the formula for particle velocity update, w^n denotes inertia weight and it is calculated by equation $w^n = w_{\max} - (w_{\max} - w_{\min})n/N$, in which N is the number of total iterations (we take 50) and w_{\max} and w_{\min} are maximum and minimum inertia weight (we take 0.9 and 0.4). This makes the particle swarm have strong global convergence ability at the beginning, but strengthens local convergence ability with the increase of iteration. c_1 and c_2 are acceleration weights (we take 0.683 for both), while r_1 and r_2 are random decimals between zero and one.

4.2.2. Procedure of PSO. According to the initialization and velocity updating strategy described above, the basic procedure of the PSO can be depicted as follows and the detailed process is shown in Algorithm 2:

Step 1: initialize particle swarm related parameters, including swarm size, position information contained in each particle, and particle velocity.

Step 2: normalize the information contained in the particles to obtain the vehicle assignment and sequence of customers which can further determine the visit time of each vehicle.

Step 3: judge whether the solution generated by each particle meets the capacity constraint. If not, the fitness value of the particle is the maximum value, and if it meets the constraint, the fitness value of the particle can be calculated.

Step 4: compare the fitness value with the personal optimal value of the particle. If it is better than the historical personal optimal value, update the personal optimal value and record the position information of the particle.

Step 5: update the particle position information and particle velocity according to the formula.

Step 6: determine whether the exit condition is satisfied (the gap is small enough or the maximum number of iterations is reached). If the condition is met, exit the loop; otherwise return to step 2.

4.2.3. Coding Rule of PSO. According to line 3 of the particle swarm flow above, parameter transformation refers to the transformation of the position information contained in the particles into the variables needed to solve the model, that is, the vehicle assignment and customer sequence. Figure 2 shows the process of assigning vehicles to customers by taking 6 customers of hazardous materials type 1 as example. Figure 2(a) shows the subset of customers for hazardous materials type 1 that need transportation services and the corresponding particle positions. Figure 2(b) shows a subset of vehicles capable of transporting type 1 hazardous materials. Figure 2(c) shows the vehicle that each customer is assigned to based on the location information contained in its particle. If all customers are assigned to one single vehicle and customer demand exceeds the capacity of the vehicle, we then need to adjust the vehicle allocation strategy. We add the requirements in ascending order according to the particle locations of the customers and assign the customers exceeding the capacity to the vehicles with smaller serial numbers in turn until the requirements are fully met.

Figure 3 depicts the generation of vehicle visiting sequence, and the number of customers is 5. Figure 3(a) represents the location information of each particle, and the

```

(1) Initialize swarm
(2) For each particle
(3)   Parameter conversion (generate customer sequence, vehicle assignment)
(4)   Check the constraints
(5)   Fitness evaluation
(6)   Update best personal solution and global solution
(7) End for
(8) While termination condition not met
(9)   Update particle position and velocity
(10)  Goto line 2
(11) End while

```

ALGORITHM 2: Particle swarm optimization (PSO).

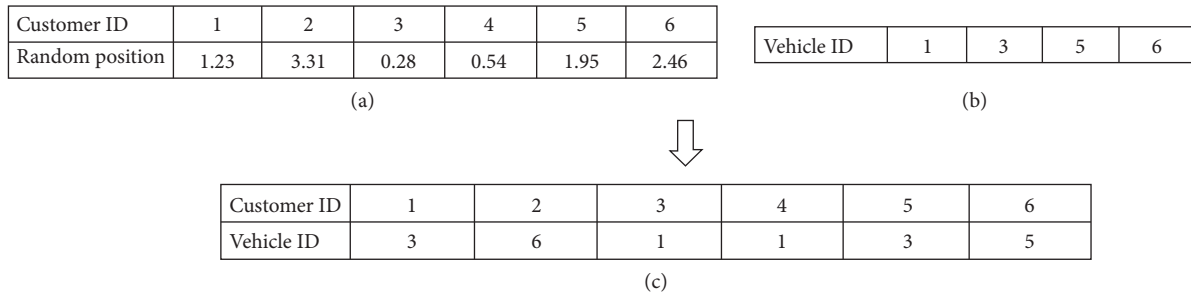
Cargo type $v = 1$:

FIGURE 2: Vehicle assignment mechanism. (a) PSO particle. (b) Vehicle subset. (c) Vehicle ID assigned to each customer.

Vehicle 1:

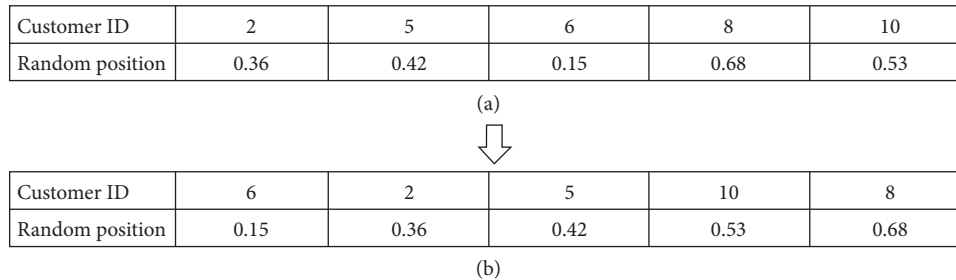


FIGURE 3: Generation of customer sequence. (a) PSO particle. (b) Customer sequence of each vehicle.

customer sequence obtained after sorting it in ascending order is depicted in Figure 3(b). Each iteration of PSO brings about a change in particle position, and a new sequence will be generated after sorting. Better paths can be found by constantly adjusting the sequence.

5. Numerical Experiments

In this part, we use extensive numerical experiments to verify whether the heuristic proposed in this article is efficient as well as the effectiveness of the model. All the experiments were conducted on a workstation with two Xeon E5-2643 CPUs (24 cores) of 3.4 GHz and 128 G of memory. All programs were compiled with C#, and the mixed integer programming model was solved by CPLEX.

5.1. Parameter Setting. We have designed multiple groups of randomly generated numerical examples for both small-scale and large-scale problems. In the small-scale example, we calculated the situation of 5, 10, and 15 customers with 2 or 3 types of hazardous materials, respectively, while in larger scale, we calculated the problem with customer number from 20 to 70 of three different hazardous materials. Since a customer could have multiple transportation demands with heterogeneous vehicles, each customer can be viewed as multiple customers, which makes solving on a larger scale more difficult. The remaining parameters were generated under some specific rules; for example, each customer had at least 1 type of hazardous materials transportation demand and the quantity was limited to a certain range. Each vehicle had a fixed latest working time. The three

TABLE 2: Comparison of two heuristics and CPLEX in small-scale problems.

Cases ID	CPLEX		Bituning		Comparison		PSO		Comparison	
	Z_C	T_C	Z_B	T_B	T_B/T_C	GAP_B (%)	Z_P	T_P	T_P/T_C	GAP_P (%)
A5-3-2-1	10237	0.2	10237	0.2	1.0	0.00	10265	9.1	45.5	0.27
A5-3-2-2	10198	0.2	10198	0.2	1.0	0.00	10233	4.3	21.5	0.34
A5-3-2-3	9421	0.2	9421	0.1	0.5	0.00	9461	4.2	21.0	0.42
A5-3-3-1	14739	2.0	14739	1.3	0.7	0.00	14921	7.3	3.7	1.23
A5-3-3-2	14249	1.4	14249	0.2	0.1	0.00	14287	24.8	17.7	0.27
A5-3-3-3	13789	1.7	13789	0.6	0.4	0.00	13946	57.3	33.7	1.14
A10-3-2-1	12041	144.1	12041	6.3	0.0	0.00	12248	24.9	0.2	1.72
A10-3-2-2	12217	335.0	12217	1.1	0.0	0.00	12457	31.3	0.1	1.96
A10-3-2-3	12527	606.3	12527	4.3	0.0	0.00	12636	29.7	0.0	0.87
A10-3-3-1	15570	358.14	15570	22.2	0.1	0.00	15703	42.7	0.1	0.85
A10-3-3-2	16655	602.9	16692	63.8	0.1	0.22	16889	32.7	0.1	1.40
A10-3-3-3	19774	610.3	19774	98.9	0.2	0.00	20003	39.1	0.1	1.16
A15-6-2-1	23551	1576.2	23551	215.2	0.1	0.00	24079	61.1	0.0	2.24
A15-6-2-2	21115	1352.1	21115	144.5	0.1	0.00	21577	60.2	0.0	2.19
A15-6-2-3	24790	2161.1	24790	218.7	0.1	0.00	25169	70.4	0.0	1.53
Avg.	—	516.8	—	51.8	0.3	0.01	—	33.3	9.6	1.17

TABLE 3: Comparison of two heuristics and CPLEX in large-scale problems.

Case ID	CPLEX (7200)	Bituning		PSO			Comparison	
	Z_C	Z_B	T_B	Z_P	T_P	T_P/T_B	GAP_B (%)	GAP_P (%)
A20-6-3-1	36507	35915	809.1	36404	124.4	0.15	1.65	0.28
A20-6-3-2	35412	34148	815.3	34972	119.7	0.15	3.70	1.26
A30-10-3-1	49132	—	>7200	46878	341.3	—	—	4.81
A30-10-3-2	53486	—	>7200	50540	438.5	—	—	5.83
A40-15-3-1	69194	—	>7200	63715	923.4	—	—	8.60
A40-15-3-2	71866	—	>7200	66598	732.2	—	—	7.91
A50-20-3-1	86210	—	>7200	78238	1229.2	—	—	10.19
A50-20-3-2	100626	—	>7200	88068	1521.2	—	—	14.26
A60-25-3-1	115044	—	>7200	99065	2671.9	—	—	16.13
A60-25-3-2	114854	—	>7200	97965	3016.5	—	—	17.24
A70-30-3-1	172731	—	>7200	134730	7407.3	—	—	28.21
A70-30-3-2	139333	—	>7200	113436	6175.8	—	—	22.83

cost coefficients are estimated under China's hazardous material transportation market. $c_{\text{tran}} = 200$ means the transportation cost in an hour which is calculated by speed, fuel consumption, and fuel cost ($60 \text{ km/h} \times 0.66 \text{ L/km} \times 5.05 \text{ yuan/L}$), $c_{\text{risk}} = 10000$ is a large number based on practical accident losses, and $c_{\text{time}} = 50$ represents additional work cost per hour.

5.2. Performance of Two Heuristics. The comparison between two heuristics and CPLEX is illustrated in Table 2, where Z means the objective value, T means computation time, and GAP is calculated by equation $(Z - Z_C)/Z_C$. The subscripts C, B, P represent CPLEX, bidirectional tuning heuristic, and PSO, respectively.

We tested fifteen cases in small scale, shown in Table 2. The case ID A5-3-2-1 means 5 customers, 3 vehicles, and 2 hazardous material types of instance 1. When the customer turns to 20, the proposed model cannot obtain a feasible solution in 7200 seconds. It is obvious that CPLEX can only maintain high calculation efficiency in extremely small-

scale examples which contain less than 10 customers. As the customer number and types of hazardous materials increase, the calculating time spent by CPLEX grows exponentially. Meanwhile, the bidirectional tuning heuristic can also obtain the optimal solution without increasing time significantly. In terms of small-scale problem, bidirectional tuning performed the best among the three methods with an average of 51.8 seconds of computation time and average gap of about 0.01%. The PSO spent the shortest time with an average of 33.3 seconds and average gap of about 1.17%, which could be considered as a near-optimal solution.

We list a group of large-scale instances of customer number from 20 to 70 to further verify the effectiveness of the algorithm. The result is shown in Table 3, where each case ID means the same as in Table 2. The GAP is calculated by equation $(Z_C - Z)/Z$. Due to the complexity of the model, when the customer number is over 30 with vehicle number of 10, the bidirectional tuning heuristic cannot find optimal solution in 7200 seconds. We find that even if it takes more time to search, the solution is

not improved. Therefore, we had to use the results computed by solving the problem after relaxation to judge the effectiveness of algorithm. The method proved to be feasible in other research [21]. However, even after relaxation, CPLEX still takes a long time to find a feasible solution. A long calculation time means no practical value, so it is acceptable that we set the computation time as 7200 seconds to get the solution within the effective time. From the data of Table 3, we can see an evident improvement of the PSO's gap value. Moreover, the improvement increases with the scale of problem. The computation time increases linearly as the problem size increases, which is much better than CPLEX and bidirectional tuning heuristic.

To summarize, in small-scale problem, bidirectional tuning heuristic can obtain the same optimal solution as CPLEX, while PSO can obtain near-accurate solution. On the contrary, PSO spent less time to find the solution, followed by the bidirectional tuning heuristic. In large-scale problem, PSO shows great capability with high-quality solution in a short time.

6. Conclusion

The heterogeneous vehicle routing problem for hazardous materials is a practical problem that deserves to be studied. It is difficult to balance the total cost under the premise of considering many influencing factors. In this paper, we study the multidepot VRPHM with risk analysis and propose mixed integer programming to transform the actual problem into a model, which can be solved by mathematical methods. The main contributions we made in this research are listed as follows: First, we consider comprehensive factors including transportation costs, multiple depots, heterogeneous vehicles, risks, and multiple accident scenarios which make the study closer to the actual situation. Furthermore, the risk is defined as uncertainty in different scenarios, and we consider heterogeneous vehicles which are rarely studied. For solving the problem in an efficient way, we design a bidirectional tuning heuristic and particle swarm optimization (PSO) to be applied to different scales of problem.

The numerical experiments show that our proposed algorithm can be used in small-scale problem with faster speed than CPLEX. And in large-scale problem (70 customers, 30 vehicles, and 3 types), PSO can find high-quality feasible solution in acceptable time.

However, this model can still be improved in terms of, for example, the definition of risk with uncertainty. The data we tested in numerical experiments are randomly generated in a reasonable range, which may be a little difference from the actual situation. And more algorithms can be tried to solve this complex problem.

Data Availability

The raw/processed data required to reproduce the findings cannot be shared at this time as the data also form a part of an ongoing study.

Conflicts of Interest

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

Acknowledgments

This study was supported by Beijing Nova Program of Science and Technology (no. Z191100001119029) and Research on Multi-Depot Heterogeneous Vehicle Routing Problem.

References

- [1] K. G. Zografos and K. N. Androutsopoulos, "A decision support system for integrated hazardous materials routing and emergency response decisions," *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 6, pp. 684–703, 2008.
- [2] G. A. Bula, C. Prodhon, F. A. Gonzalez, H. M. Afsar, and N. Velasco, "Variable neighborhood search to solve the vehicle routing problem for hazardous materials transportation," *Journal of Hazardous Materials*, vol. 324, pp. 472–480, 2017.
- [3] C. ReVelle, J. Cohon, and D. Shobry, "Simultaneous siting and routing in the disposal of hazardous wastes," *Transportation Science*, vol. 25, no. 2, pp. 138–145, 1991.
- [4] C. D. Tarantilis and C. T. Kiranoudis, "Using the vehicle routing problem for the transportation of hazardous materials," *Operational Research*, vol. 1, no. 1, pp. 67–78, 2001.
- [5] G. A. Bula, F. A. Gonzalez, C. Prodhon, H. M. Afsar, and N. M. Velasco, "Mixed integer linear programming model for vehicle routing problem for hazardous materials transportation**universidad nacional de colombia. universite de technologie de troyes," *IFAC-PapersOnLine*, vol. 49, no. 12, pp. 538–543, 2016.
- [6] V. Cuneo, M. Nigro, S. Carrese, C. F. Ardito, and F. Corman, "Risk based, multi objective vehicle routing problem for hazardous materials: a test case in downstream fuel logistics," *Transportation Research Procedia*, vol. 30, pp. 43–52, 2018.
- [7] M. Rabbani, R. Heidari, and R. Yazdanparast, "A stochastic multi-period industrial hazardous waste location-routing problem: integrating NSGA-II and Monte Carlo simulation," *European Journal of Operational Research*, vol. 272, no. 3, pp. 945–961, 2019.
- [8] L. Zhen, "Tactical berth allocation under uncertainty," *European Journal of Operational Research*, vol. 247, no. 3, pp. 928–944, 2015.
- [9] R. Pradhananga, E. Taniguchi, T. Yamada, and A. G. Qureshi, "Bi-objective decision support system for routing and scheduling of hazardous materials," *Socio-Economic Planning Sciences*, vol. 48, no. 2, pp. 135–148, 2014.
- [10] J. Du, X. Li, L. Yu, R. Dan, and J. Zhou, "Multi-depot vehicle routing problem for hazardous materials transportation: a fuzzy bilevel programming," *Information Sciences*, vol. 399, pp. 201–218, 2017.
- [11] 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.
- [12] A. Pessoa, R. Sadykov, and E. Uchoa, "Enhanced Branch-Cut-and-Price algorithm for heterogeneous fleet vehicle routing problems," *European Journal of Operational Research*, vol. 270, no. 2, pp. 530–543, 2018.

- [13] W. Sun, Y. Yu, and J. Wang, "Heterogeneous vehicle pickup and delivery problems: formulation and exact solution," *Transportation Research Part E: Logistics and Transportation Review*, vol. 125, pp. 181–202, 2019.
- [14] Y. Yu, S. Wang, J. Wang, and M. Huang, "A branch-and-price algorithm for the heterogeneous fleet green vehicle routing problem with time windows," *Transportation Research Part B: Methodological*, vol. 122, pp. 511–527, 2019.
- [15] Y.-J. Kwon, Y.-J. Choi, and D.-H. Lee, "Heterogeneous fixed fleet vehicle routing considering carbon emission," *Transportation Research Part D: Transport and Environment*, vol. 23, pp. 81–89, 2013.
- [16] P. H. V. Penna, A. Subramanian, L. S. Ochi, T. Vidal, and C. Prins, "A hybrid heuristic for a broad class of vehicle routing problems with heterogeneous fleet," *Annals of Operations Research*, vol. 273, no. 1-2, pp. 5–74, 2019.
- [17] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proceedings of the Sixth International Symposium on Micro Machine and Human Science MHS '95*, Nagoya, Japan, October 1995.
- [18] 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.
- [19] L. Zhen, C. Ma, K. Wang, L. Xiao, and W. Zhang, "Multi-depot multi-trip vehicle routing problem with time windows and release dates," *Transportation Research Part E: Logistics and Transportation Review*, vol. 135, Article ID 101866, 2020.
- [20] L. Zhen, Z. Xu, K. Wang, and Y. Ding, "Multi-period yard template planning in container terminals," *Transportation Research Part B: Methodological*, vol. 93, pp. 700–719, 2016.
- [21] L. Zhen, "Modeling of yard congestion and optimization of yard template in container ports," *Transportation Research Part B: Methodological*, vol. 90, pp. 83–104, 2016.

Research Article

Model for Design of Portfolio Venture Investment Contract When Taking Moral Hazards into Account

L. Yin,¹ Y. Liu ,² and Z. Wang²

¹School of Financial Technology, Shanghai Lixin University of Accounting and Finance, Shanghai 201209, China

²School of Management, Wuhan University of Technology, Wuhan 430070, China

Correspondence should be addressed to Y. Liu; liuying@whut.edu.cn

Received 10 June 2020; Revised 28 June 2020; Accepted 3 July 2020; Published 25 August 2020

Academic Editor: Lu Zhen

Copyright © 2020 L. Yin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Portfolio investment is adopted by the venture capital to diversify those risks involved in project selection, investing or operating so that the venture capitalist can expect a relatively stable income and lower financing risks. Based on the design of portfolio investment contract with unlimited funds developed by Kanninen and Keuschnigg, and Inderst et al., this article makes a modification and presents a model given the limitation of funds available for the venture capitalist. It is demonstrated that the marginal benefit of efforts paid by the entrepreneurs exceeds the marginal cost, given the limitation of funds available, which will conduce to a high-level engagement of the entrepreneurs. Thus, by adopting the design of renegotiation contract, the venture capitalist can manage to stimulate the entrepreneurs to make efforts, which is to result in moral hazard reduction.

1. Introduction

Enhanced risks in the venture capital investments may vary due to influences of the internal and external environment. The high uncertainty makes it essential for the venture capital to adopt an appropriate portfolio investment. It is intended to spread and diversify risks and in turn to ensure a high average yield on the investments. This is especially true with the high-tech investments. Generally, the venture capital should be invested in different regions, industries, companies, or projects instead of investing in a single one. By this means, losses from the failed investments can be made up for by the yields from those succeeded. For the venture capital investments of high uncertainty, adoption of a portfolio is becoming the choice of most venture capitalist.

By diversifying the risks involved in the project selection, investing, or operating, portfolio investments ensure the venture capitalist a relatively stable income and lower financing risks. Usually, in addition to making up for the losses resulted from the failed projects accounting for up to 70% of the number of assets included in a portfolio, the yields generated by the other 30% succeeded can bring the venture capitalist a return high enough.

In the US, the venture capital is typically invested in a dozen and even dozens of projects. Generally, the venture capitalist is not the only controlling party. They usually account for 10%–30% of the invested firm's total shares. According to a survey made in 2002 covering 218 venture projects of the US, a total loss was reported for 14.7% of all the projects and partial loss for 24.8%. However, of all the 218 projects, up to 60% earned a return of over 100%. According to the statistics, over the preceding dozens of years, the venture capital investments have produced an average yield of approximately 25%, which is far more than other types of investments. Adoption of portfolio investment contributes much to this high yield within a stable context of contracts.

The venture investments are characterized as risky and high yielding. Instead of merely taking risks, the venture capitalist always tries his best to lower the potential risks. To achieve this, most times, a portfolio is set up. Success of a venture firm is the result of the efforts made by both capitalist and entrepreneur. Even for one single asset venture investment, some moral hazard can arise from the information asymmetry between the capitalist and entrepreneur and the resulted lack of efforts. Obviously, in the context of a

portfolio the moral hazard problem associated is just more complex. It is necessary to figure out an appropriate design of portfolio contract for the purpose of avoiding moral hazard risk related to portfolio investments.

This research is intended to contribute to the findings of what makes an appropriate design of portfolio contract. Either from a theoretic perspective or a practical one, the portfolio investment decision is more likely to be constrained by limited fund availability to the venture capitalist. Taking this fund availability constraint into account, our article develops a model for the design of the portfolio venture investment contract, by which means an insight can be got into the underlying mechanism and how it works on bilateral moral hazard reduction.

2. Literature Review

With a portfolio investment contract, the venture capitalist invests in many ventures simultaneously, which gives rise to the information asymmetry between the investor and each entrepreneur and the resulted moral hazard. Study on this type of moral hazard problem can trace back to that emerges between one principal and multiple agents. Holmstrom [1] conducted the earliest study on the moral hazard problem arising from the information asymmetry between one principal and multiple agents. He concluded that when the efforts made by each agent cannot be observed by the principal and subsequently each one's contribution to the outcome cannot be identified, the agents will not try their best. Instead, they may take advantage of information asymmetry and behave as a free rider.

When a venture investment contract involves two or more entrepreneurs, information about the efforts made by each is usually unavailable. It is hard if it is not impossible to identify the contribution made by each one to the outcome specifically. As a result, moral hazard problem and its negative results can be expected. Therefore, a good design of portfolio investment contract should conduce to avoiding or reducing moral hazard between the one principal and multiple agents.

Kaplan et al. [2, 3] explored how the lack of efforts occurs when the venture capitalist, as the principal, interacts with two entrepreneurs (the agents) given complete information or incomplete information. It is suggested that moral hazard may arise in case the agents conspire with each other to undermine the venture investor's interests, or otherwise, one of the agents conspires with the principal and damages the interests of the other.

Cestone [4] concluded that, for a portfolio investment comprising two venture firms, the value of the ventures is positively related to the abilities of both the venture capitalist and the entrepreneur in charge, as well as their efforts. Additionally, the author presented an optimal income allocation plan to motivate both sides and eliminate the bilateral moral hazard.

Other research studies on moral hazard associated with the portfolio contract include figuring out the optimal number of the venture projects invested (or in other words, deciding on an optimal scale for the portfolio in question) given a certain level of efforts to be paid by all stakeholders.

Kanniainen and Keuschnigg [5], from the perspective of the bilateral moral hazard problem emerging between the venture capitalist and entrepreneur, analyzed theoretically how the capitalist should determine an optimal number of the start-up firms included in a portfolio investment. They hold that there is a trade-off between the number of start-up firms and the amount of assistance the capitalist could grant to each firm. That is, the more investees the less assistance to each. Such a trade-off determines there exist an optimal scale for a portfolio venture investment. Furthermore, they gave an interpretation about why only limited number of venture projects can be included in a portfolio investment when aimed at an optimal yield level totally. It was also supposed that as the investees increase and the assistance granted to each venture firm decreases, the entrepreneur would require more return from his/her firm. Bengtsson and Sensoy [6] investigated how contract design is related to VC abilities to monitor and provide value-added services to the entrepreneur and found that previous estimates of the amount entrepreneurs pay for affiliation with high-quality VCs are overstated.

Khanna and Mathews [7] tried to figure out the number of start-up firms to be invested based on the extent to which the venture capitalist can provide each firm with consultancy services. They believed there exist a reasonable and balanced level of the consultancy services each venture firm can expect and of the portfolio scale. Having referred to the research studies by Kanniainen et al. [5, 8], Khanna and Mathews argued that it is not enough to determine a portfolio scale merely based on finding the equilibrium of portfolio scale and the upper limit for services each firm may expect. Instead, they argued that the optimal portfolio scale can be reasonably determined only when the abilities of the venture capitalist are taken into account and the intensity of consultancy services each firm can have is predictable. Other similar research studies include that by He [9], Hori and Osano [10], Lin [11], Fu et al. [12], and An and Schneider [13]. They conducted studies on the equilibrium of efforts allocation among the venture investees and portfolio scale from different perspectives and tried to summarize the mechanism underlying a portfolio investment contract motivating the entrepreneurs. Chernenko et al. [14] found that having to carefully manage their own liquidity pushes mutual funds to require stronger redemption rights, suggesting contractual choices consistent with mutual funds' short-term capital sources. Burchardt et al. [15] provided a comprehensive theoretical and empirical literature review of venture capital contracts and highlighted the major discrepancies between theory and practice.

In brief, it can be concluded from the above literature that when bounded by a portfolio investment contract, both the capitalist and the entrepreneurs have to make efforts. It is especially true with the former. For the venture capitalist, reasonable allocation of the limited consultancy services that he/she is able to offer among the investees is critical to yield maximization. With the value-added services the capitalist has to provide being taken into account, he/she has to balance the more yields brought by a larger size of portfolio against the limitation of his/her ability to deliver value-added services.

Overall, how much effort would both the capitalist and each of the entrepreneurs like to make is critical to development of the start-up firms, especially in the framework of a portfolio investment contract. Since portfolio investment involves multiple entrepreneurs, the venture capitalist is faced with a more complex situation and it is more important to decide on a reasonable allocation of efforts. Portfolio investment has become an increasingly important means to invest in the start-up firms. The design of portfolio contract needs to be analyzed so that an optimal level of efforts be ensured from mechanisms. Based on the previous research studies, this article develops a model to optimize the design of portfolio investment contract which is expected to reach an equilibrium of incentives for the venture capitalist and each of the entrepreneurs. With this design adopted, moral hazards can be reduced and both sides can be stimulated to make essential efforts.

Also, it can be seen that, on the basis of the work involving just one entrepreneur, some researchers analyzed the possible means by which the capitalist can urge more entrepreneurs to make efforts and reduce the bilateral moral hazards in terms of portfolio investment contract. They assumed that, with unlimited availability of funds, the capitalist has enough money to make a second invest in both start-up firms. Thus, no one will fail to get further invested. That means, without fund constraint, the two start-ups do not have to compete with each other. However, that is usually not the truth. The portfolio investment decision is actually more likely to be constrained by limited fund availability to the venture capitalist. So, this article makes a modification. It is supposed that the decision about a second investment is constrained by limited fund availability to the venture capitalist. In this case, the entrepreneurs involved in the portfolio contract have to compete with each other to obtain the second investment.

3. Model Introduction

Consider a case of investing in multiple venture firms simultaneously. Each firm has a project which needs management assistance from the capitalist. A four-stage model will be developed for the venture investment contract.

Suppose each start-up project is to be invested in twice. The first investment amounts to I_1 and the second I_2 . After the first investment, both venture capitalist and entrepreneurs can observe the immediate state of the project which is relied upon to make a decision on whether to make a second investment. The whole process can be divided into four stages, which is demonstrated in Figure 1.

The logics underlying this four-stage model is that, in reality, most venture investment contracts cannot provide that the fund needs be satisfied once and for all. What is more reasonable and more usual in venture investment practices is to have a shot, wait to see the result, and then decide. With this in mind, we make the assumption that the venture capitalist will invest an amount necessary to get through the beginning phase of the selected venture projects. Then, the entrepreneurs are engaged in operating their start-ups and achieve some result based on which the capitalist is

to make the decision about whether to continue with investment. Once a further-investment decision is made, a renegotiation will be settled and the reinvestment contract is decided on. After this second investment, any yields of the start-up project will be allocated between the capitalist and the entrepreneur according to what is specified in the contract. This process can be logically simplified as the four-phase model shown in Figure 1.

Stage 1: sign the initial investment contract. With no loss of generality, assume that one venture capitalist invests in two start-ups. The amounts initially invested are both I_1 . As a return, the capitalist owns shares of two start-ups. Start-up firm i has shares amounted to s_i , $i = 1, 2$.

Stage 2: after the initial investment, each of the two entrepreneurs is engaged in operating his/her start-up. The extent to which the entrepreneur would be engaged depends on the firm's state following the initial investment, θ . After the initial investment, the entrepreneurs take up their line of duties. Assume that an entrepreneur may choose to have a high or low level of engagement. This level of engagement is denoted as e , $e \in \{e_h, e_l\}$. Costs will be incurred under different levels of engagement, which are denoted as $C(e_h)$ and $C(e_l)$. It is specified that $C(e_h) > C(e_l)$. And the marginal cost of level of engagement is ΔC , $\Delta C = C(e_h) - C(e_l)$. Suppose that the extent to which the entrepreneur is engaged determines the probability that the firm falls in a certain state. The entrepreneur makes efforts, the firm can have one of the two states, $\{\theta_g, \theta_m\}$, and θ_g and θ_m represent the good state and the normal state, respectively. When the entrepreneur is highly engaged, the start-up has the good state (θ_g) with probability of q_h , or the normal state (θ_m) with $(1 - q_h)$. If the entrepreneur takes a low level of engagement, the start-up has the good state (θ_g) with probability of q_l , or the normal state (θ_m) with $(1 - q_l)$. It is believed that $1 > q_h > q_l > 0$.

Stage 3: based on the observed state of the start-up, θ , the venture capitalist renegotiates with each of the entrepreneurs to determine whether to make a second investment and how to reallocate the income between the two sides. The total yield of each venture project can be Π or 0. The probability of being Π or 0 is dependent on the efforts made by the entrepreneur as well as the reinvestment decision made by the capitalist. If the capitalist, based on the observed state (θ), chooses to quit, the start-up will achieve the yield of Π with a much smaller probability due to the absence of further investment and assistance from the capitalist. This low probability is denoted as P_0 . And the probability of having no yield is $(1 - P_0)$. Obviously, the expected income with no further investment from the capitalist can be computed as

$$R_0 = P_0 \cdot \Pi. \quad (1)$$

If the venture capitalist chooses to make a second investment, the start-up will achieve the yield of Π with a

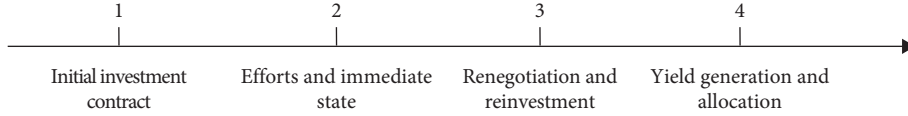


FIGURE 1: Sequence of the game involved in a portfolio investment contract.

larger probability. In the case that θ_g is observed with the project, the probability of achieving \prod is P_g and no yield ($1 - P_g$). In the case of θ_m , the probability of having \prod is P_m and no yield ($1 - P_m$). It is no surprise that the expected total yield with the state θ_g is more than that with θ_m , and the expected total yield with further investment is more than that with no more. So, it is stated that $P_g > P_m > P_0$. Apparently, when θ_g is observed with the start-up, the expected income can be calculated as

$$R_g = P_g \cdot \prod. \quad (2)$$

When θ_m is observed, the expected income is

$$R_m = P_m \cdot \prod. \quad (3)$$

Assume the additional investment is I_2 . Provided that the start-up receives further investment, the marginal income is r_g with the observed state of θ_g or r_m with θ_m . Then,

$$r_g = R_g - R_0 - I_2, \quad (4)$$

$$r_m = R_m - R_0 - I_2. \quad (5)$$

If $r_g > r_m > 0$, that is, there is a positive marginal income of further investment whether the observed state is θ_g or r_m . Therefore, further investment is always profitable.

The venture capitalist and each entrepreneur would renegotiate based on the observed firm state and the total yield expectation. By renegotiation of income allocation and redesign of investment contract, the capitalist manages to urge the entrepreneur to make efforts in Stage 2 so that the venture investment can be continued.

Stage 4: the expected yield having been earned, income is allocated between the capitalist and each entrepreneur. After the venture project has achieved the expected earnings, the capitalist and the entrepreneur will have earnings allocated according to the agreement reached in advance. It is assumed that the marginal income of efforts made by the entrepreneur exceeds the marginal cost. So, a rational entrepreneur would be highly engaged.

When the entrepreneur is indeed highly engaged, the start-up will have a state following the initial investment of θ_g with a probability of q_h , or θ_m with $(1 - q_h)$. Then, the expected marginal income of a second investment is

$$q_h \cdot r_g + (1 - q_h) \cdot r_m. \quad (6)$$

Or if the entrepreneur is not so highly engaged, the start-up will have a state following the initial investment of θ_g with a probability of q_l , or θ_m with $(1 - q_l)$. Then, the expected marginal income of further investment is

$$q_l \cdot r_g + (1 - q_l) \cdot r_m. \quad (7)$$

The entrepreneur will be urged to be highly engaged only when the marginal income of further investment with a high level of engagement is no less than that with a low level combined with the marginal cost of efforts made. That is,

$$q_h \cdot r_g + (1 - q_h) \cdot r_m \geq q_l \cdot r_g + (1 - q_l) \cdot r_m + \Delta C, \quad (8)$$

which can be reduced as

$$(q_h - q_l) \cdot (r_g - r_m) \geq \Delta C. \quad (9)$$

Equation (9) gives the constraints that ensure a high level of engagement by the entrepreneur.

3.1. Design of Portfolio Investment Contract Given the Limitation of Funds. Constrained by limited fund availability, the venture capitalist invests I_1 in each start-up initially, and after a certain period of time, there is only I_2 available for a second investment. That means, this capitalist manages to raise funds totaling $(2I_1 + I_2)$. Having invested I_1 in each of two start-ups, the capitalist has to choose the one more competitive to make further investment. Evidently, the observed state of each start-up determines which one will get a second investment. The one seeming more competitive wins.

As supposed earlier, a start-up can have the state of θ_g or θ_m after the initial investment. When the two investees have different states, one θ_g and the other θ_m , the venture capitalist would like to invest another I_2 in the one demonstrating θ_g . If both of the two have the state of θ_g , or both θ_m , the capitalist will invest I_2 in one of the two with equal probability.

With no loss of generality, assume Start-up 1 demonstrates θ_g after the initial investment and Start-up 2 θ_m . In this case, Start-up 1 will receive further investment and Start-up 2 will not have a chance. According to the renegotiation scheme of income allocation, Start-up 1 will have earnings of $(s_1 \cdot R_0 + \rho \cdot r_g)$, where ρ is the shares Start-up 1 owns based on the design of reinvestment contract. On the contrary, Start-up 2 will earn $s_2 \cdot R_0$.

In the case that the two start-ups both demonstrate the state of θ_g or otherwise θ_m , the capitalist will make further invest in any of them randomly. Then, Start-up 1 will have earnings of $(s_1 \cdot R_0 + (1/2)\rho \cdot r)$, and Start-up 2 will earn $(s_1 \cdot R_0 + (1/2)\rho \cdot r)$, where $r = \{r_g, r_m\}$.

From the perspective of Start-up 1, since Start-up 1 itself and Start-up 2 may decide to have a high or low level of engagement, the yield Start-up 1 can expect varies. It is analyzed as follows.

A Nash equilibrium exists, as demonstrated in Figure 2, when the two start-ups make their choices about how much

		Start-up 2	
		Highly engaged	Lowly engaged
Start-up 1	Highly engaged	(U_{1_a}, U_{2_a})	(U_{1_b}, U_{2_b})
	Lowly engaged	(U_{1_d}, U_{2_d})	(U_{1_c}, U_{2_c})

FIGURE 2: Expected income under different strategy combinations.

effort to make. U_{ij} ($i = \{1, 2\}$, $j = \{a, b, c, d\}$) is the yield that Start-up 1 can expect when the strategy combination adopted by entrepreneurs is j (a, b, c or d), as shown in Figure 2.

3.2. The Case That Entrepreneur 2 Chooses to Be Highly Engaged. In this case, when Entrepreneur 1 also chooses to be highly engaged, Start-up 1 may earn $(s_1 \cdot R_0 + \rho \cdot r_g)$ or $s_1 \cdot R_0$ with equal probability of $q_h \cdot (1 - q_h)$, or earn $(s_1 \cdot R_0 + (1/2)\rho \cdot r_g)$ with probability of $q_h \cdot q_h$, or $(s_1 \cdot R_0 + (1/2)\rho \cdot r_m)$ with $(1 - q_h) \cdot (1 - q_h)$. So, the expected yield of Start-up 1 when Entrepreneur 1 is highly engaged is

$$U_{1_a} = q_h(1 - q_h)(s_1 \cdot R_0 + \rho \cdot r_g) + q_h(1 - q_h)(s_1 \cdot R_0) + q_h^2\left(s_1 \cdot R_0 + \frac{1}{2}\rho \cdot r_g\right) + (1 - q_h)^2\left(s_1 \cdot R_0 + \frac{1}{2}\rho \cdot r_m\right), \quad (10)$$

which can be reduced as

$$U_{1_a} = s_1 \cdot R_0 + \rho \cdot r_g \cdot \left(q_h - \frac{1}{2}q_h^2\right) + \frac{1}{2}\rho \cdot r_m \cdot (1 - 2q_h + q_h^2). \quad (11)$$

If otherwise (i.e., Entrepreneur 1 demonstrates a low level of engagement), Start-up 1 may earn $(s_1 \cdot R_0 + \rho \cdot r_g)$ with probability of $q_l \cdot (1 - q_h)$, or earn $s_1 \cdot R_0$ with probability of $q_h \cdot (1 - q_l)$, or $(s_1 \cdot R_0 + (1/2)\rho \cdot r_g)$ with $q_l \cdot q_h$, or $(s_1 \cdot R_0 + (1/2)\rho \cdot r_m)$ with $(1 - q_l) \cdot (1 - q_h)$. So, the expected yield of Start-up 1 if Entrepreneur 1 is lowly engaged is

$$U_{1_d} = q_l(1 - q_h)(s_1 \cdot R_0 + \rho \cdot r_g) + q_h(1 - q_l)(s_1 \cdot R_0) + q_l \cdot q_h \left(s_1 \cdot R_0 + \frac{1}{2}\rho \cdot r_g\right) + (1 - q_l)(1 - q_h) \left(s_1 \cdot R_0 + \frac{1}{2}\rho \cdot r_m\right), \quad (12)$$

which can be reduced as

$$U_{1_d} = s_1 \cdot R_0 + \rho \cdot r_g \cdot \left(q_l - \frac{1}{2}q_lq_h\right) + \frac{1}{2}\rho \cdot r_m \cdot (1 - q_l - q_h + q_lq_h). \quad (13)$$

To ensure that moral hazard be eliminated and Entrepreneur 1 choose to be highly engaged, the capitalist has to

make sure that Entrepreneur 1 will earn more only when he/she makes more efforts. So, equation (14) holds.

$$U_{1_a} - U_{1_d} \geq \Delta C. \quad (14)$$

Integrate equations (11), (13), and (14), and we obtain

$$\rho(q_h - q_l) \left\{ (r_g - r_m) + \frac{1}{2} [r_m - q_h(r_g - r_m)] \right\} \geq \Delta C. \quad (15)$$

The case that Entrepreneur 2 chooses to be lowly engaged.

As analyzed in Case (1), the expected yield of Start-up 1 when Entrepreneur 1 chooses to be highly engaged is

$$U_{1_b} = s_1 \cdot R_0 + \rho \cdot r_g \cdot \left(q_h - \frac{1}{2}q_hq_l\right) + \frac{1}{2}\rho \cdot r_m \cdot (1 - q_h - q_l + q_l^2). \quad (16)$$

The expected yield of Start-up 1 if Entrepreneur 1 is lowly engaged is

$$U_{1_c} = s_1 \cdot R_0 + \rho \cdot r_g \cdot \left(q_l - \frac{1}{2}q_l^2\right) + \frac{1}{2}\rho \cdot r_m \cdot (1 - 2q_l + q_l^2). \quad (17)$$

The constraint that ensures a high level of engagement by Entrepreneur 1 is

$$\rho(q_h - q_l) \left\{ (r_g - r_m) + \frac{1}{2} [r_m - q_l(r_g - r_m)] \right\} \geq \Delta C. \quad (18)$$

As long as equations (15) and (18) hold true, Entrepreneur 1 would always take a high level of engagement to maximize his/her earnings. With this in mind, the venture capitalist, by designing the renegotiation contract, can manage to urge Entrepreneur 1 to make efforts and reduce moral hazards.

The discussion about Entrepreneur 1 can be safely generalized to Entrepreneur 2, since Entrepreneur 1 is just one picked out at random. So, the conclusion we draw for Entrepreneur 1 is completely adaptable to Entrepreneur 2.

It can be seen that constrained by the portfolio investment contract given limitation of funds, the entrepreneurs have to compete with each other for a second investment. And provided that the portfolio investment has unlimited availability of funds, the entrepreneurs can be hardly motivated due to the lack of competition. The constraint from fund availability gives rise to an increase of marginal income of efforts made. Therefore, a portfolio investment contract with limited availability of funds can urge the entrepreneurs to be highly engaged by bringing in competitive stimulation mechanism.

4. Conclusions

In the context of portfolio investment contract, the problem concerning the principal and agents motivation gets more complex because of more stakeholders involved. More diversified moral hazards arising also require more complicated design of the portfolio contract. The information

asymmetry makes either the venture capitalist or the entrepreneurs can hardly identify the engagement level of the counterpart or the resulted contributions. And, it is even harder to figure out the allocation of efforts by the capitalist among different venture projects included in a portfolio investment contract.

Kanniainen and Keuschnigg [5] and Bengtsson and Sensoy [6], based on research studies on one entrepreneur motivation, analyzed the possible means by which the capitalist can urge more entrepreneurs to make efforts and reduce the bilateral moral hazards in terms of portfolio investment contract. They assumed that, with unlimited availability of funds, the capitalist has enough money to make a second investment in both start-up firms. Thus, no one will fail to get further invested. That means, without fund constraint, the two start-ups do not have to compete with each other. Instead, both of them can get the second investment by renegotiating with the capitalist about income allocation.

A modification is made in this article. It is supposed that the decision about a second investment is constrained by limited fund availability to the venture capitalist. In this case, the entrepreneurs involved in the portfolio contract have to compete with each other to obtain the second investment. The constraint from fund availability gives rise to an increase of marginal income of efforts made. Therefore, a portfolio contract with limited availability of funds can urge the entrepreneurs to be highly engaged by bringing in competitive stimulation mechanism. It is demonstrated that the marginal benefit of efforts paid by the entrepreneurs exceeds the marginal cost, given the limitation of funds available, which will conduce to a high-level engagement of the entrepreneurs. Thus, by adopting the design of renegotiation contract, the venture capitalist can manage to stimulate the entrepreneurs to make efforts, which is to result in moral hazard reduction.

This research contributes to extension of those pioneered by Kanniainen and Keuschnigg and some other scholars [5]. Either from a theoretic perspective or a practical one, the portfolio investment decision is more likely to be constrained by limited fund availability to the venture capitalist. Taking this fund availability constraint into account, our article develops a model for design of the portfolio venture investment contract, which is intended to get an insight into the underlying mechanism and how it works on bilateral moral hazard reduction. Thus, this article contributes to the venture capital finance literature.

One drawback of this research is the lack of comparative analysis of the outcomes provided limitation and no limitation of funds available. Compared with that of unlimited fund availability, how big a marginal effect can the stimulation have that a portfolio contract with limitation of funds imposes on the entrepreneurs. In addition, constrained by limited fund availability, the capitalist has to choose only one of the initially invested venture firms to make further investment at the cost of losing the other, which is likely to inspire the capitalist to contribute more in the selected venture firm development. And the benefits this inspiration effect on the capitalist brings about are also ignored in this

article. A comparative analysis and the equilibrium issue involved will be the focus of further research studies.

Data Availability

No data are required to support the findings of this study.

Conflicts of Interest

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

References

- [1] B. Holmstrom, "Moral hazard in teams," *The Bell Journal of Economics*, vol. 13, no. 2, pp. 324–340, 1982.
- [2] S. N. Kaplan and P. Strömberg, "Characteristics, contracts, and actions: evidence from venture capitalist analyses," *The Journal of Finance*, vol. 59, no. 5, pp. 2177–2210, 2004.
- [3] E. Ball, H. H. Chiu, and R. Smith, "Can VCs time the market? An analysis of exit choice for venture-backed firms," *Review of Financial Studies*, vol. 24, no. 9, pp. 3105–3138, 2011.
- [4] G. Cestone, "Venture capital meets contract theory: risky claims or formal control?" *Review of Finance*, vol. 18, 2001.
- [5] V. Kanniainen and C. Keuschnigg, "The optimal portfolio of start-up firms in venture capital finance," *Journal of Corporate Finance*, vol. 9, no. 5, pp. 521–534, 2000.
- [6] O. Bengtsson and B. Sensoy, "Investor abilities and financial contracting: evidence from venture capital," *Journal of Financial Intermediation*, vol. 20, no. 4, pp. 477–502, 2011.
- [7] N. Khanna and R. D. Mathews, "Can herding improve investment decisions?" *The RAND Journal of Economics*, vol. 42, no. 1, pp. 150–174, 2011.
- [8] N. Khanna, "Optimal contracting with moral hazard and cascading," *Review of Financial Studies*, vol. 11, no. 3, pp. 559–596, 1998.
- [9] Z. He, "A model of dynamic compensation and capital structure," *Journal of Financial Economics*, vol. 100, pp. 351–366, 2010.
- [10] K. Hori and H. Osano, "Managerial incentives and the role of advisors in the continuous-time agency model," *Review of Financial Studies*, vol. 26, no. 10, pp. 2620–2647, 2013.
- [11] L. Lin, "Contractual innovation in China's venture capital market," *European Business Organization Law Review*, vol. 10, pp. 101–138, 2020.
- [12] H. Fu, J. Yang, and Y. An, "Contracts for venture capital financing with double-sided moral hazard," *Small Business Economics*, vol. 53, no. 1, pp. 129–144, 2019.
- [13] A. An and J. Schneider, "Minimum return guarantees, investment caps, and investment flexibility," *Review of Derivatives Research*, vol. 19, 2015.
- [14] S. Chernenko, J. Lerner, and Y. Zeng, "Mutual funds as venture capitalists? Evidence from unicorns," *SSRN Electronic Journal*, vol. 10, pp. 21–39, 2017.
- [15] J. Burchardt, U. Hommel, and D. Kamuriwo, "Venture capital contracting in theory and practice: implications for entrepreneurship research," *Entrepreneurship Theory and Practice*, vol. 40, no. 1, pp. 25–48, 2018.

Research Article

The Model of the Relationship between Urban Rail Transit and Residential Location

Lin Zhang 

School of Economics & Management, Nanchang Hangkong University, Nanchang 330063, Jiangxi, China

Correspondence should be addressed to Lin Zhang; 773651827@qq.com

Received 15 May 2020; Revised 22 June 2020; Accepted 3 July 2020; Published 4 August 2020

Academic Editor: Tingsong Wang

Copyright © 2020 Lin Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to investigate the intrinsic relationship between residence choice and urban rail transit, this paper establishes a housing valuation model, explores the interface link between the rail transit and other transport modes by the establishment of a model, and also obtains the family transportation impedance. According to the balanced housing price, the various districts' hedonic cost, and the generalized transportation impedance, the attractiveness of the various districts with respect to each mobile home is obtained. Satisfaction of any resident is received by establishing a close degree model. Due to the satisfaction and the price, we construct a largest consumer surplus model and then obtain the residence of the greatest consumer surplus for mobile home. Numerical example's result indicates that all high-income mobile homes will chose the residence for the commute destination district, especially the one in the suburbs. Furthermore, the low-income families chose the residence for the commute destination district, which has the rail transit if the income is allowed, or the nearest district to the destination with rail transportation if not. This illustrates that whether a road having urban rail transit plays a significant impact only on the low-income family residence choice when the commuter routes pass through the road and almost has no influence for other families. Hence, it is shown that the reasonable urban planning is important and that urban rail transit should form a network that will play a key role.

1. Introduction

There are different types of land use, including industrial land, commercial land, and residential land, where the residential land is the main factor that influences the traffic volume and is also the generated point of traffic demand for industrial land and commercial land in the rush hours. Since the contradiction between supply and demand of urban traffic is more intensive along with the rapid urbanization, urban rail transit becomes the preferred large-capacity public transport to relieve the traffic pressure in large city due to its characteristics of being safe, speedy, highly efficient, energy-saving, and environment-friendly. Noting that the large city pays much more attention to the scale of rail transit construction, it is meaningful to study the model of the inner relationship between the choice of residential location and the rail transit, which could provide quantitative methods to guide the orderly growth of the urban space and thus guarantee the relationship between urban

space and urban traffic to be interaction and coordination. It is well-known that there are four kinds of land-use transport integration model: the space interaction model (Enjian and Takayuki; Li et al.) [1, 2], the mathematical programming model (Chang and Mackett; Tao et al.; Zhen) [3–5], the random utility model (Chandra and Jessica; Pinjari et al.; Earnhart; Zhen) [6–9], and the bid-rent model (Zhang) [10]. The spatial interaction model distributes the employment and residential location into different zones of the region by considering the distances between the zones and their attractiveness by the force of gravity or entropy. There have been huge literature studies on the application of the spatial interaction models. Wagener [11] established an operational gravity model of spatial interactions, which can be used for urban region, and also defined a framework for the classification and evaluation of urban land use. Roy and Thill [12] demonstrated the interactions between various economic fields and land use, encompassing the environmental, technological, economic, and regulatory factors that affect

land use with a larger scope. Silveira and Dentinho [13] have discussed the relationships between activities and zones based on relative accessibilities, bid-rents, capacities, and technical coefficients through a spatial interaction model. They adjusted the bid-rents for each activity per zone and achieved the equilibrium between supply and demand for each activity per zone. Later, Dentinho and Silveira [14] divided each zone of the city into fourteen soil classes and applied the calibration of the bid-rents to depict conflicts of different activities through the estimated factors of attractiveness in the same soil class. Residents of the different zones were attracted to each other through the weight factor of various soil classes. The advantage of the space interaction model is that it has a simple concept and then is easy to understand. However, this type of model did not illustrate both the features of the location and the decision process of the family. Moreover, this model was descriptive in nature and did not consider explicitly the interaction between transportation and land use.

The mathematical programming model is designed to generate the optimal position of the family, which has been studied intensively in the past. For example, Anas et al. [15–17] developed the mathematical programming model to discuss the congestion tolls in city. The authors divided the urban space into multiple zones and used the general equilibrium to evaluate the influence with and without the tolls policy. They presented that a congestion toll would have an impact on labor supply and also the productivity by changing the labor-leisure trade off. Such a toll could cause workers to work less and thus allocate more time to leisure. Murphy [18] has utilized linear programming to provide insights into the relationship between the location of commuting and housing. The results show that the pattern of relative location advantage has altered sharply for off-peak trip-making but has remained more or less the same for trip-making in the peak period. The mathematical programming model is the combination of the space interaction model and the traffic distribution model and emphasizes the integration between transportation and land use. But the credibility of the results of the mathematical programming model is somewhat questionable due to the lack of the behavioral factors of the decision-makers and the simplicity assumptions on the huge number of modeling conditions.

The random utility model describes the interactions between transportation and the household active area by the household utility maximization principle. The theoretical foundation generally stems from the microeconomic theory; e.g., Jun [19] evaluates the impact of medium- and high-income households' preference for apartments on residential location choice by constructing a random utility-based land-use simulation model of the Seoul metropolitan area. Then, Jun [20] investigates the redistributive effects of Seoul's bus rapid transit system on development patterns and property values by using an urban simulation model. This model is the combination of the Seoul metropolitan input-output model with random utility-based location choice models and endogenous real estate markets. For the improved transportation system with new household and employment distribution patterns, Kelvin et al. [21] discussed residential

and job allocations in the system through the reliability-based network design problem and made a reliability-based land-use and transportation model for the integrated residential and job allocations and transportation network design problem. Coppola and Nuzzolo [22] gave an activities-location choice model with endogenous price which simulates, based on expected random utility principle, the behavior of several agents of the urban system to estimate the spatial distribution of socioeconomic activities within the study area as well as the impact of differential changes in accessibility on the dwelling price. A disequilibrium-based random utility modeling framework is developed for the built space markets by Farooq and Miller [23]. This framework is then used for the Greater Toronto and Hamilton Area's owner-occupied housing market within integrated land-use transportation and environment modeling system. In summary, the random utility model describes the location characteristics and the behavior utility effectively, but the model needs to be extended with more detailed household categories, and the description of the interaction process between transportation and land use is not clear.

The bid-rent model obtains the ideal location by the auction bidding process. For example, Kantor et al. [24] extend a traffic congestion external effect that adds a degree of freedom to the shape of the bid-rent curves and allows them to coincide over contiguous locations in the city where commuting takes place. Based on the bid-rent theory, Zhou and Kockelman [25] examine microscopic equilibrium of the single-family residential land development. Ma and Lo [26] formulated a nested multinomial logit model that combined with the bid-rent process to model residents' location and travel choices. Chong and Shui [27] examine how the rental differential between two locations in a metropolis is determined by the time value of a household. Models of this type are under the assumption that firms and residents compete for land, which is occupied by the agent offering the highest bid. Land rent is then equal to the highest bid. The bid mechanism is coordinated to the decision-making behavior about family location, but it is difficult to establish a clear functional relationship between land use and transportation.

The purpose of this paper is to explore the relationship between the rail transit and residential location. We establish some mathematical models to combine the optimization with the bid-rent. Based on the remaining maximum principle of the household consumption, the influence degree of rail transit where the residents choose residential location is analyzed quantitatively, and the details of the behavior where the residents choose residential location are also explained.

2. Housing Valuation Model

2.1. Estimation of Land Price. The housing is a commodity that can be consumed, and its price is mainly related to facility structures, traffic accessibility, structure characteristics of housing, the public service quality of the area, adjacent region characteristics, local residents' consumption

level, etc. One can divide a city into several districts, which is denoted by the following set:

$$D = \{1, 2, \dots, d_0\}, \quad (1)$$

where d_0 is the number of the districts in the city. Some of these districts contain (only one) rail transit station, and others do not contain the rail transit station. Let

$$D' \subseteq D, \quad (2)$$

where D' is set of the districts, containing rail transit station. Any district $k \in D$ is made up of two types of families: one type is family who is ready to move their residence and wish to purchase or rent house in another district, and the other type is family who does not hope to move its residence to other district, which contains moving families within the district.

At the period time of t , such as one year or two years, let

$$N_0^{kt} = N_1^{kt} \cup N_2^{kt}, \quad (3)$$

where N_0^{kt} is the set of present living families in the k -th district. N_1^{kt} is the set of families in the k -th district that do not move their residence to other districts and N_2^{kt} is the set of families that are ready to move their residence to other districts.

At the period time of t , we assume that the amount of money for purchasing a house is allocated to each day (according to 20-year terms), which is the same as the daily rent of a house. In what follows, the land price and house price are also allocated to daily amount. Let N_3^{kt} be the set of homeowners in the k -th district, renting or selling house to other district residents. For any $j \in N_3^{kt}$, H_j^k is the set of houses ready to be rent or sold and belongs to homeowner j in the k -th district, and, for any $m \in H_j^k$, the transaction price by which moving family i purchases the house m should satisfy

$$p_{jm}^{kt} \geq V_{jm}^{kt}, \quad (4)$$

where p_{jm}^{kt} is the transaction price of house m belonging to homeowner j in the k -th district at the period time of t and V_{jm}^{kt} is the cost of house m belonging to homeowner j in the k -th district at the period time of t . V_{jm}^{kt} is consisted by land price of the k -th district and construction cost of house m (including publicity expenses); that is,

$$V_{jm}^{kt} = V^k + J, \quad (5)$$

where V^k is the land price of the k -th district and J is the construction price of house m .

Let h^k be the completeness of hedonic facilities in the district k that satisfies

$$0 < \varepsilon^0 \leq h^k \leq 1, \quad (6)$$

where $h^k = \varepsilon^0$ is the hedonic facilities completeness of the worst district, which is generally the one at the edge of the city. Let $h^k = 1$ be the completeness of hedonic facilities in the center of city. h^k is related to the business area of shopping center, famous schools within the district, the completeness of hedonic facilities in the adjacent district,

and other factors. Under the condition of the vehicle with free-flow speed, we denote the traffic impedance from the district k to the center of city by u^k . The land price V^k is determined by both h^k and u^k of the district k ; that is, V^k is the function of h^k and u^k . Since the land price of district is positively proportional to the hedonic facilities completeness of the district and the traffic accessibility from the district to the center of city, we thus define V^k as follows:

$$V^k = V^0 h^k - c_0 \beta u^k, \quad (7)$$

where V^0 is the land price of city center which is associated with the consumption level of residents, c_0 is the proportionality coefficient, and β is the urban per capita income level per hour (calculated by five workdays in a week and 8 hours in a weekday).

Supposing that there are only two kinds of houses for sale or for rent, one is high-end and the other is mid-range. Assuming that the land price of same district and the construction costs J of same grade housing are constants and the different grades housing construction costs are different from each other, denote

$$J = \delta J^1 + (1 - \delta) J^2, \quad (8)$$

where J^1 and J^2 are the construction costs of high-end house and mid-grade one, respectively, and δ is the 0-1 function satisfying $\delta = 1$ if homeowner j 's house for sale or for rent is high-end and $\delta = 0$ if not. In summary, we can conclude that

$$V_{jm}^{kt} = V^k + J = V^0 h^k - c_0 \beta u^k + J. \quad (9)$$

2.2. Evaluation Model of the House. At the period time of t , assume that there are a certain number of families in any district that hope to move their residence to another district for improvement on the overall effectiveness of family housing, environment, traffic conditions, and so on. It means that each moving family wishes to buy an ideal residence. For any family i , which is ready to move its residence, it will give an evaluation M_{ijm}^{kt} to sale or rent house m that belonging to homeowner j in the district k , and the family i and the homeowner j will reach an agreement under the condition of $p_{jm}^{kt} \leq M_{ijm}^{kt}$. Using formula (9), we can get M_{ijm}^{kt} as follows:

$$M_{ijm}^{kt} = \frac{[(\alpha M^0 - J)h^k - c_0 \beta u^k + J]}{\alpha}, \quad (10)$$

where M^0 is the evaluation of house in the city center and α is the occupancy of the house cost within M^0 which satisfies $0 < \alpha \leq 1$; that is,

$$\alpha M^0 = V^0 + J. \quad (11)$$

3. Traffic Impedance

In order to relax the people for living and employment from the city center, we should promote rail transit station to link up with conventional public transport, motor vehicles, nonmotor vehicles, and pedestrians and thus form an urban traffic network,

where rail transit is complement with other transportation systems and shares complementary advantages. From the point of view of traffic impedance, choice of residence is traffic impedance compromise, that is, the main activities of family members, but it is not necessarily the optimal choice of each member of the family. Let A_i be the set of i -th family members; for any $a_i \in A_i$, $\Omega_{a_i}^{ih}$ is the average hourly income of member a_i , \mathfrak{F} is the ratio of commuting cost per unit time to $\Omega_{a_i}^{ih}$ for a_i , s_{a_i} is the main destinations, and ζ_{a_i} is the frequency by which a_i goes to s_{a_i} every day. Moreover, r_i is the location of the i -th family, μ_1 represents the fare for the conventional public transport, and μ_2 represents the fare for rail transit. μ_3 represents a bicycle storage costs, μ_4 represents the parking fee for a car, and μ_5 represents the fuel costs per kilometer for a car. Finally, assuming that the family member a_i goes to s_{a_i} at the rush hours every day, the traffic impedance of one return is the same, if a_i arrives the conventional public transport station or rail transit station outside in the destination, then it is arriving the destination s_{a_i} .

3.1. Traffic Impedance of Self-Driving. Let

$$G = (V, E) \quad (12)$$

be the urban traffic network, where V is a finite-points set and E is the directed-road set. Regarding each district as a point in the network, we denote by $R \subset V$ and $S \subset V$ the origin set and the destination one, respectively. For any $r \in R$ and $s \in S$, we use q^{rs} to be the traffic demand from the origin r to the destination s and then have

$$q^{rs} = \sum_{p \in p^{rs}} f_p^{rs}, \quad (13)$$

where p^{rs} is the path set from r to s , p is one of the paths, and f_p^{rs} is the OD pair traffic flow volume from r to s on the path p . For any directed road $a \in E$, the OD pair traffic flow volume from r to s on the directed road a can be obtained as

$$f_a^{rs} = \sum_{p \in p^{rs}} f_p^{rs} \delta_{ap}^{rs}, \quad (14)$$

where δ_{ap}^{rs} is the 0-1 function that satisfies $\delta_{ap}^{rs} = 1$ if $a \in p$ and $\delta_{ap}^{rs} = 0$ if not. Hence, the traffic flow volume on the directed road a can be expressed as

$$f_a = \sum_{r \in R} \sum_{s \in S} f_a^{rs}. \quad (15)$$

Let $v_a(f_a)$ be the vehicle speed on the road a when the traffic flow volume is f_a ; the time for the vehicle passing the path p can be obtained as

$$t_{p,rs} = \sum_{a \in p} \left(\frac{l_a}{v_a(f_a)} + \tau_a \right), \quad (16)$$

where p_{rs} is a path of the OD pair between r and s , l_a is length of the road a , and τ_a is the time the vehicle needs to pass the intersection at the end of the directed road a . Thus, the total generalized commuting transport cost for the i -th family members a_i driving the car from r_i to s_{a_i} on the path p can be obtained as follows:

$$c_{p_{r_i s_{a_i}}} = 2\zeta_{a_i} t_{p_{r_i s_{a_i}}} \mathfrak{F} \Omega_{a_i}^{ih} + 2\mu_5 \sum_{a \in p} l_a. \quad (17)$$

Here (and in what follows), we have calculated the cost of commuting for the back and forth on the road.

3.2. Traffic Impedance of Conventional Public Transport. Let

$$G' = (V', E'), \quad (18)$$

be the network for conventional public transport in the city, where V' is the set of conventional public transport stations and E' is the set of the directed roads for conventional public transport between two adjacent stations (could be a curve). Let P' be the conventional public transport routes set; we denote by $V'_{p'}$ the set of all stations on p' for any conventional public transport route $p' \in P'$ and by $p'_{r',s'}$ the conventional public transport path between station r' and station s' on the bus route for any stations $r', s' \in V'_{p'}$; that is, $p'_{r',s'} \subset p'$.

For any bus road $a' \in E'$ with a' being the bus road between two adjacent conventional public transport stations, we denote the length of a' by $l_{a'}$. Furthermore, let $V'_{p',q'}$ be the set of all conventional public transport stations set on the bus path p' between r' and s' , let $B_{p',q'}$ be intersections set on the bus path p' between r' and s' , and let τ_θ be the time required for vehicle to pass the intersection θ for any $\theta \in B_{p',q'}$. If $r' \in p'$, $s' \in q'$ for any $p', q' \in P'$, and $p' \cap q' = g'$ with $g' \in V'$; $p'_{r',g'} \cup q'_{g',s'}$ is thus the conventional public transport path between station r' and station s' , which is still named as the conventional public transport path and denoted by $p'_{r',s'}$ for simplicity. Now, we propose some assumptions as follows: citizens in the city can reach their destinations at most once transfer by the conventional public transport; the buses' traveling speed is the same as the other vehicles' on the road between two stations at this time; the total delay of the conventional public transport is τ'_ρ minutes for deceleration, up and down the passengers and acceleration at the station $\rho \in V'$. Under the assumptions stated above, the traffic impedance of conventional public transports between r' and s' can be deduced as

$$t_{p_{r^l/s^l}} = \begin{cases} \sum_{a^l \in p_{r^l/s^l}} \left(\frac{l_{a^l}}{v_{a^l}} \right) + \sum_{\theta \in B_{p_{r^l/s^l}}} \tau_\theta + \sum_{\rho=1}^{|V_{p_{r^l/s^l}}| - 1} \tau_\rho^l + \tau^{l/l}, & r^l, s^l \in V_{p^l}, \\ \sum_{a^l \in p_{r^l/g^l} \cup q_{g^l/s^l}} \left(\frac{l_{a^l}}{v_{a^l}} \right) + \sum_{\theta \in B_{p_{r^l/g^l} \cup B_{q_{g^l/s^l}}}} \tau_\theta + \sum_{\rho=1}^{|V_{p_{r^l/g^l}}| + |V_{q_{g^l/s^l}}| - 2} \tau_\rho^l + \tau^{l/l} + \tau^l, & r^l \in V_{p^l}, \\ & s^l \in V_{q^l}, \\ & V_{p^l} \cap V_{q^l} = g^l, \end{cases} \quad (19)$$

where $|\cdot|$ is the element number of set, τ^l is the waiting time for the passenger transfers between the conventional public transports, and $\tau^{l/l}$ is the time for waiting conventional public transport at the origin. Thus, the total generalized commuting traffic cost of the i -th family members a_i by the conventional public transport from r_i to s_{a_i} can be expressed as

$$c_{P_{r_i s_{a_i}^l}} = 2\zeta_{a_i} t_{P_{r_i s_{a_i}^l}} \mathfrak{S} \Omega_{a_i}^{ih} + 2\zeta_{a_i} \mu_1. \quad (20)$$

3.3. Traffic Impedance of Rail Transit. Let

$$G^{ll} = (V^{ll}, E^{ll}) \quad (21)$$

be the network for rail transit in the city, where V^{ll} is the set of rail transit stations and E^{ll} is the set of directed channels for rail transit between two adjacent stations. Let P^{ll} be routes set of the rail transit, for any rail transit route $p^{ll} \in P^{ll}$, $V_{p^{ll}}$ is the set of all stations on the rail transit route p^{ll} , for any stations $r^{ll}, s^{ll} \in V_{p^{ll}}$, $p_{r^{ll}/s^{ll}}^{ll}$ is the rail transit path between stations r^{ll} and s^{ll} on the rail transit route p^{ll} , and

$V_{p_{r^{ll}/s^{ll}}^{ll}}$ is the set of all stations on the rail transit route $p_{r^{ll}/s^{ll}}^{ll}$. Assume that passenger can change at most once from one station to another in the rail transit network, and the station satisfies the shortest transfer which is achieved by the same station transfers, the same hall transfers, or the channel transfer, etc. Moreover, let τ^{ll} be the time for the transfer waiting when changing the rail transit. If stations $r^{ll} \in p^{ll}$ and $s^{ll} \in q^{ll}$ for any $p^{ll}, q^{ll} \in P^{ll}$ and $p^{ll} \cap q^{ll} = g^{ll}$ with $g^{ll} \in V^{ll}$, $p_{r^{ll}/g^{ll}}^{ll} \cup q_{g^{ll}/s^{ll}}^{ll}$ is thus the rail transit path between rail transit stations r^{ll} and s^{ll} . Let $v_{a^{ll}}$ be the travel speed of train between each two adjacent rail stations on the directed channel a^{ll} , let $l_{a^{ll}}$ be the length of track channel a^{ll} , and let τ_σ^{ll} be the stopping time of rail transit at the station σ ; if a_i enters the subway platform from the entrance of the rail transit station r^{ll} , then it takes the rail transit to reach the station s^{ll} by the line p^{ll} and finally gets out of the station s^{ll} ; we can get the traffic impedance of a_i from entering the rail transit station r^{ll} to getting out of the rail transit station s^{ll} as

$$t_{p_{r^{ll}/s^{ll}}^{ll}} = \begin{cases} \sum_{a^{ll} \in p_{r^{ll}/s^{ll}}^{ll}} \frac{l_{a^{ll}}}{v_{a^{ll}}} + \sum_{\sigma=1}^{|V_{p_{r^{ll}/s^{ll}}^{ll}}| - 1} \tau_\sigma^{ll} + \hbar_1 + \hbar_3 + \hbar_2, & r^{ll}, s^{ll} \in V_{p^{ll}}, \\ \sum_{a^{ll} \in p_{r^{ll}/g^{ll}}^{ll} \cup q_{g^{ll}/s^{ll}}^{ll}} \frac{l_{a^{ll}}}{v_{a^{ll}}} + \sum_{\sigma=1}^{|V_{p_{r^{ll}/g^{ll}}^{ll}}| + |V_{q_{g^{ll}/s^{ll}}^{ll}}| - 2} \tau_\sigma^{ll} + \tau^{ll} + \hbar_1 + \hbar_3 + \hbar_2, & r^{ll} \in V_{p^{ll}}, \\ & s^{ll} \in V_{q^{ll}}, \\ & V_{p^{ll}} \cap V_{q^{ll}} = g^{ll}, \end{cases} \quad (22)$$

where \hbar_1 is the time in which passengers reach the station from the entrance of the rail transit, \hbar_3 is the time for waiting for the train, and \hbar_2 is the time for exiting the station. Thus, the total generalized commuting traffic cost of the i -th family members a_i by the rail transit line p^{ll} from entering r^{ll} to going out of s^{ll} can be expressed as

$$c_{P_{r^{ll}/s^{ll}}^{ll}} = 2\zeta_{a_i} t_{P_{r^{ll}/s^{ll}}^{ll}} \mathfrak{S} \Omega_{a_i}^{ih} + 2\zeta_{a_i} \mu_2. \quad (23)$$

3.4. Rail Transit Linking Up with Other Transportation Modes. Urban rail transit plays key role in the urban transport system, whose high accessibility promotes that the regional

economic activity will be attracted to rail transportation corridor. Urban rail transit should have an effective connectivity and coordination with other transportation modes. This may form an integrated urban passenger transport system in which the rail transit is the backbone.

3.4.1. The Traffic Connection Where One End of Rail Transit Is Walking and the Other End Is Conventional Public Transport. Good walking environment must be effective within a certain walking scale. In general, a circle, whose center is the station and the semidiameter is 600 m, could be regarded as attraction area of walking. Let $r^{//}$ be a rail transit station; the impedance that the i -th family walks to $r^{//}$ within the walking attraction area of rail transit station $r^{//}$ can be read as

$$t^{ib} = \frac{l}{v_1}, \quad (24)$$

where l is the distance between the i -th family and rail transit station and v_1 is the average speed of walking. We can get from formula (23) that the total generalized commuting transport cost for the i -th family members a_i waling to rail transit station $r^{//}$ followed by going to $s^{//}$ by rail transit and finally exiting the station $s^{//}$ can be expressed as

$$c_{r^{//}s^{//}}^{iba_i} = 2\zeta_{a_i} t^{ib} \mathfrak{F}\Omega_{a_i}^{ih} + c_{p^{//}s^{//}}. \quad (25)$$

If the location of the i -th family is within the walking attraction area of rail transit station $r^{//}$, using formulas (20) and (25), we can get the total generalized commuting transport cost for the i -th family members a_i walking to rail transit station $r^{//}$ from r_i followed by going to $s^{//}$ by rail transit and exiting the station $s^{//}$ and finally going to s_{a_i} by conventional public transport, which can be expressed as

$$c_{r_i r^{//} s^{//} s_{a_i}}^{iba_i} = c_{r_i r^{//} s^{//}}^{iba_i} + c_{p^{//} s_{a_i}}. \quad (26)$$

3.4.2. The Traffic Connection Where One End of Rail Transit Is Bicycle and the Other End Is Conventional Public Transport. Since the attraction distance of bicycle is 600~2000 m from the rail transit station, we will focus our study on the ring area whose center is the rail transit station and the semidiameter is 600~2000 m, and thus we conclude that the traffic impedance from the i -th family to rail transit station $r^{//}$ by bicycle is

$$t^{ic} = \frac{l}{v_2}, \quad (27)$$

where v_2 is the average speed for riding a bicycle. If the location of the i -th family is within the bicycle attraction area of rail transit station $r^{//}$, we can conclude from formula (23) that the total generalized commuting transport cost for the i -th family members a_i to ride a bicycle to rail transit station $r^{//}$ followed by going to $s^{//}$ by rail transit and finally exiting the station $s^{//}$ can be expressed as

$$c_{r^{//}s^{//}}^{ica_i} = 2\zeta_{a_i} t^{ic} \mathfrak{F}\Omega_{a_i}^{ih} + \mu_3 \zeta_{a_i} + c_{p^{//}s^{//}}. \quad (28)$$

According to formulas (20) and (28), we can get that the total generalized commuting transport cost for the i -th family members a_i to ride a bicycle from home to rail transit station $r^{//}$ followed by going to $s^{//}$ from $r^{//}$ by rail transit and finally going to s_{a_i} by conventional public transport on the bus path p' can be expressed as

$$c_{r^{//}s^{//}s_{a_i}}^{ica_i} = c_{r^{//}s^{//}}^{ica_i} + c_{p^{//}s_{a_i}}. \quad (29)$$

3.4.3. The Traffic Connection Where Both Ends of Rail Transit Station Are Conventional Public Transport. By virtue of formulas (20) and (23), it is known that the total generalized commuting transport cost for the i -th family members a_i to reach rail transit station $r^{//}$ from r_i by conventional public transport followed by going to $s^{//}$ by rail transit and finally exiting the station $s^{//}$ can be obtained as

$$c_{r_i r^{//} s^{//}}^{ida_i} = c_{p_{r_i r^{//}}} + c_{p^{//} s^{//}}. \quad (30)$$

Using formulas (20) and (30), the total generalized commuting transport cost for the i -th family members a_i to get to rail transit station $r^{//}$ from r_i by conventional public transport followed by going to $s^{//}$ from $r^{//}$ by rail transit and finally going to s_{a_i} by conventional public transport can be deduced:

$$c_{r_i r^{//} s^{//} s_{a_i}}^{ida_i} = c_{r_i r^{//} s^{//}}^{ida_i} + c_{p^{//} s_{a_i}}. \quad (31)$$

3.4.4. The Traffic Connection Where One End of Rail Transit Is Car and the Other End Is Conventional Public Transport. According to formulas (17) and (23), we can know that the total generalized commuting transport cost for the i -th family members a_i to reach rail transit station $r^{//}$ from r_i by driving car followed by going to $s^{//}$ by rail transit and finally exiting the station $s^{//}$ can be derived as

$$c_{r_i r^{//} s^{//}}^{iea_i} = c_{p_{r_i r^{//}}} + \mu_4 + c_{p^{//} s^{//}}. \quad (32)$$

By formulas (20) and (32), the total generalized commuting transport cost for the i -th family members a_i to get to rail transit station $r^{//}$ from r_i by driving car followed by going to $s^{//}$ from $r^{//}$ by rail transit and then exiting the station $s^{//}$ and finally going to s_{a_i} by conventional public transport can be expressed as

$$c_{r_i r^{//} s^{//} s_{a_i}}^{iea_i} = c_{r_i r^{//} s^{//}}^{iea_i} + c_{p^{//} s_{a_i}}. \quad (33)$$

3.5. Generalized Transportation Costs of the Family. In general case, the private car owners will drive to the destination unless the generalized commuting transport cost of private car is much more than the generalized commuting transport cost of conventional public transport. Hence, assume that the car does not transfer to conventional public transport separately, and the private car owners will switch to other transportation modes to go the destination when the generalized commuting transport cost of private car is ω

times in comparison to the other transportations ($\omega > 1$); we will consider the generalized transport cost of the i -th family members a_i going to the main destinations s_{a_i} from r_i , which is denoted by $c_{r_i s_{a_i}}$, in the following four different cases:

$$c_{r_i s_{a_i}} = (1 + (\omega - 1)\delta_{a_i}/\delta_{a_i}) \cdot \min \left\{ c_{r_i s_{a_i}}^{i b a_i} \delta_{i b} / + c_{r_i s_{a_i}}^{i c a_i} \delta_{i c} / , \frac{c_{p_{r_i s_{a_i}}} \delta_{a_i}}{\omega}, c_{p_{r_i s_{a_i}}} \right\}, \quad (34)$$

where $\delta_{i b} /$ is the 0-1 function satisfying $\delta_{i b} / = 1$ if the location of the i -th family is within the walking attraction range of 600 m and $\delta_{i b} / = 0$ if not, $\delta_{i c} /$ is the 0-1 function satisfying $\delta_{i c} / = 1$ when the location of the i -th family is within the ring region of the bicycle attraction range of 600~2000 m of rail transit and $\delta_{i c} / = 0$ otherwise, δ_{a_i} is the 0-1 function satisfying $\delta_{a_i} = 1$ if a_i has private car and $\delta_{a_i} = 0$ if not,

- (1) If the district where r_i locates contains rail station $r^{//}$ and the district where s_{a_i} locates contains rail station $s^{//}$, it follows from formulas (17), (20), (25), and (28) that

and $\delta_{a_i} /$ is the 0-1 function satisfying $\delta_{a_i} / = 1$ if the i -th family members a_i drive to the destination and $\delta_{a_i} / = 0$ if not.

- (2) If the district where r_i locates contains rail station $r^{//}$ and the district where s_{a_i} locates does not contain rail station $s^{//}$, we can get from formulas (17), (20), (26), and (29) that

$$c_{r_i s_{a_i}} = (1 + (\omega - 1)\delta_{a_i}/\delta_{a_i}) \cdot \min \left\{ c_{r_i s_{a_i}}^{i b a_i} \delta_{i b} / + c_{r_i s_{a_i}}^{i c a_i} \delta_{i c} / , \frac{c_{p_{r_i s_{a_i}}} \delta_{a_i}}{\omega}, c_{p_{r_i s_{a_i}}} \right\}. \quad (35)$$

- (3) If the district where r_i locates does not contain rail station $r^{//}$ and the district where s_{a_i} locates does not

contain rail station $s^{//}$, we obtain after using formulas (17), (20), (31), and (33) that

$$c_{r_i s_{a_i}} = (1 + (\omega - 1)\delta_{a_i}/\delta_{a_i}) \cdot \min \left\{ c_{r_i s_{a_i}}^{i d a_i} , \frac{c_{p_{r_i s_{a_i}}} \delta_{a_i}}{\omega}, c_{r_i s_{a_i}}^{i e a_i} \delta_{a_i} , c_{p_{r_i s_{a_i}}} \right\}. \quad (36)$$

- (4) If the district where r_i locates does not contain rail station $r^{//}$ and the district where s_{a_i} locates does not

contain rail station $s^{//}$, it follows from formulas (17), (20), (30), and (32) that

$$c_{r_i s_{a_i}} = (1 + (\omega - 1)\delta_{a_i}/\delta_{a_i}) \cdot \min \left\{ c_{r_i s_{a_i}}^{i d a_i} , \frac{c_{p_{r_i s_{a_i}}} \delta_{a_i}}{\omega}, c_{r_i s_{a_i}}^{i e a_i} \delta_{a_i} , c_{p_{r_i s_{a_i}}} \right\}. \quad (37)$$

With the generalized transport cost of i -th family members a_i obtained in formulas (34) to (37) at hand, we can choose the coordinate function from formulas (34) to (37) with respect to the difference of both the family location and the destinations of the family members and finally get the generalized transport cost of the i -th family as follows:

$$c_i = \sum_{a_i \in A_i} c_{r_i s_{a_i}}. \quad (38)$$

4. Hedonic Cost of the Family

Each family has a percentage of household income used for enjoyment and will add its hedonic cost with the increase of income. At the same time, the level of average family hedonic cost is different in different cities, and different locations family in the same city also has different hedonic cost. Let $\bar{\Omega}$ be all family average income levels every day in the city; the hedonic cost of the family is related to the household income

level, and the hedonic cost in the city center is thus a function of $\bar{\Omega}$, which is denoted by $\chi^0(\bar{\Omega})$. Furthermore, the hedonic cost function is $\varepsilon\chi^0(\bar{\Omega})$ at the edge of the city, where ε is the basic cost coefficient for maintaining family's living.

If the transportation for the city is only the conventional public transport, hedonic cost of the family is associated with family's position in the city. More precisely, the hedonic cost is the highest in city center and is inversely proportional to the distance to the city center.

Let d^k be the distance between the city center and the k -th district, d^0 is the semidiameter of the city, and the hedonic cost function χ^k for the k -th district is a point on the line, which links $\chi^0(\bar{\Omega})$ with $\varepsilon\chi^0(\bar{\Omega})$. This means χ^k is a convex combination between $\chi^0(\bar{\Omega})$ and $\varepsilon\chi^0(\bar{\Omega})$. Thus, the hedonic cost of the district without rail transit station can be obtained as

$$\chi^k = \left(1 - \frac{d^k}{d^0}\right)\chi^0(\bar{\Omega}) + \varepsilon\chi^0(\bar{\Omega}) \cdot \frac{d^k}{d^0}. \quad (39)$$

Since rail transit will increase the regional accessibility along the rail way, economic activities in the region are attracted to rail transportation corridor. That is to say, the facilities completeness of the district with rail transits station is higher than the districts in the same location without rail transit station. Hence, the family's hedonic cost of the district with rail transit station is much higher due to the increasing accessibility and thus the shorter distance between the district and the city center. In order to illuminate this fact, the distance d^k must be multiplied by a rail transit impact factor γ if the k -th district contains a rail transit station, which is expressed as $(1 - \gamma d^k/d^0)\chi^0(\bar{\Omega}) + \varepsilon\chi^0(\bar{\Omega}) \cdot \gamma d^k/d^0$ with $0 < \gamma \leq 1$. The impact factor γ is inversely proportional to the ratio between the speed of rail transit and conventional traffic speed along the rail way. Indeed, γ is the smaller when the ratio is bigger and that rail transit impacts accessibility of the k -th district more. When $\gamma = 1$, the rail transit does not impact the district's accessibility. Besides, the completeness of district facilities also impacts the hedonic cost of the district families to some extent. This leads to the fact that χ^k must be amended again, which is achieved by multiplying a factor that is related to the hedonic facilities completeness h^k . Recalling that the hedonic facilities completenesses of the urban fringe district and the city center are ε^0 and 1, respectively; we can calculate the hedonic facilities completeness of the k -th district as

$$h_1^k = \left(1 - \frac{d^k}{d^0}\right)1 + \varepsilon^0 \cdot \frac{d^k}{d^0}. \quad (40)$$

However, this is not the real hedonic facilities completeness of the k -th district and is also named as the calculated completeness. Hence, the ratio between the real completeness h^k and the calculated completeness h_1^k can be used as the adjustment of the hedonic cost for hedonic facilities completeness. With the discussion stated above, the hedonic cost of the district with a rail transit station can be obtained as

$$\chi^k = \left[\left(1 - \gamma \frac{d^k}{d^0}\right)\chi^0(\bar{\Omega}) + \varepsilon\chi^0(\bar{\Omega}) \cdot \gamma \frac{d^k}{d^0} \right] \cdot \frac{h^k}{h_1^k}. \quad (41)$$

5. Choice of Residence

For the choice of residence, consumers always regard the family as a unit, and where to live is the optimal decision of the whole family. The factors that are considered when family chooses residential location include the total traffic impedance of all family members between the main activity location and residential location, hedonic facilities completeness of the district, residential housing prices of the district, and the family income. Each family will wish to choose the coordinate residence, which coincides with its income level. In other words, for some fixed income level, the family housing conditions are pursued as well as possible, the consumption of the hedonic aspects is as high as possible, and family daily traffic impedance is as low as possible. For the period time of t , it thus follows from the housing prices p_{jm}^{kt} , the hedonic cost χ^k , and the total generalized transport cost c_i of any family i that the attraction of the k -th district to the family i is

$$z^{ik} = \frac{p_{jm}^{kt}\chi^k}{c_i}. \quad (42)$$

Let the gross income of family i be Ω^i every day; it is rate y for household total income, which composes of the families daily consumption for housing, hedonic, and transportation; μ is the sum of daily transportation expenses of family, which includes transportation fares and driving fuel consumption; it holds that

$$p_{jm}^{kt} + \chi^k + \mu \leq \eta\Omega^i, \quad (43)$$

is the premise for family i , which chooses the k -th district to be the residential location; i.e., the k -th district is feasible to family i . Thus, we regard all of these districts as a feasible zone for family i which is denoted by K_i . The family i has its own satisfaction degree μ_i^k for the family housing of any district $k \in K_i$. The satisfaction degree is 1 for attractiveness largest housing and is decreased with reducing attractiveness.

Denoting

$$z^i = \max_{k \in K_i} \{z^{ik}\}, \quad (44)$$

we will calculate the close degree of each district's housing attractiveness z^{ik} and z^i in K_i , which is regarded as the satisfaction degree of family i to the district. The close degree is denoted by $\mathbb{R}(z^{ik}, z^i)$ which is expressed as

$$\mathbb{R}(z^{ik}, z^i) = 1 - \frac{z^i - z^{ik}}{z^i}, \quad (45)$$

obviously, where $0 \leq \mathbb{R}(z^{ik}, z^i) \leq 1$. When district's housing attractiveness satisfies $z^{ik} = z^i$, the close degree reaches the maximum value of 1. The difference between them is greater

as the close degree is smaller. Thus, the satisfaction degree of that family i can be obtained as

$$\mu_i^k = \mathbb{R}(z^{ik}, z^i), \quad (46)$$

for the housing in the k -th district.

It is well known that family i has a satisfaction degree μ_i^k to the k -th district, which has the property that the satisfaction degree is larger as the value of the housing will be much higher in the k -th district. On the contrary, the value for family i is smaller. Moreover, the satisfaction degree influences the extent of family i 's value, which has positive correlation with the house value. Therefore, if family i has purchased one house in the k -th district, family i will then obtain the surplus value:

$$U_i^k = M_{ijm}^{kt} - p_{jm}^{kt}(1 - \mu_i^k)^2. \quad (47)$$

Based on each district's housing prices and the housing satisfaction degree, the moving family chooses the residential location with the largest consumer surplus; i.e.,

$$k_i = \arg \max_k (U_i^k - p_{jm}^{kt}). \quad (48)$$

Assuming that all families in the city have individual rationality, family i chooses the k -th district to be the residential location under the condition that

$$U_i^k \geq p_{jm}^{kt}, \quad (49)$$

meaning

$$M_{ijm}^{kt} - p_{jm}^{kt}(1 - \mu_i^k)^2 \geq p_{jm}^{kt}. \quad (50)$$

The prerequisite of transaction success needs to meet the price greater than or equal to costs; at the same time, transaction success also needs to meet the condition that the transaction price cannot be higher than the valuation of homebuyers for the house; that is,

$$V_{jm}^{kt} \leq p_{jm}^{kt} \leq M_{ijm}^{kt}. \quad (51)$$

Finally, it is also subject to family economic conditions; i.e.,

$$p_{jm}^{kt} + \chi^k + \mu \leq \eta\Omega^i. \quad (52)$$

Based on the analysis stated above in formulas (48)–(52), the nonlinear programming model can thus be read as

$$\begin{cases} k_i = \arg \max_k (U_i^k - p_{jm}^{kt}), \\ M_{ijm}^{kt} - p_{jm}^{kt}(1 - \mu_i^k)^2 \geq p_{jm}^{kt}, \\ p_{jm}^{kt} + \chi^k + \mu \leq \eta\Omega^i, \\ V_{jm}^{kt} \leq p_{jm}^{kt} \leq M_{ijm}^{kt}. \end{cases} \quad (53)$$

Now, we will solve the nonlinear programming model (53) and obtain the residential location k^* with the largest surplus value for family i and also the corresponding transaction prices $p_{jm}^{kt^*}$ of house. We conclude that family i will not move its home if there is no district that meets the condition.

6. Algorithm

The algorithm for rail transit and urban residential location choice relation model is as follows:

Step 1: given the initial value of all the various types of nonmoving residential location family numbers $|N_1^{kt}|$ and the moving residential location family numbers $|N_2^{kt}|$ in the district, according to $|N_1^{kt}|$ and $|N_2^{kt}|$, we will calculate all the OD demands q^{rs} , assign initial value to p_{jm}^{kt} , and set the traffic volume $f_a := q_a^0$ on each road, where $k \in D$, $i \in N_2^{kt}$, and $a \in E$.

Step 2: if the OD demand is $q^{rs} > 0$, calculate the traffic speed $v_a(f_a)$ of each road on all of the paths between r and s ; seek the path with the smallest traffic impedance between r and s , expressed as

$$p_0 = \arg \min_p c_{p_{rs}}. \quad (54)$$

Set

$$q^{rs} := q^{rs} - 1. \quad (55)$$

Step 3: if $a_0 \in p_0$, then

$$f_{a_0} := f_{a_0} + 1, \quad (56)$$

and calculate the passengers number of various traffic modes; if all $q^{rs} \leq 0$, then turn to step 4; else turn back to step 2.

Step 4: if $|N_2^{kt}| > 0$, then

$$|N_2^{kt}| := |N_2^{kt}| - 1, \quad (57)$$

calculate

$$\mu_i^k = \mathbb{R}(\alpha_i^k, \alpha_i), \quad (58)$$

and solve formula (53) of the nonlinear programming model, so k^* and $p_{jm}^{kt^*}$ will be obtained.

Step 5: set

$$\begin{cases} |N_1^{k^*t}| := |N_1^{k^*t}| + 1, \\ p_{jm}^{kt} := p_{jm}^{kt^*}. \end{cases} \quad (59)$$

We will stop the calculation if all districts are subject to $|N_2^{kt}| \leq 0$ and will turn back to step 2 if not.

The algorithm contains two loops: one is inside and the other is outside. The inside loop is made up of step 1 and step 2, which is to realize the OD distribution, road traffic impedance calculation, and transport means selection. If an OD demand distribution is end, then it will skip the OD allocation until all OD distributions are end in the cycle. The outside loop is all kinds of families that choose the residential location with the largest family surplus value by their own characteristics, until the selection of all categories of moving families in the district is over.

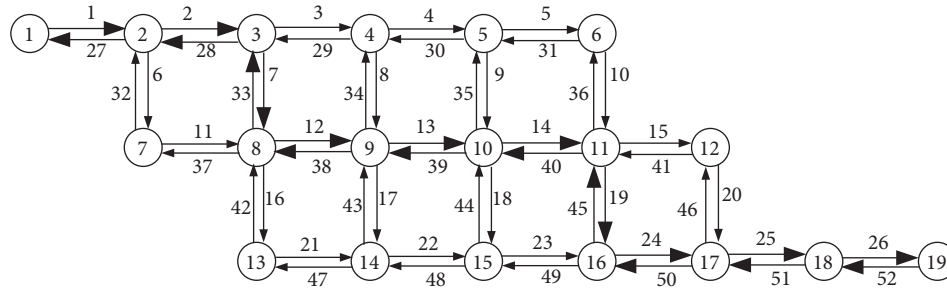


FIGURE 1: Simple traffic network.

7. Example

As is shown in Figure 1, this is a part of one city's transport network, which is divided into 19 (1 to 19) residential areas. There are four main destinations, 1, 19, 6, and 9, where 9 is the city center area. There is a two-way road between the residential areas, the black line in the road network is also a rail transit, and the volume and free-flow speed of two opposite directions on each road are the same. The static data for the length, volume, and free-flow speed of each road are shown in Table 1.

For any residential area, there are commuters that go to the four destinations and then compose 76 OD demands. Take the family as a research unit and do not consider the commute inside the district, such as walking commuter, riding bicycle commuter, and other circumstances. Because the choice of family residential location concerns commuting distance, we only simulate it from the view of public transport and driving car. Each family has two members to go to the major destinations by a kind of traffic modes or the combination of multiple traffic modes. In addition to holiday, the frequency for each family member going to the major destinations is 1.

According to the difference of the destinations for the family's two members, the families of each district are divided into 10 types: destination 9, destination 1, destination 19, destination 6, destinations 9 and 1, destinations 9 and 19, destinations 9 and 6, destinations 1 and 19, destinations 1 and 6, and destinations 19 and 6. These 10 types of families are distributed evenly among all types of families and distinguished by high-income family that has a private car and middle-low-income one that does not have private car. If the two destinations of the high-income family members are the same, then their daily commute is to carpool own car. If the destinations are different, then the car can only be used by the member of outlying destination district, and the other member commutes by public transportation. The car fuel fee is $0.65 \text{ yuan} \cdot \text{km}^{-1}$. If the general cost of total travel is 2 times than that by rail transit, the owners of private car will take rail transit to go to the destination. Furthermore, private cars do not transfer with conventional public transport, and private car drivers do not choose conventional public transport commuter mode.

All members of middle-low-income families commute by public transportation, and all of the paths between each OD pair have conventional public transport that can meet

the traffic demand. Let each bus go ahead 600 m for conventional public transport station. The passenger capacity per conventional public transport at the rush hour is 60 people, and the ticket price is 1 yuan. The average delay of both the conventional public transport stops at the station and the intersection is 60 s. Let the train be made up of 8 block trains, and one-way passenger capacity in the rush hour is 30,000 passengers per hour, and the ticket price is 2 yuan. The bicycle storage cost is 1 yuan per vehicle at the rail station. In order to attract the car driver to transfer the rail transit, the parking fee at the rail station is 10 yuan per vehicle. The time for passengers walking from the rail transit station's entrance to the rail station's platform is 6 minutes, which includes buying the ticket. It takes 4 minutes for passenger to exit the rail transit station. The departure frequency of the rail transit at the rush hour is $12 \text{ veh} \cdot \text{h}^{-1}$. Next, we will assume some known data as follows: the average speed of walking is $1.25 \text{ m} \cdot \text{s}^{-1}$, the average speed of riding bicycle is $20 \text{ km} \cdot \text{h}^{-1}$, rail transit stations located in a going-through residential area, the running speed of rail vehicle between two rail adjacent stations is $120 \text{ km} \cdot \text{h}^{-1}$, and the average delay that rail transit vehicle decelerates, docks the platform, and accelerates at one rail station is 60s. Finally, the completeness of hedonic facilities, the total number of families, the proportion of high-income families, and the number of all types of families in each district are shown in Table 2.

The numerical value corresponding to the relationship between traffic flow/volume and average speed/free flow speed is shown in Table 3 [28].

Fixing the study period time to be one year, the monthly housing costs evaluation M^0 in the city center is 4,000 yuan (calculated by 20 years). The proportionality coefficient c_0 is $1/30$. The monthly per capita income level of the city is 2,400 yuan, and the work per weekday is 8 hours. The monthly incomes of high-income family and middle-low-income one are 12,000 yuan and 6,000 yuan, respectively. Assume that the ratio of moving demand families in each residential area is 20%, it is rate 70% for household total income. The hedonic cost χ^0 in the city center is 2,000 yuan; the basic living cost coefficient for maintaining family is $\varepsilon = 0.5$. The housing supply is enough in each district, which contains a certain number of high-grade housings to meet the high-income families' demand whose price is 1.2 times than that of the ordinary housing. Let the initial traffic saturation be 0.8 on the outlet inlet main roads 25, 26, 27, 1, 51, and 52 in

TABLE 1: Static attributes of the simple traffic network.

Road	1,27	2,28	3,29	4,30	5,31	6,32	7,33	8,34	9,35	10,36	11,37	12,38	13,39
Length (km)	8	6	2.9	3.5	4	5	3.87	4	3	5	5	3.2	3
Volume (veh·h ⁻¹)	3000	3000	4000	3000	2000	2000	3000	4000	3000	2000	2000	4000	4000
Free-flow speed (km·h ⁻¹)	50	50	60	50	30	30	50	60	50	30	30	60	60
Road	14,40	15,41	16,42	17,43	18,44	19,45	20,46	21,47	22,48	23,49	24,50	25,51	26,52
Length (km)	3	6	4	3.5	3.72	3.5	4.5	3	3.8	5	4	4.5	10
Volume (veh·h ⁻¹)	4000	3000	3000	4000	3000	4000	3000	3000	2500	4000	4000	4000	3000
Free-flow speed (km·h ⁻¹)	60	50	50	60	50	60	50	50	30	60	60	60	50

the rush hours, and the initial saturation of the other road traffic is 0.4. According to the different cases for whether there is rail transit between districts 18 and 19 or not, we use the consumer surplus model to simulate the choice of moving families as regards where to live, and the results are shown in Tables 4 and 5, respectively.

When there is rail transit between districts 18 and 19, the result that moving families choose the residential location is given in Table 4. It shows that both the high-income and middle-low-income moving families with their single commuter destination being 1, 19, and 6 will choose the residential location in the destination district. The district is located in suburban area, and the housing price is cheap. If they choose the destination district as the residential location, the commute traffic impedance is the smallest and the consumer surplus value is the largest for the families. For the high-income moving families with their destination being 9, since the family economic conditions can undertake the high housing price in the city center, families choose to live in the commuter destination district 9, where their traffic impedance is the smallest, and the consumer surplus value is the largest. However, for the middle-low-income moving families with their destination being 9, since they cannot afford the high housing price in the city center, these families will choose the residential location with high accessibility rail transit and much lower housing price to make their consumer surplus value be the largest, such as destination 8 or 10.

If the family has two commuter destinations, according to consumer surplus maximum principle, both high-income families and middle-low-income ones have chosen to live in one of the two destinations. If the two destinations of high-income families are in the suburban area and in the city center, respectively, according to the consumer surplus maximum principle, they will choose the residential location in the suburban area; for example, the moving families whose destinations are 9 and 1, 9 and 6, and 9 and 19 will choose districts 1, 6, and 19, respectively. If the two destinations of middle-low-income families are in the suburban area and in the city center, respectively, according to the consumer surplus maximum principle, if suburban commuter destination contains high accessibility rail transit, then most of families will choose their residence in the city center with the same rail transit, and fewer families will choose the commuter destination in the suburbs. If the destination of suburban area has no rail transit station, then a large proportion of families choose to live in the commuter destination district in the suburban area. If the two commuter destinations of families are in the suburban area, both

the high-income and the middle-low-income moving families will choose to live in one of the two destinations, where the rail transit station is located. For example, the moving family's residential location whose destination is 1 and 19 will choose district 19, the moving family's residential location whose destination is 1 and 6 will choose district 1, and the moving family's residential location whose destination is 6 and 19 will choose district 19.

When there is no rail transit between districts 18 and 19, the result where moving families choose the residential location is shown in Table 5. Compared with Tables 4 and 5 showing the following results, when there is no rail transit between districts 18 and 19, there is no effect on the residential location choice for high-income families in the city. If moving families have one single commuter destination or two commuter destinations located in the suburb, there is no effect on the residential location choice of middle-low-income moving families too. For the middle-low-income families whose two destinations are located in the suburban area and city center, respectively, if the commuter route do not pass the road between districts 18 and 19, then it has a little impact on the moving families to choose their residential location, for example, when the family's commuter destinations are 9 and 6 and 9 and 1. If the commuter route passes the road between districts 18 and 19, whether there is rail transit between districts 18 and 19 or not, then it has appreciable impact on the moving family as to how to choose its residential location. If there is a rail transit between districts 18 and 19, then most of middle-low-income moving families with their destinations being 9 and 19 will choose their residence located in city center; indeed, there are 1460 middle-low-income moving families choosing district 9 as their residence, and 942 middle-low-income moving families choosing district 19 as their residence. If there is no rail station between districts 18 and 19, most of middle-low-income moving families with their destinations being 9 and 19 will choose their residence located in the suburban area; indeed, there are 133 middle-low-income moving families choosing district 9 as their residence and 2181 middle-low-income moving families choosing district 19.

In a word, all of high-income moving families choose their residential location located in the commuter destination district and prefer to choose the commuter destination as their residence in the suburban area. This is due to the higher income of these families, and the families own private cars; thus the accessibility that they go to other districts is much higher, so they obtained the largest consumer surplus value. In addition to the middle-low-income family of one

TABLE 2: The static data of the residential area.

Residential area	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Completeness of hedonic facilities	0.8	0.7	0.8	0.9	0.7	0.7	0.6	0.8	1	0.9	0.8	0.5	0.6	0.8	0.7	0.7	0.7	0.7	0.8
The number of families	7000	5000	5000	7000	5000	5000	5000	5000	20000	7000	7000	5000	5000	5000	5000	5000	5000	5000	7000
Proportion of high-income families (%)	20%	20%	20%	25%	15%	20%	15%	20%	30%	25%	20%	10%	15%	20%	15%	20%	15%	15%	20%
	9	560	400	1000	2100	1000	500	400	1500	1400	700	500	750	1000	1000	500	250	250	350
	1	2100	1750	500	350	250	1500	250	1000	350	350	250	250	250	250	100	100	100	140
	19	210	150	250	350	250	100	100	1000	350	700	500	250	500	250	500	1000	1500	2800
	6	210	150	250	350	500	1000	150	1000	700	1400	500	500	500	250	500	500	250	350
	9,1	2100	1500	1500	1050	250	1750	1250	3000	350	350	250	1500	500	250	50	50	50	70
The distribution of main destinations for the family	9,19	140	100	250	350	250	100	500	3000	1050	700	500	500	1000	1250	1250	1000	1000	1400
	9,6	140	100	250	1050	1500	150	500	3000	1400	1400	1000	250	250	500	500	500	500	700
	1,19	700	400	350	350	250	400	350	1000	350	350	250	500	500	500	500	500	500	350
	1,6	700	350	500	700	500	350	350	1000	350	350	250	250	250	250	100	100	100	140
	19,6	140	100	150	350	250	100	100	1000	700	700	1000	250	250	500	1000	1000	750	700

TABLE 3: The relationship between vehicle’s average speed and path flow.

Traffic flow/volume	≤0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	≥1.3
Average speed/free-flow speed	1	34/35	31/35	24/35	19/35	14/35	9/35	4/35	3/35	2/35	0

TABLE 4: Mobile families choose the result of residential location with rail transport between districts 18 and 19.

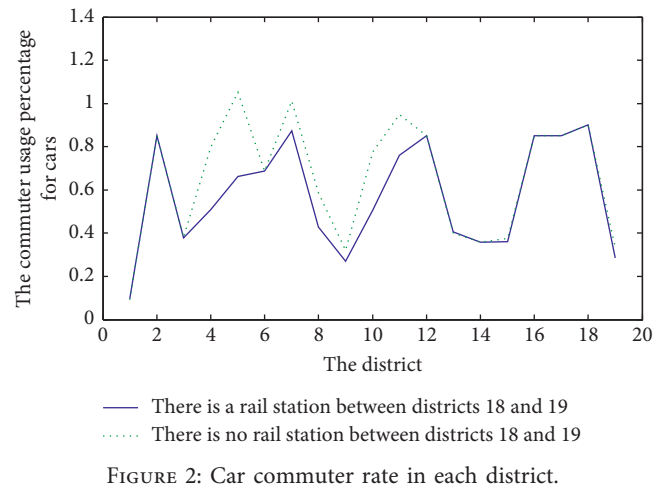
Class of family destination	9	1	19	6	9 and 1	9 and 19	9 and 6	1 and 19	1 and 6	6 and 19					
The number of middle-low-income families	1725	0	1254	1620	1781	1469	156	2389	1850	464	1460	942	1342	1127	1507
The number of high-income families	0	857	0	404	427	368	677	0	0	609	640	0	331	296	356
Selected district	8	9	10	1	19	6	1	9	9	19	6	9	19	1	19

TABLE 5: Mobile families choose the result of residential location with no rail transport between districts 18 and 19.

Class of family destination	9	1	19	6	9 and 1	9 and 19	9 and 6	1 and 19	1 and 6	6 and 19					
The number of middle-low-income families	1725	0	1254	1620	1781	1469	96	2449	133	2181	1473	929	1342	1127	1507
The number of high-income families	0	857	0	404	427	368	677	0	0	609	640	0	331	296	356
Selected district	8	9	10	1	19	6	1	9	9	19	6	9	19	1	19

single commuter in destination 9, the middle-low-income moving families also choose the commuter destination as their residence located in the suburban area. If the middle-low-income families’ two commuter destinations are located in the suburban area, then they prefer to choose the commuter destination district with the rail transit of high accessibility. The middle-low-income families with one single commuter in destination 9 choose their residence near the city center with rail transit of high accessibility. The reason for this is that the lower-income families do not have private cars and thus go to other districts only by the public transport. The completeness of public traffic facilities is the highest in the city center, and it has high accessibility, while the districts with rail transit station have high accessibility. Therefore, they integrated family income, house prices, district facilities completeness, hedonic cost, and other factors and thus choose residence located with the largest consumer surplus value. Whether a road has rail transit or not has a significant impact on the middle-low-income families choosing their residence located when the commuter route passes the road, while it almost does not have impact on the other families.

Figure 2 shows the contrast of the commuter usage percentage for cars in each district, when there is rail transit between districts 18 and 19 or not. We can know the following: the commuter usage percentage of cars has obviously changed in districts 4, 5, 7, 8, 9, 10, 11, and 19, and there is a smaller change in the remaining districts. More precisely, when there is not rail transit between district 18 and 19, these significant changes happened district (4, 5, 7, 8, 9, 10, 11 and 19), and high-income families whose commuter destination district is 19, the commuter usage of cars increase 29%, 39%, 13.6%, 16%, 5%, 27%, 19% and 5% respectively, comparing with the case that there is rail transit between district 18 and 19.comparing with the case that there is rail transit between district



18 and 19, comparing with the case that there is rail transit between district 18 and 19. Indeed, this is because the distance between destination 19 and these districts is much bigger; if there is rail transit between districts 18 and 19, the accessibility that the family goes to destination district 19 from these districts by conventional public transport and then transfers to rail transit or direct rail transit is much higher than that by driving private car. However, if there is no rail transit between districts 18 and 19, the accessibility that the family goes to destination district 19 by conventional public transport will decrease, and the commuters of the high-income families will thus drive private cars in this case for the lowly comfortable degree and other reasons. Since districts 1, 2, and 3 are very far from 19 district, the general transportation cost of driving commuter is much higher. Notice that districts 1, 2, and 3 have rail transit station; whether there is rail transit between districts 18 and 19 or not, the high-income earners whose commuter

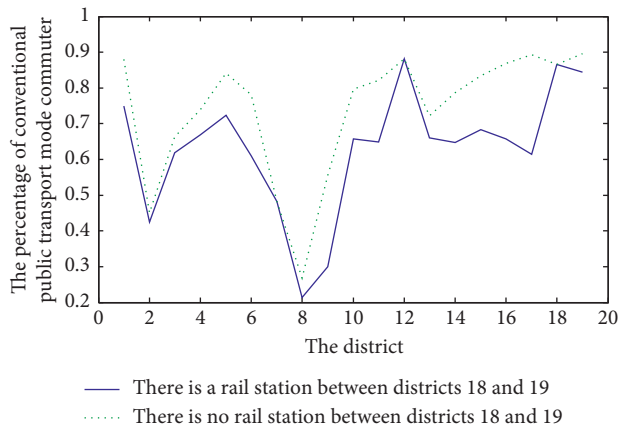


FIGURE 3: Conventional public transport commuter rate of residents in each district.

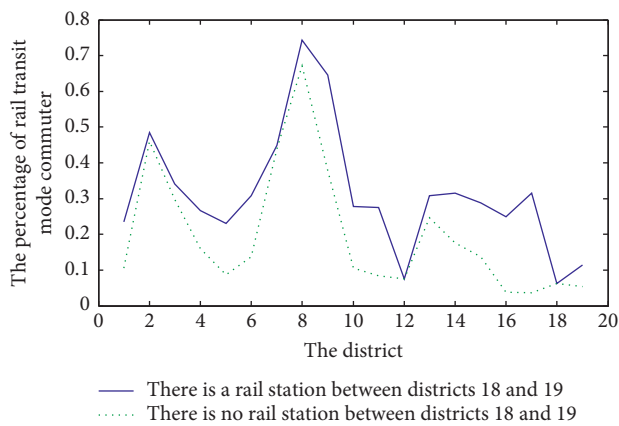


FIGURE 4: Rail transport commuter rate of residents in each district.

destination is district 19 will choose rail transit or rail transit transfer instead of conventional public transport to commute. The remaining districts are not in rail transit line but are close to district 19; the number of high-income earners commuting district 19 by the rail transit modes is much smaller; therefore, whether there is rail transit between districts 18 and 19 or not, the commuter usage percentage of cars has a small change in these districts.

It is shown from Figures 3 and 4 that as the rail transit between districts 18 and 19 is changed from nonexistence to existence, the percentage of rail transit mode commuter has increased almost in all districts, and the percentage of conventional public transport mode commuter has decreased correspondingly.

If there is no rail transit between districts 18 and 19, the rail transit commuting percentage of residents in districts 4, 9, 11, 16, and 17 has decreased by 21%, 27%, 29%, 21%, and 28%, respectively, in comparison to the case where there is rail transit between districts 18 and 19. The rail transit commuting percentage of residents in districts 1, 5, 6, 10, 14, and 15 has decreased by 10~20%, the rail transit mode commuter number in the other districts has decreased by 1~10% at the same time, and all of these residents change

their transport mode from rail transit to conventional public transport. Thus, the reasonableness for the design of the part of an urban road network will influence not only residents choosing their transport mode in a part of the city but also transport mode and traffic behavior of the whole city's residents.

The example shows the importance of reasonable urban planning, as well as the important role of rail transit in urban public transport systems. At the same time, we find that one single rail transit line cannot meet the requirements of the transport mode backbone. The primary reason is the limitations of the passenger flow attracted scope and the line route directions. Thus, rail transit has to form a network and thus will play a key role in the whole traffic network.

8. Conclusions

The relationship model between residential location selection and rail transit system is discussed in the paper. We have established the impedance model with the traffic connects each other between various transportation modes, so the general traffic impedance of each family is obtained. The choice of living location is the optimal decision of the whole family. Considering the factors of the household income, the workplace characteristics of members, the generalized traffic impedance of the family, the facility structure of each district, and housing price, each mobile family will choose destination district as the residential location with the largest consumer surplus value. Urban traffic structure has been optimized by rail transit. Indeed, the residents prefer to live in a specific location if the destination district has rail station, and this may decrease the traffic demand of the city and optimize the urban spatial structure. At the same time, the city manager could guide residents to choose their residence location by attracting more businesses and enterprises to the district with rail station, which improved the structure of urban space to be more perfect.

Data Availability

The data used to support the study are included within the article.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (no. 51368046) and Natural Science Foundation of Jiangxi Province, China (no. 20151BAB201028).

References

- [1] Y. Enjian and M. Takayuki, "A study of an integrated intercity travel demand model," *Transportation Research Part A: Policy and Practice*, vol. 39, pp. 367–381, 2005.

- [2] L. B. Li, M. Wang, Z. Dong et al., "Method of parking index based on parking function and location conditions," *China Journal of Highway and Transport*, vol. 23, no. 1, pp. 111–115, 2010.
- [3] J. S. Chang and R. L. Mackett, "A bi-level model of the relationship between transport and residential location," *Transportation Research Part B: Methodological*, vol. 40, no. 2, pp. 123–146, 2006.
- [4] F. Tao, J. Y. Zhang, F. Akimasa et al., "An integrated model system and policy evaluation tool for maximizing mobility under environmental capacity constraints: a case study in Dalian City, China," *Transportation Research Part D: Transport and Environment*, vol. 15, pp. 263–274, 2010.
- [5] R. B. Chandra and Y. G. Jessica, "A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels," *Transportation Research Part B: Methodological*, vol. 41, pp. 506–526, 2007.
- [6] L. Zhen, "Modeling of yard congestion and optimization of yard template in container ports," *Transportation Research Part B*, vol. 90, pp. 83–104, 2016.
- [7] A. R. Pinjari, C. R. Bhat, and D. A. Hensher, "Residential self-selection effects in an activity time-use behavior model," *Transportation Research Part B: Methodological*, vol. 43, no. 7, pp. 729–748, 2009.
- [8] D. Earnhart, "Combining revealed and stated data to examine housing decisions using discrete choice analysis," *Journal of Urban Economics*, vol. 51, no. 1, pp. 143–169, 2002.
- [9] L. Zhen, "Tactical berth allocation under uncertainty," *European Journal of Operational Research*, vol. 247, no. 3, pp. 928–944, 2015.
- [10] L. Zhang, W. Du, and Q. Q. Guo, "Negotiation mechanism model of relationship between residential location and traffic system," *Journal of Traffic and Transportation Engineering*, vol. 10, no. 6, pp. 102–109, 2011.
- [11] M. Wegener, "Operational urban models state of the art," *Journal of the American Planning Association*, vol. 60, no. 1, pp. 17–29, 1994.
- [12] J. R. Roy and J.-C. Thill, "Spatial interaction modelling," *Fifty Years of Regional Science*, vol. 83, pp. 339–361, 2004.
- [13] P. Silveira and T. Dentinho, "Spatial interaction model of land use - an application to Corvo Island from the 16th, 19th and 20th centuries," *Computers, Environment and Urban Systems*, vol. 34, no. 2, pp. 91–103, 2010.
- [14] T. P. Dentinho and P. Silveira, "Spatial interaction model with land use to analyse the impact of designed accessibilities," in *Proceedings of the 8th world congress of RSAI 2008 an Application to Corvo Island from XVI, XIX and XX Centuries*, Sao Paulo, Brazil, 2008.
- [15] A. Anas and R. Xu, "Congestion, land use, and job dispersion: a general equilibrium model," *Journal of Urban Economics*, vol. 45, no. 3, pp. 451–473, 1999.
- [16] A. Anas, "The optimal pricing, finance and supply of urban transportation in general equilibrium: a theoretical exposition," *Economics of Transportation*, vol. 1, no. 1-2, pp. 64–76, 2012.
- [17] A. Anas and T. Hiramatsu, "The economics of cordon tolling: general equilibrium and welfare analysis," *Economics of Transportation*, vol. 2, no. 1, pp. 18–37, 2013.
- [18] E. Murphy, "Urban spatial location advantage: the dual of the transportation problem and its implications for land-use and transport planning," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 1, pp. 91–101, 2012.
- [19] M.-J. Jun, "Redistributive effects of bus rapid transit (BRT) on development patterns and property values in Seoul, Korea," *Transport Policy*, vol. 19, no. 1, pp. 85–92, 2012.
- [20] M.-J. Jun, "The effects of housing preference for an apartment on residential location choice in Seoul: a random bidding land use simulation approach," *Land Use Policy*, vol. 35, pp. 395–405, 2013.
- [21] K. W. Kelvin, S. C. Yim, A. Wonga, A. Chen et al., "A reliability-based land use and transportation optimization model," *Transportation Research Part C*, vol. 19, pp. 351–362, 2011.
- [22] P. Coppola and A. Nuzzolo, "Changing accessibility, dwelling price and the spatial distribution of socio-economic activities," *Research in Transportation Economics*, vol. 31, no. 1, pp. 63–71, 2011.
- [23] B. Farooq and E. J. Miller, "Towards integrated land use and transportation: a dynamic disequilibrium based micro-simulation framework for built space markets," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 7, pp. 1030–1053, 2012.
- [24] Y. Kantor, P. Rietveld, and J. van Ommeren, "Towards a general theory of mixed zones: the role of congestion," *Journal of Urban Economics*, vol. 83, pp. 50–58, 2014.
- [25] B. Zhou and K. Kockelman, "Micro-simulation of residential land development and household location choices: bidding for land in Austin, Texas," *Transportation Research Record*, vol. 2077, pp. 106–112, 2008.
- [26] X. Ma and H. K. Lo, "Modeling transport management and land use over time," *Transportation Research Part B: Methodological*, vol. 46, no. 6, pp. 687–709, 2012.
- [27] T. T.-L. Chong, K. C.-W. Shui, and V. H. Wong, "The nexus between labor wages and property rents in the Greater China area," *China Economic Review*, vol. 30, pp. 180–191, 2014.
- [28] S. Markose, A. Alentorn, D. Koesrindartoto, P. Allen, P. Blythe, and S. Grosso, "A smart market for passenger road transport (SMPRT) congestion: an application of computational mechanism design," *Journal of Economic Dynamics and Control*, vol. 31, no. 6, pp. 2001–2032, 2007.

Research Article

Analysis of the Impact of Interest Rate Liberalization on Financial Services Management in Chinese Commercial Banks

Tian Meng, Mengnan Sun, Yixuan Zhao, and Bo Zhu 

Shanghai Sci-Tech Finance Institute, Shanghai University, Shanghai 200072, China

Correspondence should be addressed to Bo Zhu; zhubo@staff.shu.edu.cn

Received 29 May 2020; Revised 28 June 2020; Accepted 7 July 2020; Published 3 August 2020

Academic Editor: Lu Zhen

Copyright © 2020 Tian Meng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the advancement of China's interest rate marketization reform, commercial banks' net interest margin has narrowed. This paper selects 16 representative listed banks as the research object and conducts an empirical analysis from the two dimensions: profit level and profit structure. The study finds that the marketization of interest rates promoted the narrowing of net interest margins caused by the narrowing of net interest margins, and the profitability of commercial banks was suppressed. The narrowing of net interest spreads forced commercial banks to actively expand their intermediate business activities and adjust business structure correspondingly. The narrowing of net interest spreads has different impacts on the profitability of commercial banks of different sizes.

1. Introduction

Interest rate liberalization means that interest rates are determined by the market factors rather than being set by the regulators. Through market competition mechanisms, financial institutions are allowed to set the price interest rates independently. Monetary authorities can only affect the level of interest rates through macrocontrol prudential policies indirectly. China's interest rate marketization process has adopted a form of gradual reform. With the cancellation of the deposit interest rate ceiling, the marketization of interest rates was accelerated in 2015. In August 2019, the People's Bank of China announced an improvement in the formation mechanism of the basic loan interest rate (LPR), which further deepened the market-based interest rate reform. Interest rate liberalization helps strengthen the market's leading role in resource allocation, optimizes the allocation of financial resources, and further promotes economic growth. Interest rate liberalization will also intensify competitions among financial institutions. The research found that in the process of interest rate liberalization, there is a clear phenomenon of narrowing the deposit and loan spreads. The profitability of commercial banks is affective since the main source income of the traditional business

model of China's commercial banks depends on the net interest. As the main intermediary institution in the financial market, China's commercial banks play an important role in the process of capital financing in China. In the context of interest rate liberalization, the spread of deposits and loans is continuously and gradually narrowing, and the profit level and income structure of commercial banks are affected to a certain extent. How commercial banks adjust their business structure and maintain sustained profitability become more and more important and worth to be analyzed. This article will select the panel data of 16 listed commercial banks in China from 2008 to 2018, construct a regression model, reveal the different ways of changing the business structure of commercial banks through empirical researches, and make several recommendations.

2. Literature Review

Regarding net interest margins, interest rate liberalization, and commercial bank profitability, although there have been relevant research results, domestic and foreign scholars still have different research focuses.

The study of net interest spread by foreign scholars started earlier and mainly focused on the affecting factors of

deposit and loan spreads. Ho and Saunders [1] started the research by constructing a market maker model at the beginning and used this empirical research to find the level of interest spread influential by factors such as the bank's own transaction size, the external market structure of the banking industry, and the central bank's interest rate policies. The related model is continuously improved on the basis of the model by scholars including Angbazo [2], Carbó and Rodríguez [3], Claeys and Vander [4], Maudos and Solís [5], and Islam [6]. It is proved in the expanded model that the level of interest spread is affected by monetary policy, institutional factors, market structure, and the bank's own factors. There are not many foreign research materials focusing on the impact of interest spreads on bank profitability, however. Saunders and Schumacher [7], Claassen, Coleman, and Donnelly [8] studied from the perspective of the impact of macro interest rate fluctuations on bank performance. Regarding the impact of interest rate liberalization, foreign scholars have more research results from the perspective of financial deepening and financial liberalization. The existing literatures have not formed a unified conclusion on the advantages and disadvantages of financial market liberalization on a country's economy. As the literature on its merits increases, Dwumfour [9] proposed that financial liberalization helps to diversify risks, monitor agents, and promote the financial system to provide financial services more effectively and believes that in a region with comprehensive economic growth, less competitive reduction will instead lead to inefficiency in the banking market. Brissimis [10] and Shim [11] found out that low interest spreads have an impact on the bank's main income structure. They explored the path of interest rate fluctuations that led to fluctuations in the financial environment and thus affected bank income.

The domestic research on the deposit and loan spreads of commercial banks started relatively late due to China's early strict economic and financial regulations. The marketization of interest rates was launched until the end of the last century, and the issue of net interest spreads gradually attracted the attention of scholars. Regarding the study of bank spreads, scholars' discussions mainly focused on the influencing factors of bank spreads and the trend of spreads. Chinese scholars believe that the factors influencing spreads can be divided into risk factors, quality factors, market environment, and financial policy factors, and they are also related to the heterogeneous characteristics of banks. Zhou [12] believes that for many domestic commercial banks, changes in commercial banks' net interest spreads are mainly affected by capital adequacy regulation, benchmark deposit and loan spreads, and deposit interest rate controls. Wang and Guo [13] found that the relevant interest rate policies have led to a significant reduction in the net interest margin of the banking industry, rather than the development of interest income, which helps banks to cope with the adverse impact of interest rate liberalization on the net interest margin. Zhou [14], Zhang [15], Sui and Xing [16] proposed a bank spread decision model suitable for China's national conditions based on the HS market maker model. The study found that the factors that influence the spread

include default risk, interest rate risk, liquidity level, degree of risk aversion, asset size, last year's net interest spread, management model, market power, and policy support and restrictions. Ding Ning [17] conducted a structural analysis of the high deposit and loan spreads in China's banking industry and believed that there are two main reasons for the high deposit and loan spreads in China. There is a contradiction between supply and demand in the capital market of "large banks and small enterprises." Yin et al. [18] believe that there is a negative correlation between relatively liberalized economic environment and bank spreads. Domestic scholars have not yet formed a unified view regarding the trend of interest spreads. Xiao and Wu [19] and others believe that market-oriented reforms will lead to a continuous narrowing of net interest margins. However, other scholars, Ba et al. [20], believe that although the current and future net interest spreads are narrowing, from a longer-term perspective, there will be long-term stability or expansion after narrowing trend. Chinese scholars have constructed a system of factors affecting the profitability of banks from different perspectives regarding the profitability of commercial banks.

Scholars such as Zhang [21], Liu [22], and Wang [23] believe that the bank's return on assets is the best variable to measure the bank's profit level. They have reached similar conclusions based on different data and models—the individual characteristics of the banking institution and the macroeconomic and financial structure of the bank's environment are significant factors influencing its profitability. Niu and Qiu [24] believe that the net interest spread has a significant impact on the sustainable development of the bank. The size of the net interest spread affects the bank's ability to continue to obtain cash flow in the future, and thus have an impact on the bank's risk exposure, liquidity level, and capital adequacy factors. Regarding the impact of interest rate liberalization on the operation and development of China's commercial banks, scholars generally agree that interest rate liberalization has brought challenges and opportunities to the traditional business models of commercial banks. Liu [25] believes that the marketization of interest rates will inevitably pose serious challenges to the business model of traditional commercial banks. Banks must consider potential risks, actively adjust the income structure, and implement on diversification strategy after setting up the basis of a traditional asset and liability businesses. Zhu et al. [26] believe that the marketization of interest rates, the entry of foreign banks, and the competition of local peers have made domestic banking operations unprecedentedly grim, but empirical research proves that the improvement of the marketization level and financial marketization level of the locality has a positive impact on bank performance. Wang and Guo [13] proved that the market-based interest rate reforming has lowered the bank's net interest margin and that national banks can meet the challenges by developing noninterest income, but regional banks still have difficulty using this strategy.

Foreign scholars have relatively abundant research materials on the effect of interest spreads, but there are insufficient literature papers on the effect of interest spreads

on the income structure of banks and the differences between commercial banks of different natures. There is a shortage of domestic literature research on how interest spreads are affecting the operating structure of commercial banks. This paper uses a comparative analysis method to analyze the otherness affected by different commercial banks; the research variables are relatively comprehensive, and the corresponding indicators are selected from the two perspectives of business structure and the rate of return, to a certain extent, enhances the comprehensiveness of the bank's profitability research and reliability.

3. Theoretical Analysis and Research

Hypothesis of Net Interest Margin and Profit Structure

3.1. Net Interest Spread Narrows in Volatility. In general, interest spreads include spreads between deposit interest rates and loan interest rates of commercial banks, spreads between central bank rediscount rates and commercial bank loan interest rates, and spreads between domestic financial market interest rates and international financial market interest rates. The interest spreads in this article refer to the narrowly defined spreads, that is, the ratio of the bank's interest income to all interest-earning assets minus the ratio of interest expenditure to all interest-paying assets.

The net interest spreads of different commercial banks in China show a narrowing trend in volatility. From the perspective of loans, with the advancement of interest rate liberalization and the enrichment and innovation of financing tools in the capital market, commercial banks' competition in the loan market is getting fiercer. As for deposit interest rates, deposits are important sources of funds for commercial banks and are the basis for banks to carry out all other businesses, resulting in a certain price rigidity in the adjustment of bank deposit rates. Comparing the data of various banks, we can find that the change in the net interest spread of large state-owned commercial banks is smaller than that of joint-stock banks and city commercial banks. The changing trend is similar to that of national joint-stock commercial banks (see Figure 1).

3.2. Revenue Structure. Operating income reflects the operating results of the bank. From the perspective of income structure, the proportion of noninterest income is an important indicator to measure the sustainable development of the bank. From 2008 to 2018, the proportion of the noninterest income of commercial banks in China showed an increasing trend, rising from 15% to 30% on average. China's commercial banks adopt a separate business model and are subject to stricter supervision. Noninterest income mainly comes from basic businesses such as settlement and bank cards and rarely involves complex products or services. Moreover, business development in investment banking, private banking, and financial derivatives is still in its infancy. Therefore, the proportion of noninterest income in banks is relatively small. As the mixed operations dominate foreign commercial banks, intermediary business products

and services are rich in variety and have a wide range of coverage, and noninterest income accounts for more than 40% to 60%. China's commercial banks are quite different from mature international banks in terms of income structure, and there is still room for development in the noninterest income business of Chinese banks. After analyzing the data, the proportion of the noninterest income of domestic commercial banks has continued to expand. Among them, joint-stock commercial banks generally have a higher proportion of noninterest income, followed by large state-owned banks, and noninterest income of regional city commercial banks. It is the most dependent part on deposit and loan spread income (see Figure 2).

According to the financial constraints theory, since the government's interest rate control is actually a monopoly privilege granted to banks, this regulation increases the bank's net interest margin. The government's interest rate liberalization is used to make the market play a decisive role in determining the level of interest rates. With the advancement of interest rate marketization reform, commercial banks' independent pricing space has expanded, and market competition has become fierce. According to the theory of interest rate determination, in the money market, the floating range of the benchmark interest rate set by government intervention is cancelled, and the actual interest rate will definitely return to the equilibrium interest rate, which causes the operating cost of commercial banks to rise. However, due to lower barriers to entry and entry of foreign banks, commercial banks have limited room to increase both loan interest rates and business rates as suppliers. In order to attract depositors, expand credit business, high interest rate lending reserves and competing for price lending may also become the normal state of future commercial banks. Therefore, interest rate liberalization has reduced the bank's net interest margin in terms of cost and revenue. At the same time, fierce market competition will force commercial banks to expand noninterest income, reduce their dependence on interest spreads, and optimize the business structure. This paper presents the following research hypotheses:

H1: according to the theory of financial constraints and the theory of interest rate determination, the market-oriented interest rate policy has changed the pricing mechanism of the money market. The monopoly rents obtained by commercial banks have gradually decreased, the deposit and loan spreads have narrowed, and the profitability of commercial banks has been suppressed.

H2: according to the financial deepening theory, the marketization of interest rates promotes the narrowing of net interest spreads and the narrowing of net interest spreads, forcing commercial banks to actively expand the intermediate business and increase the proportion of noninterest income, which has a positive effect on the optimization of commercial banks' business structure.

H3: different commercial banks have different scales. Their development strategies, market positioning, and

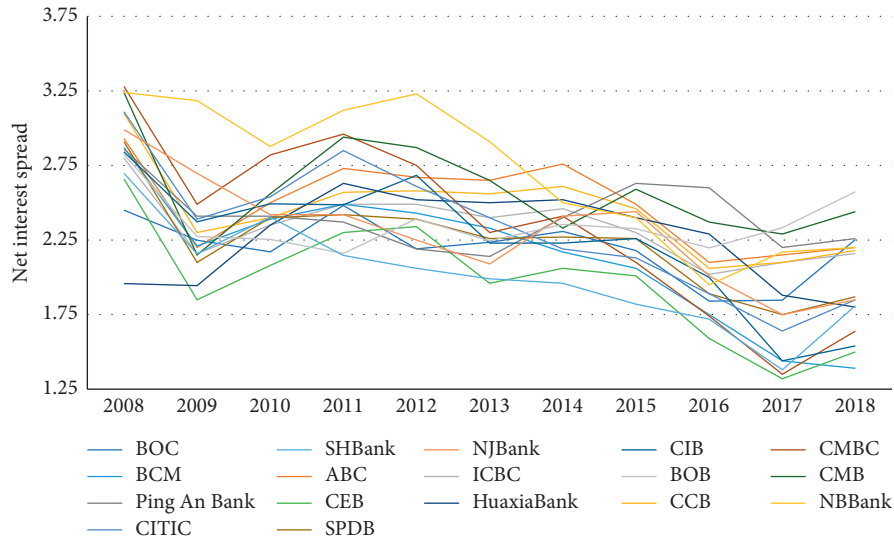


FIGURE 1: Trend chart of narrowing net interest spread fluctuations of commercial banks. Source of data: the annual reports of banks from 2008 to 2018.

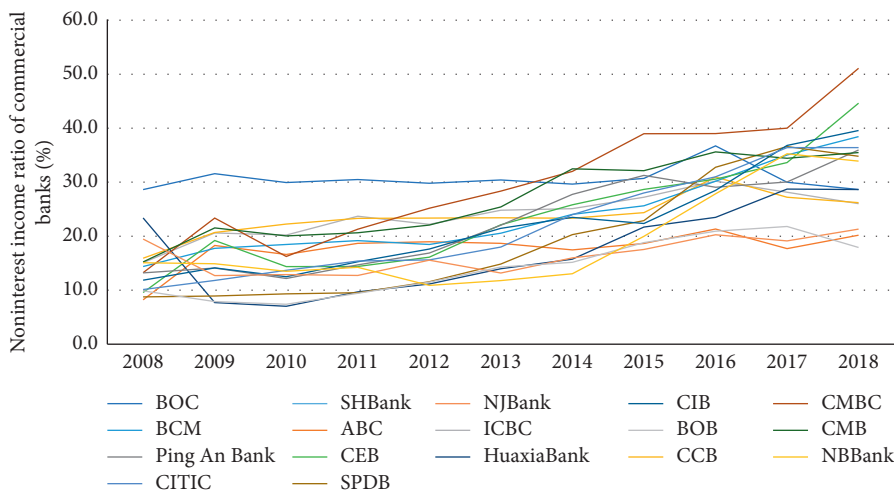


FIGURE 2: Change trend of noninterest income ratio of commercial banks from 2008 to 2018. Source of data: The Bank-scope database and bank annual reports.

business models are heterogeneous. As interest rate liberalization advances, the net interest margin narrows, and there are differences in the degree of impact on the business of commercial banks of different sizes.

4. Selection of Samples, Variables, and Models

4.1. Research Sample. This paper selects 16 listed banks, such as Bank of China and Agricultural Bank of China to be a research sample. The sample data is collected from 2008 to 2018, including the process of China’s interest rate marketization. Three different types of institutions were selected to be the samples: large state-owned banks, joint-stock commercial banks, and medium-sized city commercial banks. Financial data such as income, cost, and capital

structure of commercial banks are collected from the Bank-scope database. Economic data such as GDP, total deposits in the banking sector, and the market value of the stock market are collected from the website of the National Bureau of Statistics and the EPS database. Internet financial data such as third-party payment scale and P2P transaction scale come from the “2018 China Third-Party Payment Industry Report” published by i-Research.

4.2. Variable Selection and Description

4.2.1. Explained Variables

(1) *Return on Total Assets (ROA)*. This index is selected according to the core index of DuPont analysis system. The

indicator of the total asset return rate reflects the profitability of an enterprise more comprehensively. The higher the return on total assets shows, the higher the efficiency of the bank's asset utilization and the ability to represent the profitability indicators of commercial banks.

(2) *Proportion of Noninterest Income (NIIR)*. The profitability of a commercial bank is not only reflected in how much profit it generates from the use of assets, but also in its income structure. With the deepening of interest rate marketization, bank interest margin income has gradually decreased. In order to maintain long-term stable development, we must reduce dependence on traditional interest income and explore other business income, and noninterest income variables reflect whether the bank's business structure needs to be optimized or not.

4.2.2. *Explanatory Variables. Net interest margin (NIS)*. Net interest Spread = yield of interest-earning assets - interest-bearing debt interest rate = (interest income/interest-earning assets) - (interest expense/interest-earning assets). The most concerned financial indicators are banks' interest income, interest expenditure, and related assets. Net interest spread is the difference between the average loan interest rate and the average deposit interest rate of the bank, so it is more accurate when examining the efficiency of the bank.

4.2.3. Control Variables

(1) *Asset Size (SIZE)*. Select the natural logarithm of the bank's total assets as a variable to measure the scale of assets and test whether the profitability of larger banks in the interest rate marketization process is better than that of small-scale commercial banks.

(2) *Deposit-Loan Ratio (LD)*. The greater the value of the loan-to-deposit ratio, the greater the profitability of commercial banks. However, if the ratio exceeds a certain range, it will lead to operational risks for the bank, but it will have a negative effect on profitability. Especially in the context of interest rate liberalization, the decline in loan interest rates has a greater impact on loan interest income, causing banks to deviate from expected returns.

(3) *Cost-to-Income Ratio (CIR)*. The cost-to-income ratio is the ratio of operating expenses to operating income of a commercial bank. The lower the ratio, the better the effect of cost and cost control, which in turn means that the bank's operating efficiency is higher and the bank's profitability is higher.

(4) *NPL Ratio (RNPL)*. The nonperforming loan ratio refers to the proportion of subordinated loans, suspicious loans, and loss loans of commercial banks in total loans. This ratio is used to reflect the operating risks of banks.

(5) *Capital Adequacy Ratio (CAR)*. Capital adequacy ratio is the ratio of commercial bank's capital to risk assets, which is

used to measure the capital it has to withstand risks such as credit risk, interest rate risk, and political risk in the course of its operations.

(6) *NPL Provision Ratio (NPL)*. The nonperforming loan provision ratio is the ratio of the loan loss provision actually accrued by the bank to the nonperforming loan. This ratio is an important indicator to measure whether the commercial bank can effectively offset the bad debts.

(7) *National Economic Development (GDP)*. The operating conditions of commercial banks will be affected by the different economic cycles they are in, and the GDP growth rate can well reflect the operation of China's macroeconomy.

(8) *Bank Deposit Contribution Rate (BDR) and Securitization Ratio (SR)*. They reflect the importance of the banking industry in the entire economy. For instance, a high ratio of bank deposits to GDP may reflect a high demand for banking services. The securitization rate is the ratio of the total market value of stocks to GDP, which measures the scale of the stock market in the entire economy and reflects the depth of financial market development.

(9) *Third-Party Payment (TPP) and P2P Transactions*. The securities market is a direct financing market, and there are competing and complementary relationships with commercial banks. The emergence of third-party payments and P2P transactions has weakened the financial intermediary role of commercial banks, and funds have increasingly flowed to other channels, which has hit commercial banks' sources of funds and interest income.

Variable type and description are shown in Table 1.

4.2.4. *Model Building*. When exploring the determinants of the profitability of commercial banks, refer to the method of Asli Demirgüç-Kunt and Harry Huizinga [27], construct linear equations, and study the impact of various variables such as the bank's own characteristics and external environmental characteristics on the bank's income and return rate.

The macroeconomic operating environment and financial market influencing factors of various commercial banks are basically similar at different time periods. Therefore, the macroeconomic operating indicators and financial market influencing factors were not taken into account the personality characteristics of banks. Based on this, this article draws on the following model:

$$y_{it} = \alpha_0 + \alpha_i A_{it} + \beta_t B_t + \gamma_t C_t + \mu_{it}. \quad (1)$$

In equation (1), i represents different sample banks and t represents the period. The interpreted variable y_{it} represents the profitability index of i bank in period t . The explanatory variable A_{it} represents the internal influencing factors of i bank in period t , B_t represents the index of the external macroeconomic operation of the commercial bank in period t , and C_t represents the external financial market influencing factors of commercial bank in period t ; Further, α_0 is a

TABLE 1: Research variables and their meanings.

Variable type	Symbol	Name	Formula	Meaning
Explained variable	ROA	Return on total assets	Net profit/average total assets	Earnings performance
	NIIR	Proportion of noninterest income	Noninterest income/operating income	Profit structure
Explanatory variables	NIS	Net interest spread	Interest income/interest-earning assets-interest expense/interest-paying assets	Interest rate liberalization
Control variable	SIZE	Asset size	LN (total assets)	Commercial bank size
	LD	Loan deposit ratio	Loan amount/deposit amount	Loan proportion
	CIR	Cost-to-income ratio	Operating expenses/operating income	Costs and expenses
	RNPL	NPL ratio	Nonperforming loans/total loans	Costs and expenses
	CAR	Capital adequacy ratio	Total capital/risk assets	Risk prevention
	NPL	NPL provision ratio	Loan loss provision/nonperforming loan	Risk prevention
	GDP	National economic development	GDP annual growth rate	Macroeconomic conditions
	BDR	Bank deposit contribution rate	Total deposits in the banking sector/GDP	Financial industry development
	SR	Securitization rate	Stock market value/GDP	Financial industry development
	TPP	Third-party payment	Third-party payment scale/banking assets	I-finance
P2P	P2P transaction	P2P transaction volume/banking loan balance	I-finance	

constant, and α_i , β_t , and γ_t are coefficients. μ_{it} is a random error term.

According to the research focus and hypothesis of this article, some corrections were made to the above equation. Since this article only considers the domestic macro-economy, deletes the relevant variables of international trade and international capital markets, and combines the correlation between indicators and the availability of data, this article constructs the following linear regression equation model.

Since the total operating income of commercial banks mainly comes from interest spread income and noninterest income, the profit formula of commercial banks can be expressed as follows:

$$NP = IR + NIR - C. \quad (2)$$

Among them, NP stands for net profit, IR stands for Interest spread income, NIR stands for noninterest income, and C stands for related costs. The interest income can be further expressed as follows:

$$IR = BL \times (1 - NPL) \times LR - BD \times DR. \quad (3)$$

Among them, BL, NPL, LR, BD, and DR represent loan amount, nonperforming loan rate, loan interest rate, deposit amount, and deposit interest rate. Further, if the loan interest rate is the sum of the deposit interest rate and the deposit-loan spread, the spread income can also be expressed as follows:

$$IR = BL \times (1 - NPL) \times (DR + NIS) - BD \times DR. \quad (4)$$

Accordingly, the net profit can be expressed as follows:

$$NP = BL \times (1 - NPL) \times (DR + NIS) - BD \times DR + NIR - C, \quad (5)$$

because

$$ROA = \frac{NP}{SIZE}. \quad (6)$$

SIZE represents total assets, then,

$$\begin{aligned} ROA &= \frac{BL \times (1 - NPL) \times (DR + NIS) - BD \times DR + NIR - C}{SIZE}, \\ ROA &= \frac{BL \times (1 - NPL) \times (DR + NIS) - BD \times DR + NIR - C}{SIZE} \\ &= \frac{NIS \times BL \times (1 - NPL) + BL \times DR \times (1 - NPL) - BD \times DR - C}{SIZE}. \end{aligned} \quad (7)$$

Therefore, there is a positive correlation between the deposit-loan spread NIS and the ROA level of total assets.

Similarly, the expression of noninterest income ratio (NIIR) regarding the deposit-loan spread (NIS) is

$$\begin{aligned} \text{NIIR} &= \frac{\text{NIR}}{\text{TR}} = \frac{\text{NP} - \text{IR} + \text{C}}{\text{TR}} = \frac{\text{NP} - \text{BL} \times (1 - \text{NPL}) \times (\text{DR} + \text{NIS}) - \text{BD} \times \text{DR} + \text{C}}{\text{TR}} \\ &= \frac{-\text{NIS} \times \text{BL} \times (1 - \text{NPL}) + \text{NP} - \text{BL} \times (1 - \text{NPL}) \times \text{DR} - \text{BD} \times \text{DR} + \text{C}}{\text{TR}}, \end{aligned} \quad (8)$$

where TR represents total income, because $\text{BL} > 0$, $(1 - \text{NPL}) > 0$. That is, there is a negative correlation between the deposit-loan spread NIS and the proportion of noninterest income NIIR.

Model one is

$$\begin{aligned} \text{ROA}_{it} &= \alpha_0 + \alpha_1 \text{NIS}_{it} + \alpha_2 \text{SIZE}_{it} + \alpha_3 \text{LD}_{it} + \alpha_4 \text{CIR}_{it} \\ &\quad + \alpha_5 \text{RNPL}_{it} + \alpha_6 \text{CAR}_{it} + \alpha_7 \text{NPL}_{it} + \beta_1 \text{GDP}_t \\ &\quad + \beta_2 \text{BDR}_t + \gamma_1 \text{SR}_t + \gamma_2 \text{TPP}_t + \gamma_3 \text{P2P}_t + \varepsilon_{it}. \end{aligned} \quad (9)$$

Model two is

$$\begin{aligned} \text{NIIR}_{it} &= \alpha_0 + \alpha_1 \text{NIS}_{it} + \alpha_2 \text{SIZE}_{it} + \alpha_3 \text{LD}_{it} + \alpha_4 \text{CIR}_{it} \\ &\quad + \alpha_5 \text{RNPL}_{it} + \alpha_6 \text{CAR}_{it} + \alpha_7 \text{NPL}_{it} + \beta_1 \text{GDP}_t \\ &\quad + \beta_2 \text{BDR}_t + \gamma_1 \text{SR}_t + \gamma_2 \text{TPP}_t + \gamma_3 \text{P2P}_t + \mu_{it}. \end{aligned} \quad (10)$$

Similarly, i represents the number of sections and t represents the observation period of the section; ROA_{it} is i -commercial bank's total asset return rate in the year, which measures the profit performance of commercial banks. NIIR_{it} represents the proportion of the noninterest income of i -commercial bank in year t and measures its profit structure. NIS_{it} represents the net interest margin of i -commercial bank at time t in order to affect the profitability of commercial banks' explanatory variables. In terms of control variables, SIZE_{it} , LD_{it} , CIR_{it} , RNPL_{it} , CAR_{it} , and NPL_{it} are the indicators within the bank that may affect profitability. GDP_t , BDR_t , SR_t , TPP_t , and P2P_t represent the external environmental impact factors of commercial banks in period t . μ_{it} , ε_{it} are the random error terms of the two models, respectively.

5. Empirical Analysis

5.1. Hausman Test and Correlation Analysis. Empirical research uses panel data. Before regression, we should first figure out whether to choose a fixed-effect model or a random-effect model. This article conducts the Hausman test. The results show that both models should use fixed-effect models. In addition, in order to ensure that there is no serious multicollinearity between the explanatory variables and the control variables, correlation coefficient analysis was performed, and the results showed that the correlation coefficients between the variables were all less than 0.7, basically within a reasonable range. The control variables

selected in this paper are more reasonable and can be used for relevant empirical research.

5.2. Full Sample Regression Results. Using measurement software, the fixed-effect model is used to perform regression on the explanatory variables. The regression results of the two models are shown in Table 2.

On the whole, in the empirical results in Table 2, the R^2 of the two models is 0.64 and 0.78, respectively. The higher the goodness of fitting, the better the fitting effect of the two models. The independent variables can have an effect on the dependent variables. Moreover, both models passed the F test, indicating that the two models as a whole are significant. In both models, the net interest spread (NIS) passed the test at the 1% level, indicating that in the context of interest rate liberalization, the change in spreads can indeed have a significant impact on the profitability of commercial banks. The impact of the profit structure is different. From the empirical results of Model 1, the net interest margin (NIS) has passed the test significantly and is positively correlated with the ROA. The variable coefficient is 0.1704; that is, the larger the net interest spread, the higher the total asset return rate. The relationship between the net interest spread and the return on total assets is consistent with the research assumptions. In model two, the net interest margin (NIS) passed the test at a level of 1%, and the coefficient was -8.4801, which was negatively correlated with the profit structure indicator NIIR. That is, as the net interest margin narrowed, the proportion of noninterest income increased. This result means that it is influenced by the interest rate marketization policy. That is, the relationship between net interest margin and nonincome ratio is consistent with the research hypothesis.

For control variables, in model one, the size of bank assets (SIZE) is significant at the level of 1%, and it is positively correlated with the explanatory variable (ROA), indicating that the larger commercial banks are, the more they are capable of improving profitability. The cost-to-income ratio (CIR) is significant at the 5% level, so it is negatively correlated with the return on total assets. The nonperforming loan ratio (RNPL) passes the test at a level of 10% and has a negative correlation with the return on total assets. Reducing the nonperforming loan ratio helps commercial banks to improve their profitability. Capital adequacy ratio (CAR) has passed the test significantly, and the coefficient is positive, that is, banks with sufficient capital have stronger ability to resist risks and have higher social credibility, which is conducive to the development of businesses such as deposits and loans. Higher levels of profitability are also easier to achieve. The NPL variable has not passed the test, and the explanation for the change in the

TABLE 2: Commercial bank linear regression results.

Variable	Model one (ROA)	Model two (NIIR)
NIS	0.1704*** (4.7434)	-8.4801*** (-7.3092)
SIZE	0.1075*** (3.0811)	5.3488*** (3.6392)
LD	-0.0003 (-0.2335)	0.2621*** (5.4875)
CIR	-0.0051* (-1.8174)	-0.3077*** (-3.1732)
RNPL	-0.0574* (-1.8489)	2.8694*** (2.8715)
CAR	0.0449*** (7.3378)	0.4278** (2.1741)
NPL	0.0030 (0.1675)	0.5892 (1.0202)
GDP	0.0038 (1.1434)	-0.1480 (-1.4135)
BDR	1.6426** (2.2027)	-6.1948 (-0.2544)
SR	-0.3846*** (-3.3060)	0.4081 (0.1113)
TPP	-0.0012*** (-5.8383)	-0.0035 (-0.5311)
P2P	-0.0021 (-1.3860)	-0.0447 (-0.9596)
R ²	0.6438	0.7826
F	20.1448	39.9527

The symbols ***, **, and * indicate significant at 1%, 5%, and 10% levels.

return on total assets in this model is limited. The economic growth rate (GDP) is positively correlated with the return on total assets; that is, the better the macroeconomic operates, the more beneficial it is for commercial banks to obtain profits. The indicator of the BDR is significant at the level of 5%, and the coefficient is positive, indicating that the higher the ratio of total deposits in the banking sector to GDP, the higher the profitability of commercial banks. In other words, the overall rapid development of the banking industry has created a good environment for commercial banks to increase profits. The securitization rate (SR) has passed the significance test. This indicator is also used to measure the development of the financial market. However, unlike the deposit rate, this indicator is the ratio of the total market value of stocks to GDP. The higher the ratio, the greater the share of direct financing in the financial market. It is a market that competes with indirect financing of commercial banks, which has an adverse effect on the profitability of commercial banks. Internet financial development indicators the third-party payment indicator (TPP) is significant at the level of 1% and has a negative correlation with the total asset return of commercial banks. This means that the development of third-party payment platforms has weakened the status of commercial banks as financial intermediaries, diverted some customers, and has a negative impact on the bank's total asset return. P2P transaction scale (P2P), another indicator to measure the development of Internet finance, has not passed the significance test. The empirical results of model two: the size of bank assets (SIZE) has passed the significance test and has a positive correlation

with the variable NIIR, indicating that larger commercial banks are more capable of optimizing their profit structure. The variable loan deposit ratio (LD) coefficient is positive and passes the test at the 1% level; that is, the greater the proportion of loans, the more favorable the improvement of the profit structure of commercial banks. The cost-to-income ratio (CIR) is significant at the level of 1% and has a negative correlation with the proportion of noninterest income, indicating that the lower the cost, the more beneficial it is to expand the noninterest income business and improve the profit structure of commercial banks. The nonperforming loan ratio (RNPL) has passed the significance test and has a positive correlation with the proportion of noninterest income. However, this article believes that the effect of the nonperforming rate on the proportion of noninterest income is due to the loss of loan interest. This increase in the proportion of noninterest income is indeed a dangerous signal, not an improvement in the income structure. The capital adequacy ratio (CAR) is significant at the level of 5%, and the coefficient is 0.4278. That is, a bank with sufficient capital has more human and financial resources to innovate its new business area, expands its intermediate business with a higher social reputation, and thus obtains a higher level of noninterest income, which has a positive role in promoting the improvement of the profit structure. The NPL ratio is also an indicator to measure the risk resistance of commercial banks, but this variable has not passed the significance test, so the interpretation in this model is limited. Economic growth rate (GDP), Bank deposit contribution rate (BDR), and P2P transaction scale have a negative correlation with the proportion of noninterest income. The securitization rate (SR) and third-party payment scale (TPP) are positively correlated with the proportion of noninterest income. However, none of these macroeconomic variables, financial market variables, and Internet financial development variables beyond the bank's own scope have passed the test at a 10% significance level, so the impact is not significant.

5.3. Regression Results of Different Size Samples. From the regression results in the previous section, we have drawn the basic conclusion that there is a positive correlation between asset size and the profit level and profit structure of commercial banks. Then, whether the impact of commercial banks of different sizes in the marketization of interest rates is different will be the focus of the following research. This article divides sixteen commercial banks into three groups of large state-owned commercial banks, medium-sized joint-stock commercial banks, and small urban commercial banks based on the average assets of banks from 2008 to 2018 and performs linear regression on the panel data of the three groups of samples, respectively. Model 1 regression results are shown in Table 3.

In the empirical process of classifying by asset size, the net interest margin (NIS) and the commercial bank's return on total assets still show a positive correlation. Both large state-owned commercial banks and joint-stock commercial banks passed the test significantly. The third group of small

TABLE 3: Classification sample model-regression results.

Variables	Large state-owned banks	Medium-sized joint-stock banks	Small city banks
NIS	0.2087** (2.6952)	0.1678*** (3.2273)	0.0075 (0.1155)
SIZE	0.0854 (1.0322)	0.1893** (2.2626)	0.0032 (0.0452)
LD	-0.0036* (-1.7169)	-0.0012 (-0.3561)	0.0011 (0.4351)
CIR	-0.0259*** (-4.3733)	-0.0104*** (-3.1809)	0.0136** (2.5885)
RNPL	-0.0951* (-1.7764)	-0.1233* (-1.7216)	0.3292* (2.052)
CAR	0.0023 (0.1321)	0.0222 (1.3415)	0.0323*** (3.4033)
NPL	-0.0102 (-0.2156)	0.0747** (2.128)	0.0713 (1.6115)
GDP	0.0035 (0.6871)	-0.0139** (-2.1673)	-0.0014 (-0.268)
BDR	1.1916 (0.8737)	-2.0148 (-1.5869)	-0.1032 (-0.0945)
SR	-0.183 (-0.9344)	0.1674 (0.8051)	-0.3476** (-2.1596)
TPP	-0.0026*** (-3.9951)	0.0001 (0.1934)	-0.0014*** (-4.2494)
P2P	-0.0022 (-0.5087)	-0.0009 (-0.3304)	-0.0013 (-0.5592)
R ²	0.8648	0.5889	0.8629
F	16.5931	12.1856	31.5443

The symbols ***, **, and * indicate a significant level at 1%, 5%, and 10%.

city commercial banks was not significant. This may be affected by the small sample size, but it can still prove that the narrowing of interest rate spreads in the process of interest rate liberalization. Profitability has a negative impact. After comparing the coefficient, when the net interest margin changes, the large state-owned commercial banks have the biggest impact on the yield. This may be due to the excessive scale of state-owned commercial banks, lower operating efficiency, and greater difficulty in innovation and transformation, and the main source of income is relatively more dependent on interest income. At the same time, the medium-sized joint-stock commercial banks have strong innovation capabilities and can innovate products and services according to market demand in a timely manner, reducing the impact of narrowing spreads.

In the classification regression results of Model 2 (see Table 4), the net interest margin (NIS) still passed the significance test, and the coefficient is negative, which is negatively correlated with the NIIR. It means that due to the interest rate marketization policy, as the net interest margin narrows, the proportion of the noninterest income of commercial banks gradually increases, and the income structure is gradually optimized. And the comparison found that the coefficient of NIS in the regression model of large state-owned banks is significantly larger than that of small and medium-sized bank models. This obvious difference indicates that the proportion of the noninterest income of large banks is more affected. On the one hand, due to the relatively large deposits and loans of state-owned

TABLE 4: Classification sample model two regression results.

Variables	Large state-owned bank	Medium-sized joint-stock banks	Small city bank
NIS	-13.5588*** (-5.0734)	-3.0516* (-1.7126)	-3.1769 (-1.0407)
SIZE	2.2315 (0.7817)	6.727** (2.3465)	-4.2566 (-1.2982)
LD	0.3563*** (4.9041)	-0.0611 (-0.5344)	0.2572** (2.1749)
CIR	0.3669* (1.7929)	-0.2093* (-1.8654)	0.8293*** (3.3456)
RNPL	-2.8703 (-1.5541)	-1.8772 (-0.7649)	15.7069** (2.0811)
CAR	1.3169** (2.1616)	-0.5538 (-0.9755)	0.3452 (0.7720)
NPL	-2.9247* (-1.7831)	-1.1796 (-0.9807)	6.6129*** (3.1785)
GDP	-0.0214 (-0.1229)	-0.5243** (-2.3903)	-0.3699 (-1.5030)
BDR	-10.6244 (-0.2257)	-46.6835 (-1.0731)	-12.1935 (-0.2373)
SR	-10.6243 (-1.5719)	8.9122 (1.2508)	10.8347 (1.4308)
TPP	-0.017 (-1.5509)	0.0707*** (4.3983)	0.006 (0.4012)
P2P	-0.0313* (-1.8410)	-0.0142 (-0.4511)	-0.0072 (-0.0647)
R ²	0.8526	0.7302	0.7363
F	37.9882	64.7370	10.9137

The symbols ***, **, and * indicate a significant level at 1%, 5%, and 10%.

commercial banks, the degree of dependence on loan interest income is high, resulting in a significant reduction in interest income when interest margins narrow. On the other hand, like other banks, market competition forces banks to actively develop intermediate businesses and optimize their own income structure. The profit structure of small and medium-sized city commercial banks is not significantly affected by the interest rate marketization policy: on the one hand, small and medium-sized banks have been established for a relatively short period of time and established new businesses and innovative products as market competition strategies at the beginning level; on the other hand, small and medium-sized banks, especially small city commercial banks, are mainly market-oriented in their local cities, and their profitability depends largely on local economic development.

5.3.1. Robust Test. In order to test the reliability of the above model results, the method of replacing the explained variables was used to conduct the robustness test, and the ROE and ROA were replaced by the ROE and IH Profit structure indicator noninterest income ratio (NIIR). The rate of return on net assets reflects the level of returns of shareholders' equity, and it can also reflect the level of corporate profits to a larger extent. The Herfindahl Index can be used to measure the degree of corporate diversification. In the study of the income structure of commercial banks, the index combines interest income and commission income to reflect the degree

of diversification of the income structure. The results of the robustness test reprove the robustness of the above models:

$$\begin{aligned} \text{Hefindar Index IH} = & 1 - (\text{proportion of net interest income})^2 \\ & - (\text{commission fee})^2. \end{aligned} \quad (11)$$

6. Conclusions and Recommendations

There is a significant positive correlation between the net interest margin and the return on the total assets of commercial banks. With the deepening of marketization, banks' original profit model that mainly relied on deposit and loan spreads to obtain profits has been hit; the main source of profit has been greatly reduced, the return on total assets has fallen, and the narrowing of net interest margins has also had an important impact on the profit structure of commercial banks. The proportion of noninterest income as an important indicator reflecting the profit source structure of commercial banks has gradually increased in the process of interest rate liberalization. The proportion of the noninterest income of commercial banks is increasing since it is affected by the continuous narrowing of net interest margins. Commercial banks should change their business structure and reduce their dependence on income from spreads; facing the challenges and opportunities of marketization, commercial banks should develop noninterest business and increase the proportion of noninterest income actively.

After comparing different sizes of commercial banks, the empirical results show the effect of independent variable changes on the differentiation of dependent variable changes. The net interest spread has a greater impact on large state-owned commercial banks when judged from the absolute value of the coefficient. Large state-owned banks have a larger scale and higher market share in the deposit and loan market. In the process of marketization of interest rates, the narrowing of net interest margins has had a greater impact on these banks. The nationwide joint-stock commercial bank originated from the stage of market competition, with more flexible institutions, larger asset scale, and consciousness and sufficient ability to pursue innovation to respond to market changes. Although urban commercial banks are small in scale and lack innovation capacity since they mainly serve the development of the regional economy, their profit performance depends largely on the regional economy, and the impact of net interest margins is relatively small.

Main recommendations: First, improve the interest rate pricing system and cultivate independent pricing capabilities. Commercial banks should continue to improve interest rate pricing capability as their core decision. In the future, the volatility of market interest rates may be greater. Commercial banks should improve their ability to avoid, diversify, and compensate for interest rate risks to ensure the stability of bank operations and ensure the stability of bank operations. Therefore, commercial banks need to accelerate the improvement of the pricing system and improve the banks' independent pricing capabilities. Second, they need to enhance cost management capabilities and increase

tolerance for narrowing spreads. The traditional model of relying on spread income is not sustainable. In the face of the decline in income growth caused by the narrowing of deposit and loan spreads, commercial banks must refine their extensive operations. Stabilize and expand asset sources and control capital costs; strengthen expense management, control operating costs throughout the process, and increase tolerance for narrowing spreads. Third, they should promote the transformation of integrated service models and improve the income structure. Commercial banks are facing more and more competition from cross-border competition, and accelerating the adjustment of profit structure is the focus of the bank's future transformation and development. We should pay attention to and carry out business and product innovation, integrate the traditional business of commercial banks into securities, funds, insurance, trust, leasing, and other fields, provide customers with new comprehensive financial services through market cross-border and mixed operation, and use technology to build a financial consumer ecosystem and then improve the bank's income structure. Fourth, in order to achieve differentiated operations, banks need to base on their own advantages. For large state-owned banks, they can make full use of their scale advantages and brand advantages, based on traditional business, cross-sell products to existing customers, and provide customers with comprehensive financial services. As for large state-owned commercial banks, the national joint-stock commercial banks have stronger innovation capabilities, are more concerned about market changes, and are more sensitive. They should maintain their advantages, adjust business layout and product characteristics in a timely manner according to market changes, and take the lead in entering the credit blue ocean market. For city commercial banks, their advantage lies in their regionality. The localization strategy is more efficient than the expansion strategy. They should be based on the local area and rely on government support to improve the adhesion of the customer group in the region. Small-scale banks such as city commercial banks should also actively introduce strategic investors, expand the scale of banks, strengthen risk management, improve corporate governance, and enhance the stability of bank operations.

Several other topics remain for further study. The innovation of the commercial bank's intermediary business model is an important direction of research. Further research on the application of modern technologies such as big data, cloud computing, artificial intelligence in the commercial bank's intermediary business development will further enhance the commercial bank's intermediary business capabilities.

Data Availability

All data included in this study are available from the corresponding author upon request. Financial data such as income, cost, and capital structure of commercial banks are collected from Bank-scope database. Economic data such as GDP, total deposits in the banking sector, and the market value of the stock market are collected from the website of

the National Bureau of Statistics and the EPS database. Internet financial data such as third-party payment scale and P2P transaction scale come from the “2018 China Third-Party Payment Industry Report” published by *i-Research*.

Conflicts of Interest

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

References

- [1] T. S. Y. Ho and A. Saunders, “The determinants of bank interest margins: theory and empirical evidence,” *The Journal of Financial and Quantitative Analysis*, vol. 16, no. 4, pp. 581–600, 1981.
- [2] Angbazo, “LTI Commercial bank net interest margins, default risk, interest-rate risk, and off-balance sheet Banking,” *Journal of Banking & Finance*, vol. 1, pp. 43–55, 1997.
- [3] R. Rodríguez, “The determinants of bank margins in European Banking,” *Journal of Banking and Finance*, vol. 31, pp. 2043–2063, 2007.
- [4] V. V. Claeyns, “Determinants of bank interest margins in central and eastern Europe: a comparison with the west,” *Economic Systems*, vol. 32, pp. 197–216, 2008.
- [5] J. Maudos and L. Solís, “The determinants of net interest income in the Mexican Banking system: an integrated model,” *Journal of Banking and Finance*, vol. 33, pp. 1920–1931, 2009.
- [6] S. Islam and N. Shin-Ichi, “The determinants of bank net interest margins: a panel evidence from South Asian countries,” *Research in International Business and Finance*, vol. 37, pp. 501–514, 2016.
- [7] A. Saunders and L. Schumacher, “The determinants of bank interest rate margins: an International study,” *Journal of International Money and Finance*, vol. 19, no. 6, pp. 813–832, 2000.
- [8] S. Claessens, N. Coleman, and M. Donnelly, ““Low-For-Long” interest rates and banks’ interest margins and profitability: cross-country evidence,” *Journal of Financial Intermediation*, vol. 35, no. 35, pp. 1–16, 2018.
- [9] R. A. Dwumfour, “Explaining banking spread,” *Journal of Financial Economic Policy*, vol. 11, pp. 139–156, 2019.
- [10] S. N. Brissimis, M. D. Delis, and N. I. Papanikolaou, “Exploring the nexus between Banking sector reform and performance: evidence from newly acceded EU countries,” *Journal of Banking & Finance*, vol. 32, pp. 2674–2683, 2008.
- [11] J. Shim, “Bank capital buffer and portfolio risk: the influence of business cycle and revenue diversification,” *Journal of Banking & Finance*, vol. 37, no. 3, pp. 761–772, 2013.
- [12] H. Zhou, “Research on the influencing factors of net interest margin of Chinese commercial banks—based on the empirical evidence from 1999 to 2006,” *Financial Research*, vol. 4, pp. 69–84, 2008.
- [13] H. Wang and J. Guo, “Market-oriented interest rate reform, noninterest income and bank performance,” *Financial Forum*, vol. 19, no. 8, pp. 3–12, 2014.
- [14] K. Zhou, T. Li, and X. He, “What determines the net interest margins of Chinese commercial banks?” *Economic Research*, vol. 8, pp. 65–76, 2008.
- [15] Y. Zhang and Z. Zhang, “Affecting factors of net interest margin of Chinese Banking,” *Financial Forum*, vol. 15, no. 6, pp. 22–27, 2010.
- [16] C. Sui and T. Xing, “Research on the loan pricing of Chinese commercial banks based on non-complete interest rate marketization,” *International Finance Research*, vol. 12, pp. 82–93, 2013.
- [17] N. Ding, “Analysis of the economic impact of deposit and loan spreads in the Chinese banking industry,” *Macroeconomic Research*, vol. 12, pp. 64–73, 2013.
- [18] M. Yin, K. Xu, and Y. Tai, “Analysis of influencing factors in the banking interest spread on the basis of economic liberalization,” *Investment Research*, vol. 4, pp. 60–71, 2013.
- [19] X. Xiao and Y. Wu, “The impact of America’s market-based interest rate reform on the banking,” *International Finance Research*, vol. 1, pp. 69–75, 2011.
- [20] S. Ba, Z. Hua, and Y. Zhu, “International comparison: path, performance and market structure,” *Journal of Central China Normal University (Humanities and Social Sciences Edition)*, vol. 5, pp. 33–46, 2012.
- [21] J. Zhang and P. Wang, “Research on the frontier efficiency of China’s banking industry and its influencing factors: a distance function model based on stochastic frontier,” *Financial Research*, vol. 12, pp. 1–18, 2009.
- [22] M. Liu, X. Zhang, and C. Zhang, “Research on the correlation of Chinese commercial banks’ diversification, performance and risk,” *International Finance Research*, vol. 8, pp. 59–69, 2012.
- [23] M. Wang and X. Liu, “Research on the impact of income structure of commercial banks on profitability: analysis based on panel data of 14 listed banks in China,” *Nankai Management Review*, vol. 2, pp. 143–149, 2013.
- [24] X. Niu and X. Qiu, “Interest rate and bank risk-taking: an empirical study based on China’s listed banks,” *Financial Research*, vol. 4, pp. 15–28, 2013.
- [25] S. Liu, “The savings and loan crisis and the implication on China’s market-based interest rate reform,” *International Finance Research*, vol. 4, pp. 13–21, 2013.
- [26] H. Zhu, L. Zhang, and N. Wang, “The external determinants of commercial banks’ performance in China,” *Management Review*, vol. 10, pp. 3–12, 2014.
- [27] A. Demirgüç-Kunt and H. Huizinga, “Determinants of commercial bank interest margins and profitability: some international evidence,” *The World Bank Economic Review*, vol. 13, no. 2, pp. 379–408, 1999.

Research Article

Evolutionary Game Models on Multiagent Collaborative Mechanism in Responsible Innovation

Kun Yang, Wan Wang , and Bin Hu

School of Management Studies, Shanghai University of Engineering Science, Shanghai 201620, China

Correspondence should be addressed to Wan Wang; m030218140@sues.edu.cn

Received 11 May 2020; Revised 29 May 2020; Accepted 9 June 2020; Published 1 August 2020

Academic Editor: Lu Zhen

Copyright © 2020 Kun Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Innovation is a game process; in particular, the behavior among multiple agents in responsible innovation is susceptible to the influence of benefits, risks, responsibilities, and other factors, resulting in unstable collaborative relationships. Therefore, this paper constructs a tripartite evolutionary game model including the government, enterprises, and the public, combined with system dynamics modeling to simulate and analyze the tripartite behavior strategy and sensitivity to relevant exogenous variables. The study shows that the tripartite game eventually converges to a stable state of the government active supervision, enterprises making responsible innovation, and the public's positive participation. The positive participation of the public drives rapidly the game to a steady state, while the behavioral strategies of enterprises are more susceptible to the behavior of the government. Supervision cost, penalty amount, and value compensation are the most critical factors influencing the change of the corresponding agents' behavior strategy, and the final strategic stability of tripartite is affected by multiple exogenous variables.

1. Introduction

The world is facing enormous societal challenges such as climate change, food safety, society security, energy demand, demographic change, and well-being. These enormous challenges transcend national borders and have an impact on a large number of people, cities, and the entire planet [1]. Therefore, governments around the world are emphasizing the importance of innovation, because innovation and technological development are generally regarded as a panacea for grand societal challenges [2]. However, innovation is generally considered to be inherently good, but always has the probability of having unforeseen consequences [3]. In the short term, innovation may have certain advantages, but from a long-term perspective, innovation will face questions, dilemmas, and uncertainties in its future development. Today, many researchers agree that even the most promising innovations can fail because the ethical and societal concerns they bring are not properly taken into account [4]. In this context, Responsible Research and Innovation (RRI) is seen as a way of governance innovation development to address challenges such as population

ageing, poverty, inequality, and the availability of high-quality healthcare services [5, 6]. RRI is a recent expression used by the European Union to indicate the part of its research and innovation (R&I) strategy and is highly concerned by the realm of politics and academia, aiming to make the process and results of R&I ethically acceptable and socially desirable [7]. RRI can add new elements to innovation governance by making R&I participants jointly responsible for societal embedding and potential impact [8].

RRI is usually implemented from the perspective of policy or social ethics and focuses on the academic research and development environment [6], while responsible innovation (RI) pays more attention to the innovation process itself and has become a concept close to RRI [9]. For the research purpose of this paper, this research will use the term of RI. RI research revolves around whether and how to lead technology and innovation to socially desired goals [10] and manage innovation through early “upstream” interventions rather than post-event “downstream” monitoring and “corrective” interventions [11]. Its core is the process of open research and innovation, which incorporates “new voices of science and innovation governance” [12]; on the one hand, it

aims to improve the legalization and democratization of the innovation process and output, and on the other hand, it attempts to increase the diversity of views to meet the needs of social actors [13]. To meet these needs, RI focuses on the participation and collaboration of stakeholders [14]; they need to continue to participate, not only in consultation, but also to provide information, power, and opportunities to play a role in decision-making to achieve what is called “mutually beneficial interaction” [13]. Meanwhile, this inclusiveness of upstream stakeholders and the public can help achieve collective responsibility to control and guide innovation so that it is ethically acceptable, societally desirable, and sustainable [15, 16]. However, the current view of innovation in the RI literature tends to be narrow [9], and the research of RI should more adequately consider that innovation is an endogenous process resulting from the collaborative actions of interdependent heterogeneous agents in a complex system, where the outcomes are characterized by essential uncertainty [17, 18]. At the same time, implementing RI in the business context also faces a series of important challenges. For example, first of all, focusing solely on science and technological development without considering other types of innovation can produce a narrow view on innovation [9]. Second, enterprises prioritize the economic impact achieved by innovation and focus more on commercially driven innovation processes [16]. Third, different stakeholders have various values and interests in the business context, and innovators face different constraints in terms of confidentiality and public image [4]. Furthermore, innovation is a process of the multiagent game [19]; when different stakeholders collaboratively engage in RI, the rationality of each agent is constrained by the limitations of available information, cognition, and decision-making time, which makes the collaborative relationship between multiple agents into a long-term dynamic game relationship where each agent continuously changes its strategy through imitation and learning to maximize its own benefits.

The abovementioned analysis can lead to the following questions:

How to establish a tripartite evolutionary game model for multiagent collaboration in RI?

How does the behavior of multiagent interact with each other in RI?

How to make effective collaboration mechanisms to promote multiagent cooperation and responsible innovation governance?

This paper first constructs a tripartite evolutionary game and analyzes the stability of each player’s game strategy in the model. Then, a tripartite game system dynamics (SD) model is established by introducing SD, and their interaction relationship is studied by using simulation. Through theoretical analysis and extensive simulation, it is hoped that the three questions posed above can be answered.

This paper is organized as follows. Section 2 presents a literature review on the multiagent in RI and the tripartite evolutionary game. Section 3 develops the tripartite evolutionary game model and analyzes the equilibrium point

and multiagent asymptotic stability. Section 4 advances numerical simulation analysis based on SD to illustrate the interaction relationship. Finally, conclusions and limitations are given in Section 5.

2. Literature Review

2.1. Multiagent in Responsible Innovation. RI is a new concept developed and introduced by researchers and policymakers in a top-down manner [20]. The main idea behind it is to democratize innovation [21], and to implement collaborative governance forms such as stakeholder and public participation. Incorporating different stakeholders and the public into RI, in turn, means increasing the possibilities of anticipating and discerning how research and innovation may benefit society and prevent any negative consequences from occurring [22]. This can increase the chance of innovation being adopted, better embed innovation into society, and ensure that innovation brings societal benefits. Thus, first and foremost, it is important to identify the stakeholders in RI and classify these stakeholders.

RI is a means of emphasizing multistakeholder partnerships to mobilize and share knowledge, technology, and financial resources and to encourage and promote simultaneously effective public and civil society partnerships [6]. The main stakeholders are crucial to the innovation process, and there is a lot of literature discussing different stakeholders related to or participating in RI. Von Schomberg pointed out that the main stakeholders in RI include European citizen, consumer, producers, civil society, policymakers, etc., whose scope of action is mainly concentrated on the application phase of new technologies, the innovation process, and European legislation [23]. Davies and Horst summarize the most of stakeholders mentioned in the web-based and academic aspects related to RI, such as academic institutions, key users, researchers, businesses and supply chain organizations, funders and regulators, consumers, affected parties, and the society at large, etc. [24]. Silva classifies stakeholders in RI as internal stakeholders and external stakeholders, where external stakeholders include individual researchers, research organizations, civil society actors, legislators, public bodies, etc. [6]. Blok et al. put forward the classification according to economic and noneconomic stakeholders; for example, economic stakeholders have employees and suppliers, and noneconomic stakeholders have NGOs and research institutes [25]. Mei and Chen divide stakeholders into policymakers (e.g., governments and policy institutions at all levels), experts in innovation activities (e.g., business innovation planners, innovation implementation organizations or institutions), and society actors (e.g., the public, innovative social participants, and potential service targets) [26]. Therefore, given the research purpose of this paper and for the convenience of analysis, the stakeholders are classified according to the research of Mei and Chen [26], and the government, enterprises, and the public are selected as agents from the three categories.

Innovation is a collaborative and game process of multiple parties. As a regulator, the government draws a legal red line that enterprises and their stakeholders cannot cross [27] and has the power to regulate innovation and influence the public's behavior [28]. The main purpose of enterprises is to pursue profits, and responsible activities of enterprise involvement will help it build a lasting positive image and good reputation [29]. Enterprises improve social outcomes by better coordinating with the government, the public, or other stakeholders and try to generate long-lasting "win-win" solutions [30]. Participatory planning is a way of involving the public in the decision-making process, and public participation is an important aspect of democracy and trust in the government, which helps improve the legitimacy of decisions [31]. From the perspective of evolutionary economics, the public as a consumer plays a vital role in responsible innovation and demand [18]. On the one hand, they can actively exert pressure on enterprises to conduct innovation in a responsible manner; on the other hand, they are a significant reference group for enterprises to better align their products and services with consumers' expectations and needs [32]. The three main agents of the government, enterprise, and the public constitute the most important stakeholder group, and enterprises produce and sell products resulting from innovation, but innovation will not succeed without public acceptance, market diffusion, and government supervision of innovation development.

2.2. Tripartite Evolutionary Game. A key assumption of traditional game theory is that participants are completely rational, and this rationality assumption is not consistent with the facts [33]. Evolutionary game theory is a theory that combines game theory analysis and dynamic evolution process analysis, which overcomes the hypothesis of complete rationality and complete information in a game model [34]. At present, evolutionary game theory has been introduced into a wide range of fields, especially the tripartite evolutionary game, which is suitable for analyzing the cooperative game behavior of multiple stakeholders. Liu et al. establish a multiple agent's evolutionary game model that includes carbon fiber production enterprises, carbon fiber application enterprises, and governments and analyze the influencing factors of collaborative innovation in the carbon fiber industry to alleviate the problem of industry chain disconnection and promote sustainable development of the carbon fiber industry [35]. Guo and Li construct a tripartite evolutionary game model between the government, private sector, and owners and combine SD to simulate the evolution process, which can be helpful to explore the collaboration mechanism of participants in the PPP model of the old community renovation project [36]. Based on the mixed development environment of cascade hydropower stations, Chen et al. establish a tripartite evolutionary game model, which aims to explore the directions and conditions for cooperative and noncooperative strategies evolving into steady states and promote the joint operation of cascade hydropower stations [37].

However, the current research on multiagent collaboration in RI mainly focuses on qualitative analysis, such as the cooperation of stakeholders in the port RI model [38, 39], RI of stakeholders in small producers clusters [40], and RI and stakeholders collaboration in infrastructures [41]. The literature shows that few researchers incorporate the agent's bounded rationality into RI and study the collaborative problem of multiagent in RI through quantitative analysis, while the evolutionary game is suitable for analyzing behavior change of stakeholders. Thus, it would be more practical and meaningful to closely explore the collaboration mechanism in RI with boundedly rational agents based on a tripartite evolutionary game.

3. Construction and Analysis of the Tripartite Evolutionary Game Model

3.1. Model Assumptions. In this paper, the evolutionary game method is used to analyze the collaborative mechanism of multiagent under RI, so the following model assumptions are proposed. Meanwhile, for the convenience of analysis, this research will select the government, enterprises, and the public as three types of representatives among policymakers, innovation activity experts, and society actors.

Assume that each agent in the game model has two possible strategic choices. The government's strategic choices will be active supervision (AS) and inactive supervision (IS). The possible strategic choices for enterprises are to make responsible innovation (MRI) and irresponsible innovation (NMRI), respectively. The public's strategic choices will be positive participation (PP) and negative participation (NP). The probability of the government choosing AS is x ($0 \leq x \leq 1$), the probability of enterprises choosing MRI is y ($0 \leq y \leq 1$), and the probability of the public choosing PP is z ($0 \leq z \leq 1$).

Assume that the relevant parameters related to the government are as follows. The government carries out active supervision and pays the supervision cost C_1 , including investing a lot of human, material, and financial resources. RI meets social values and needs and brings social value income R_1 to the government. At the same time, the government will also provide financial assistance M for enterprises that perform responsible innovation. However, when the government conducts passive supervision, the negative externalities of irresponsible innovation may bring about a social crisis and damage the reputation of the government, denoted by B .

Assume that the relevant parameters related to enterprises are as follows. The normal earnings on innovation for an enterprise are R_2 , whereas RI increases R&D investment C_2 . However, the enterprises' irresponsible innovation may affect ethical, ecological, and sustainable development and bring potential risks (D) to the public, and enterprises also are punished (T) by the government. At the same time, in order to show a responsible attitude, the government of negative supervision will design policies in advance to require enterprises to compensate the amount K_1 for damages to the public at potential risk.

Assume that the relevant parameters related to the public are as follows. The interaction and discussion between the public and other agents in the innovation process take opportunity cost C_3 and also bring benefits (R_3) to the public. Under the condition of public participation, the responsible government will punish the irresponsible innovation of enterprises and give the public value compensation K_2 . However, when the public does not participate, it cannot obtain interests and any compensation.

Based on the above assumptions, this research starts from the strategic choices of the government, enterprises, and the public to build a tripartite game payoff matrix, as shown in Table 1.

3.2. Multiagent Replicated Dynamics Equation. According to Table 1, the expected payoffs functions of the government's choice of AS and IS strategies are E_x and E_{1-x} , respectively:

$$E_x = yz(R_1 - C_1 - M) + y(1-z)(R_1 - C_1 - M) + (1-y)z(T - C_1 - K_2) + (1-y)(1-z)(T - C_1), \quad (1)$$

$$E_{1-x} = yzR_1 + y(1-z)R_1 + (1-y)z(-B) + (1-y)(1-z)(-B), \quad (2)$$

Based on equations (1) and (2), from the Malthusian dynamic equation, the government's replication dynamic equation $F(x)$ can be calculated:

$$F(x) = \frac{dx}{dt} = x(1-x)(E_x - E_{1-x}) = x(1-x)[T + B - C_1 - y(M + T + B) - K_2z(1-y)]. \quad (3)$$

Then, the expected payoffs functions of enterprises' choice of MRI and NMRI strategies are E_y and E_{1-y} , respectively:

$$E_y = xz(R_2 - C_2 + M) + x(1-z)(R_2 - C_2 + M) + (1-x)z(R_2 - C_2) + (1-x)(1-z)(R_2 - C_2), \quad (4)$$

$$E_{1-y} = xz(R_2 - T) + x(1-z)(R_2 - T) + (1-x)z(R_2 - K_1) + (1-x)(1-z)R_2. \quad (5)$$

Based on equations (4) and (5), from the Malthusian dynamic equation, the enterprises' replication dynamic equation $G(y)$ can be calculated:

$$G(y) = \frac{dy}{dt} = y(1-y)(E_y - E_{1-y}) = y(1-y)[x(T + M) + K_1z(1-x) - C_2]. \quad (6)$$

Finally, the expected payoffs functions of the public's choice of PP and NP strategies are E_z and E_{1-z} , respectively:

TABLE 1: Payoff matrix of tripartite collaboration.

Strategy selection		Public	
		PP (z)	NP ($1-z$)
Government AS (x)	Enterprises	$R_1 - C_1 - M$	$R_1 - C_1 - M$
	MRI (y)	$R_2 - C_2 + M$	$R_2 - C_2 + M$
	Enterprises	$R_3 - C_3$	0
	NMRI	$-C_1 + T - K_2$	$-C_1 + T$
		($1-y$)	$R_2 - T$
			$R_3 - C_3 - D + K_2$
Government IS ($1-x$)	Enterprises	R_1	R_1
	MRI (y)	$R_2 - C_2$	$R_2 - C_2$
	Enterprises	$R_3 - C_3$	0
	NMRI	$-B$	$-B$
		($1-y$)	R_2
			$R_3 - C_3 - D + K_1$

$$E_z = xy(R_3 - C_3) + x(1-y)(R_3 - C_3 - D + K_2) + (1-x)y(R_3 - C_3) + (1-x)(1-y)(R_3 - C_3 - D + K_1), \quad (7)$$

$$E_{1-z} = x(1-y)(-D) + (1-x)(1-y)(-D). \quad (8)$$

Based on equations (7) and (8), from the Malthusian dynamic equation, the public's replication dynamic equation $H(z)$ can be calculated:

$$H(z) = \frac{dz}{dt} = z(1-z)(E_z - E_{1-z}) = z(1-z)[R_3 - C_3 + K_1(1-x)(1-y) + K_2x(1-y)]. \quad (9)$$

Consequently, combining equations (3), (6), and (9) can produce a three-dimensional dynamic system including the government, enterprises, and the public:

$$\begin{cases} F(x) = x(1-x)[T + B - C_1 - y(M + T + B) - K_2z(1-y)], \\ G(y) = y(1-y)[x(T + M) + K_1z(1-x) - C_2], \\ H(z) = z(1-z)[R_3 - C_3 + K_1(1-x)(1-y) + K_2x(1-y)]. \end{cases} \quad (10)$$

3.3. Equilibrium Point and Multiagent Asymptotic Stability Analysis. According to equation (10), it is obvious that the three-dimensional dynamic system has 8 equilibrium points, namely, (0, 0, 0), (0, 1, 0), (0, 0, 1), (0, 1, 1), (1, 0, 0), (1, 1, 0), (1, 0, 1), and (1, 1, 1), which form the boundary of the domain Ω of the evolutionary game as $\{(x, y, z) | 0 \leq x \leq 1; 0 \leq y \leq 1; 0 \leq z \leq 1\}$. Also, the three-dimensional dynamic system may have a mixing strategy equilibrium point $E^*(x^*, y^*, z^*)$ that satisfies equation (11). However, when $E^* \notin \Omega$, E^* should be rejected. Furthermore, in the tripartite evolutionary game model, the relevant parameters may affect the agent's behavior choices. Therefore, the replicator dynamics equation of each agent can be derived to obtain equation (12):

$$\begin{cases} T + B - C_1 - y(M + T + B) - K_2z(1 - y) = 0, \\ x(T + M) + K_1z(1 - x) - C_2 = 0, \\ R_3 - C_3 + K_1(1 - x)(1 - y) + K_2x(1 - y) = 0, \end{cases} \quad (11)$$

$$\begin{cases} F'(x) = (1 - 2x)[T + B - C_1 - y(M + T + B) - K_2z(1 - y)], \\ G'(y) = (1 - 2y)[x(T + M) + K_1z(1 - x) - C_2], \\ H'(z) = (1 - 2z)[R_3 - C_3 + K_1(1 - x)(1 - y) + K_2x(1 - y)]. \end{cases} \quad (12)$$

3.3.1. Asymptotic Stability Analysis for the Government

- (1) Based on equation (3), if $T + B - C_1 - y(M + T + B) - K_2z(1 - y) = 0$, there is $F(x) \equiv 0$. At this time, the government's strategic selection will not change over time; that is, the government's strategy is stable.
- (2) If $T + B - C_1 - y(M + T + B) - K_2z(1 - y) > 0$, suppose $F(x) = 0$, and $x = 0, x = 1$ are its two stable solutions. According to $F'(x)$ in equation (12), $F'(0) > 0, F'(1) < 0$ can be calculated, and then $x = 1$ is the equilibrium point. It shows that if the government's penalty T and reputation loss B exceed the sum of the supervision cost C_1 , financial support M , and value compensation K_2 , then the active supervision strategy is the evolutionary stable strategy for the government.
- (3) If $T + B - C_1 - y(M + T + B) - K_2z(1 - y) < 0$, at this time, $F'(0) < 0, F'(1) > 0$ can be calculated, and then $x = 0$ is the equilibrium point. It shows that if the government's penalty T and reputation loss B are lower than the sum of the supervision cost C_1 , financial support M , and value compensation K_2 , then the inactive supervision strategy is the evolutionary stable strategy for the government.

3.3.2. Asymptotic Stability Analysis for Enterprises

- (1) Based on equation (6), if $x(T + M) + K_1z(1 - x) - C_2 = 0$, there is $G(y) \equiv 0$. At this time, enterprises' strategic selection will not change over time; that is, enterprises' strategy is stable.
- (2) If $x(T + M) + K_1z(1 - x) - C_2 > 0$, suppose $G(y) = 0$, and $y = 0, y = 1$ are its two stable solutions. According to $G'(y)$ in equation (12), $G'(0) > 0, G'(1) < 0$ can be calculated, and then $y = 1$ is the equilibrium point. It shows that if the sum of the punishment T , the financial support M , and the public compensation K_1 exceeds its RI investment C_2 , then the making RI strategy is the evolutionary stable strategy for enterprises.
- (3) If $x(T + M) + K_1z(1 - x) - C_2 < 0$, at this time, $G'(0) < 0, G'(1) > 0$ can be calculated, and then $y = 0$ is the equilibrium point. It shows that if the sum of the punishment T , the financial support M , and the

public compensation K_1 is lower than its RI investment C_2 , then the making irresponsible innovation strategy is the evolutionary stable strategy for enterprises.

3.3.3. Asymptotic Stability Analysis for the Public

- (1) Based on equation (9), if $R_3 - C_3 + K_1(1 - x)(1 - y) + K_2x(1 - y) = 0$, there is $H(z) \equiv 0$. At this time, the public's strategic selection will not change over time; that is, the public's strategy is stable.
- (2) If $R_3 - C_3 + K_1(1 - x)(1 - y) + K_2x(1 - y) > 0$, suppose $H(z) = 0$, and $z = 0, z = 1$ are its two stable solutions. According to $H'(z)$ in equation (12), $H'(0) > 0, H'(1) < 0$ can be calculated, and then $z = 1$ is the equilibrium point. It shows that if the sum of the public's benefits R_3 and compensation (K_1 and K_2) exceeds its opportunity cost of participation C_3 , then the positive participation strategy is the evolutionary stable strategy for the public.
- (3) If $R_3 - C_3 + K_1(1 - x)(1 - y) + K_2x(1 - y) < 0$, at this time, $H'(0) < 0, H'(1) > 0$ can be calculated, and then $z = 0$ is the equilibrium point. It shows that if the sum of the public's benefits R_3 and compensation (K_1 and K_2) is lower than its opportunity cost of participation C_3 , then the negative participation strategy is the evolutionary stable strategy for the public.

3.4. Equilibrium Point Stability Analysis. In the previous section, 9 equilibrium points in the three-dimensional dynamic system have been described, but it is uncertain whether these points are the evolutionary stability strategy of the system. According to the research of Wainwright [42] and Lyapunov and Mikhailovich [43], the equilibrium point is asymptotically stable ESS (called sink) only when both strict Nash equilibrium and pure strategy Nash equilibrium are simultaneously satisfied, but E^* is a mixed strategy Nash equilibrium, and it is not a sink. Therefore, we only need to analyze the asymptotic stability of the remaining 8 equilibrium points. Then, based on Friedman's research conclusion [44], the local stability of the equilibrium point is analyzed by the system's Jacobian matrix, that is, $\det(J) > 0$ and $\text{tr}(J) < 0$, thus solving for the system's Jacobian matrix as shown in

$$J = \begin{bmatrix} \frac{\partial F(x)}{\partial x} & \frac{\partial F(x)}{\partial y} & \frac{\partial F(x)}{\partial z} \\ \frac{\partial G(y)}{\partial x} & \frac{\partial G(y)}{\partial y} & \frac{\partial G(y)}{\partial z} \\ \frac{\partial H(z)}{\partial x} & \frac{\partial H(z)}{\partial y} & \frac{\partial H(z)}{\partial z} \end{bmatrix} = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{bmatrix}, \quad (13)$$

where $J_{11} = (1 - 2x)[T + B - C_1 - y(M + T + B) - K_2z(1 - y)]$; $J_{12} = x(1 - x)(K_2z - M - T - B)$; $J_{13} = x(1 - x)K_2(y - 1)$; $J_{21} = y(1 - y)(M + T - K_1z)$; $J_{22} = (1 - 2y)[x(M + T) + K_1z(1 - x) - C_2]$; $J_{23} = y(1 - y)K_1(1 - x)$; $J_{31} = z(1 - z)(1 - y)(K_2 - K_1)$; $J_{32} = z(1 - z)(K_1x - K_2x - K_1)$; $J_{33} = (1 - 2z)[R_3 - C_3 + K_1(1 - x)(1 - y) + K_2x(1 - y)]$.

The $\det(J)$ and $\text{tr}(J)$ values of the equilibrium point can be obtained from the above Jacobian matrix and judgment conditions, as shown in Table 2. According to Table 2, the size of $\det(J)$ and $\text{tr}(J)$ is determined by the relevant parameter values, but the existing conditions cannot determine the stability of the 8 equilibrium points; that is, whether there is an equilibrium point that makes the tripartite evolutionary game reach stability is unclear. System dynamics (SD) can analyze the complex dynamic evolution process of the evolutionary game model under the conditions of limited rationality and asymmetric information [45]. Hence, the next section will combine SD simulation tools to construct a multiagent evolutionary game model and analyze the effect of different parameter values on the evolutionary game process.

4. Numerical Simulation Analysis Based on System Dynamics

Regarding SD modeling, Sterman [46] pointed out that the simulation model does not lie in how real it is, but in its usefulness, and focuses on revealing the regularity of changes in things. Meanwhile, the correctness of the SD model structure is more important than the parameter value selection, so we should focus on the validity, consistency, and adaptability of the model structure [47]. SD simulation can effectively analyze the feedback behavior of complex systems and the effectiveness of strategies in order to provide practical guidance for designing, formulating, and optimizing management policies [48]. Therefore, this section will construct a multiagent SD model to research the long-term collaborative relationship and behavior strategy evolution of multiagent in responsible innovation, thus providing an experimental platform for studying different influencing factors and making effective decisions.

4.1. Simulation Model Construction. According to the above game analysis and replicator dynamics equations, this study uses Vensim PLE 8.0.9 to establish a multiagent evolutionary game SD model, which includes three subsystems, namely, the government, enterprises, and the public. As shown in

Figure 1, the multiagent evolutionary game SD model is composed of multiple variables, namely, three level variables, three rate variables, nine intermediate variables, and twelve exogenous variables. Specifically, three level variables are used to represent, respectively, the probability of the government choosing AS strategy, probability of enterprises choosing MRI strategy, and probability of the public choosing PP strategy; the change rate of supervision by the government, the change rate of enterprises making responsible innovation, and the change rate of participation by the public are described by three rate variables; and the exogenous variables correspond to the parameters in the payoff matrix of the tripartite collaboration of Table 1. Moreover, in the evolutionary game SD model, the flow rate formula and the functional relation among level variables, rate variables, intermediate variables, and exogenous variables are based mainly on the above multiagent evolutionary game analysis and replicated dynamic equation, namely, equations (1)–(10).

The SD model setting is as follows: INITIAL TIME = 0, FINAL TIME = 10, TIME STEP = 0.03125, Units for Time: Year, and Integration Type: Euler. Meanwhile, according to the case of RI practice [38, 41] and related literature value assignment method [49], the initial values of different exogenous variables in the SD model are set as $R_1 = 50$, $C_1 = 24$, $M = 4$, $T = 15$, $B = 13$, $R_2 = 30$, $C_2 = 12$, $K_1 = 4$, $D = 2$, $K_2 = 3$, $R_3 = 10$, $C_3 = 8$.

4.2. Model Simulation Analysis. According to the above game equilibrium analysis, it can be seen that an evolutionary equilibrium is bound to be reached among the government, enterprises, and the public, while the reasons, processes, and stability for achieving equilibrium are not clear. Furthermore, because the three-dimensional dynamic system is affected by different factors, which makes the equilibrium state reached in a certain situation quickly broken, therefore, this study uses SD modeling to simulate the dynamic game between the three parties in RI, in order to analyze the multiagent behavior strategy selection and its sensitivity to changes in relevant parameters.

4.2.1. Overall Simulation Analysis of the Model. In the process of RI, when the initial state of the tripartite game among the government, enterprises, and the public is pure strategy, that is, the strategy choices of each agent are divided into 0 and 1, this equilibrium state is unstable and is broken when one or more of multiagents make small changes. Therefore, the equilibrium point (1, 0, 0) is taken as an example to simulate and analyze the evolution process of multiagent.

As shown in Figure 2, according to the simulation analysis of the equilibrium point (1, 0, 0), when the government selects the AS strategy, both enterprises and the public play the game with a small probability of strategy selection (that is 0.01), whereas, once their strategy mutations can yield higher expected benefits, they adjust the strategy to bring the system to a new equilibrium state. In fact, if enterprises and the public play with a higher

TABLE 2: $\det(J)$ and $\text{tr}(J)$ of the Jacobian matrix.

Equilibrium point	$\det(J)$	$\text{tr}(J)$
(0, 0, 0)	$-C_2(T+B-C_1)(R_3-C_3+K_1)$	$T+B-C_1-C_2+R_3-C_3+K_1$
(0, 1, 0)	$C_2(-C_1-M)(R_3-C_3)$	$-C_1-M+C_2+R_3-C_3$
(0, 0, 1)	$(T+B-C_1-K_2)(K_1-C_2)(C_3-R_3-K_1)$	$T+B-C_1-K_2-C_2+C_3-R_3$
(0, 1, 1)	$(-C_1-M)(C_2-K_1)(C_3-R_3)$	$-C_1-M+C_2-K_1+C_3-R_3$
(1, 0, 0)	$(C_1-T-B)(T+M-C_2)(R_3-C_3+K_2)$	$C_1-B+M-C_2+R_3-C_3+K_2$
(1, 1, 0)	$(C_1+M)(C_2-T-M)(R_3-C_3)$	$C_1+C_2-T+R_3-C_3$
(1, 0, 1)	$(C_1-T-B+K_2)(T+M-C_2)(C_3-R_3-K_2)$	$C_1-B+M-C_2+C_3-R_3$
(1, 1, 1)	$(C_1+M)(C_2-T-M)(C_3-R_3)$	$C_1+C_2-T+C_3-R_3$

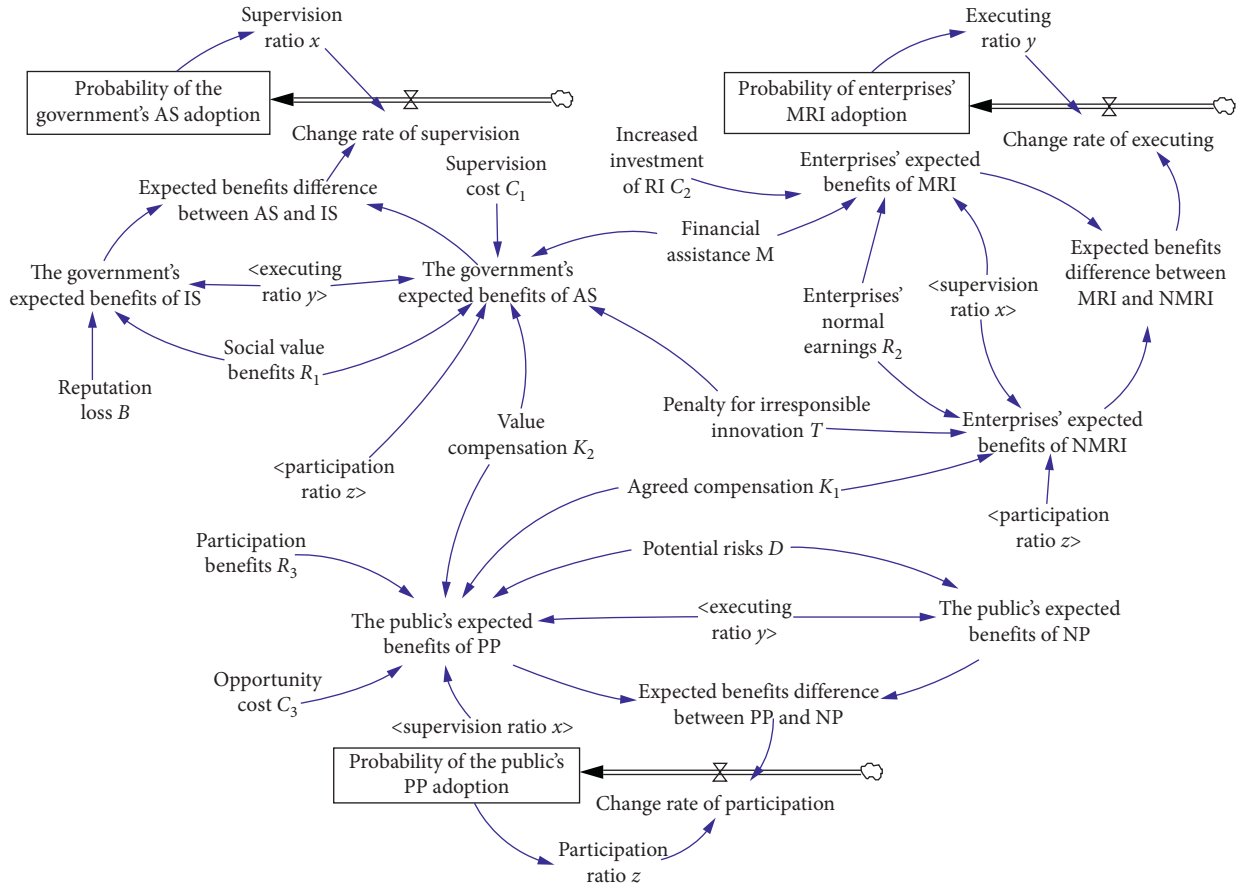


FIGURE 1: Multiagent evolutionary game SD model.

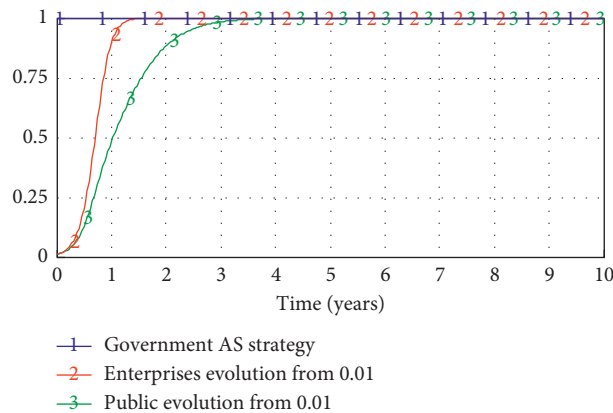


FIGURE 2: The evolution process of y and z starting from 0.01.

probability, the time for making strategy selection is shorter and the range of change is larger, and eventually the system will reach a new equilibrium state.

The external pressure from public participation in RI forced the government and enterprises to act to understand and respond to societal needs and values. Simulating the stability of other equilibrium points also verifies that public participation is its optimal choice; namely, when the public makes strategic choices with a small probability mutation, they will eventually reach an equilibrium state of 1. As shown in Figure 3, when the public is always actively participating in the process of RI, if the government chooses the AS strategy, no matter which strategy enterprises choose to mutate, its final strategy choice will reach an equilibrium state of 1; that is, enterprises make RI, which is consistent with the simulation results of the equilibrium points (1, 0, 1) and (1, 1, 1). However, once the government chooses the IS strategy, the final strategic choice of enterprises will reach an equilibrium state of 0 (see Figure 4); that is, enterprises make irresponsible innovation, which corresponds to the simulation results at the equilibrium points (0, 0, 1) and (0, 1, 1).

Similarly, under the conditions of public positive participation, if enterprises select the MRI strategy, regardless of which strategy the government chooses to play, it will ultimately choose the IS strategy. Over time, the lax state of the government will drive enterprises to select the NMRI strategy. At this time, no matter which strategy the government chooses to play, it will eventually choose the AS strategy to break the initial equilibrium state of (0, 0, 1). Through the repeated evolutionary game, the evolution of each agent strategy is ultimately stabilized in the equilibrium state of (1, 1, 1).

4.2.2. Simulation Analysis of the Government Strategy Selection. According to the local stability analysis of the equilibrium point in Table 2, the stability of the equilibrium point depends on the size of the relevant game parameters, that is, the value of exogenous variables in the SD game model. The following is still taking the strategy combination (1, 0, 0) as an example to analyze the sensitivity of the government strategy selection to the change of exogenous variables.

Assume that the government’s initial strategy choice is inactive supervision and the evolutionary game is mutated from 0 to 0.1. It can be seen from the dynamic simulation that the government’s strategic choices are primarily affected by two exogenous variables, namely, supervision cost C_1 and penalty T for irresponsible innovation to enterprises. As shown in Figures 5 and 6, when the government’s strategy selection begins to evolve from the initial supervision cost, the response time for strategy selection becomes shorter as the cost decreases, and the equilibrium state of 1 is reached faster; otherwise, if the cost increases beyond the government’s tolerance, its strategy choice evolves to an “inactive supervision” state of 0. Meanwhile, the higher the penalty threshold for enterprises, the greater the variability and probability that the government chooses the AS strategy. Moreover, comparing Figures 5 and 6, it can be seen that the

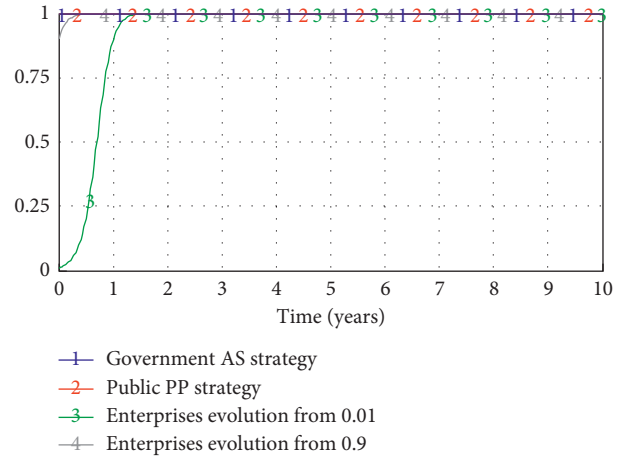


FIGURE 3: The evolutionary process of enterprises under active supervision strategy.

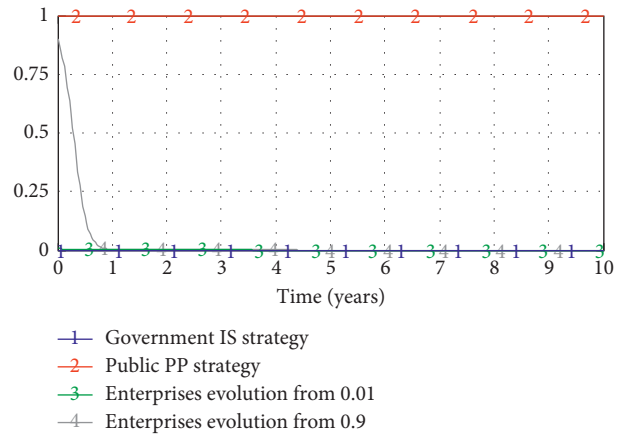


FIGURE 4: The evolutionary process of enterprises under active supervision strategy.

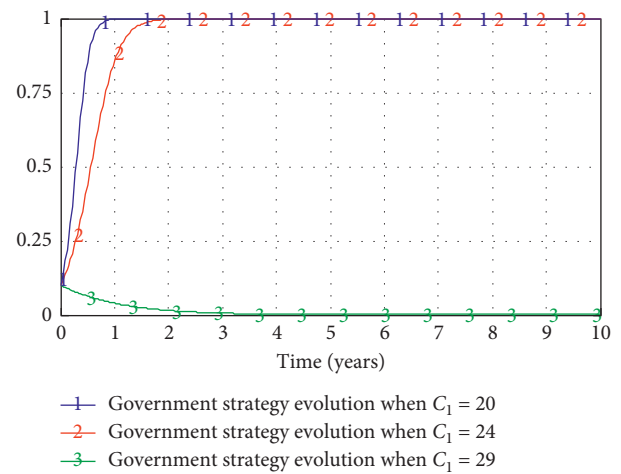


FIGURE 5: The impact of C_1 on the government strategy selection.

greater the probability that the government chooses AS strategy, that is, the higher the government’s supervision for enterprises, the greater the penalty for enterprises’

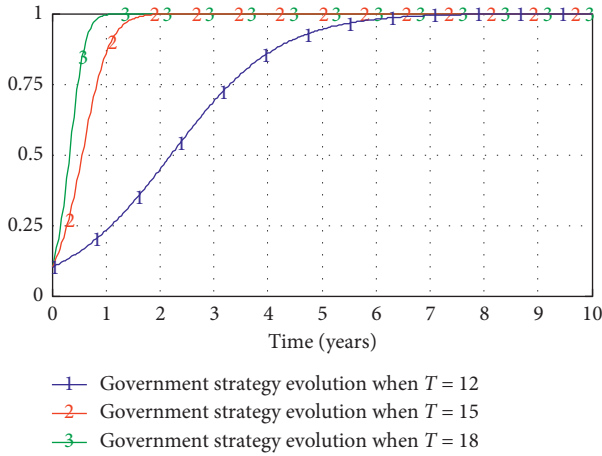


FIGURE 6: The impact of T on government strategy selection.

irresponsible innovation, which compensates for the higher cost of making an AS strategy faster.

4.2.3. Simulation Analysis of Enterprises Strategy Selection. According to the local stability analysis of the equilibrium point in Table 2, the strategy combination (1, 0, 0) is taken as an example to analyze the sensitivity of enterprises' strategy selection to the change of exogenous variables. Further analysis found that enterprises' strategic choices were mainly influenced by exogenous variables such as the increased R&D investment C_2 and penalty T .

As shown in Figure 7, enterprises start a game with a probability of 0.1; when their R&D costs are lower and penalties are higher, their response time to choose an MRI strategy is shorter and reaches an equilibrium of 1 faster. As the cost increases to the equivalent of penalty, the enterprises' response for strategy selection becomes slower, and once the cost is too high or even exceeds the penalty, enterprises may mutate to an equilibrium of 0; that is, enterprises would rather choose to accept punishment than make RI. This also indicates that a stable difference between C_2 and T must be maintained; once a variable is too high or low, this will affect the strategic choice of enterprises to achieve a stable state. This is also true in reality, where the intensity of the government penalty may affect the threshold for enterprises to bear punishment, and once the threshold is too high, it will make enterprises reduce their sensitivity to punishment and lose the motivation effect.

4.2.4. Simulation Analysis of the Public Strategy Selection. According to the local stability analysis of the equilibrium point in Table 2, the strategy combination (1, 0, 0) is taken as an example to analyze the sensitivity of the public's strategy selection to the change of exogenous variables. Further analysis found that the public's strategic choices were mainly influenced by exogenous variables such as opportunity cost C_3 , participation benefits R_3 , and value compensation K_2 .

As shown in Figure 8, when the public starts mutation with a probability of 0.1 to make an evolutionary game, it is found that three exogenous variables increase or decrease separately;

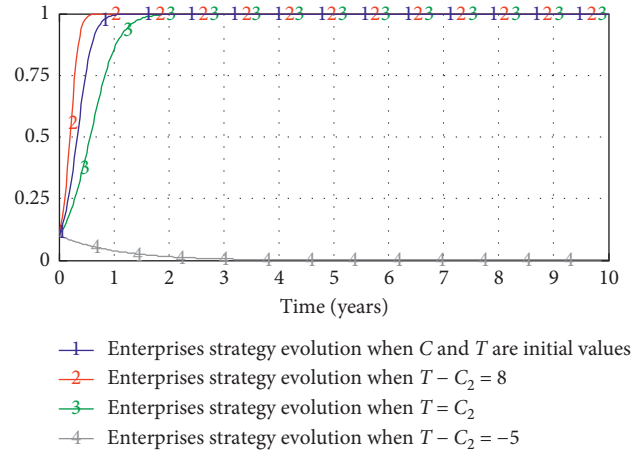


FIGURE 7: The impact of C_2 and T on enterprises strategy selection.

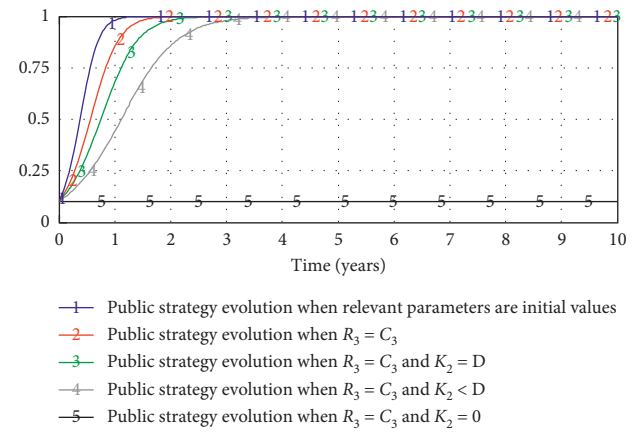


FIGURE 8: The impact of R_3 , C_3 , and K_2 on the public strategy selection.

namely, R_3 and K_2 increase, C_3 decreases, the public chooses the PP strategy with shorter reaction time and greater probability, and the trend of change is similar to that of the government and enterprises, with an s -type curve evolution. Specifically, the public responds faster to making strategic choices when each variable is the initial value, while the government's value compensation K_2 significantly affects the public's strategy selection as the cost of participation increases. Once this kind of negative external value compensation is reduced or is even 0, the public will reduce the probability of strategy choice because the risks suffered cannot be compensated for; but excessively higher value compensation will increase the government's financial burden and force it to change its strategic choices. Therefore, in the process of RI, the expected benefits of public participation and government compensation are important factors influencing the choice of strategy for the public, and the government should pay attention to the "bottom line" of public participation and set a reasonable amount of compensation.

5. Conclusions

In the process of RI, a stable collaborative relationship is a fundamental prerequisite for achieving responsible

governance in scientific and technological innovation. In this paper, the evolutionary game method is used to construct a behavioral strategy game model of multiagent participation in RI and finally draw some significant conclusions by system dynamics simulation.

First, the strategies of tripartite eventually converged to the best equilibrium state of the government's active supervision, enterprises making responsible innovation, and the public's positive participation. Second, the active participation of the public promotes the game to reach a stable state more quickly, and enterprises' behavior strategies are more susceptible to government actions. Third, supervision cost, penalty amount, and value compensation are the most important factors that affect the change of the corresponding agent's behavior strategy, and the final strategic stability of tripartite is influenced by multiple exogenous variables.

Given the above research results, the government, first of all, should play a leading role and establish an institutional system that matches responsible innovation. Then, it should establish a linkage mechanism between reward and penalty and setting a reasonable threshold for penalties. Third, the government should construct a sound public decision-making participation mechanism, which can improve the rationality of related innovation decisions.

This research contends with several limitations. First, for the convenience of analysis, all multiagent and influencing factors are not considered. Second is the lack of actual data of exogenous variables due to conditional constraints; we only set the values of exogenous variables for the evolutionary game by referring to the case of responsible innovation practice and related research assignment method. Therefore, in future research, we will combine practice investigation by adding real data and more influencing factors and agents into the model.

Data Availability

The simulation data used to support the findings of this study are included within the article.

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 Ministry of Education Humanities and Social Sciences Planning Foundation of China (Grant no. 19YJA790028), National Natural Science Foundation of China (Grant no. 71502101), and Shanghai Philosophy and Social Science Foundation (Grant no. 2019EGL017).

References

- [1] G. George, J. Howard-Grenville, A. Joshi, and L. Tihanyi, "Understanding and tackling societal grand challenges through management research," *Academy of Management Journal*, vol. 59, no. 6, pp. 1880–1895, 2016.
- [2] B. Godin, *Innovation Contested: The Idea of Innovation over the Centuries*, Routledge, Abingdon, UK, 2015.
- [3] J. Stilgoe, R. Owen, and P. Macnaghten, "Developing a framework for responsible innovation," *Research Policy*, vol. 42, no. 9, pp. 1568–1580, 2013.
- [4] B. E. Ribeiro, R. D. J. Smith, and K. Millar, "A mobilising concept? unpacking academic representations of responsible research and innovation," *Science and Engineering Ethics*, vol. 23, no. 1, pp. 81–103, 2017.
- [5] S. Khavul and G. D. Bruton, "Harnessing innovation for change: sustainability and poverty in developing countries," *Journal of Management Studies*, vol. 50, no. 2, pp. 285–306, 2013.
- [6] L. M. D. Silva, C. C. Bitencourt, K. Faccin, and T. Iakovleva, "The role of stakeholders in the context of responsible innovation: a meta-synthesis," *Sustainability*, vol. 11, no. 6, p. 1766, 2019.
- [7] R. Von Schomberg, "A vision of responsible research and innovation," in *Responsible Innovation*, J. B. R. Owen and M. Heintz, Eds., John Wiley & Sons, Chichester, UK, pp. 51–74, 2013.
- [8] E. Fisher and A. Rip, "Responsible innovation: multi-level dynamics and soft intervention practices," in *Responsible Innovation*, R. Owen, J. Bessant, and M. Heintz, Eds., pp. 165–183, John Wiley & Sons, Chichester, UK, 2013.
- [9] V. Blok and P. Lemmens, "The emerging concept of responsible innovation: three reasons why it is questionable and calls for a radical transformation of the concept of innovation," in *Responsible Innovation 2*, B.-J. Koops, I. Oosterlaken, H. Romijn, T. Swierstra, and J. V. D. Hoven, Eds., Springer, Cham, Switzerland, 2015.
- [10] T. Brand and V. Blok, "Responsible innovation in business: a critical reflection on deliberative engagement as a central governance mechanism," *Journal of Responsible Innovation*, vol. 6, no. 1, pp. 4–24, 2019.
- [11] A. Genus and M. Iskandarova, "Responsible innovation: its institutionalisation and a critique," *Technological Forecasting and Social Change*, vol. 128, pp. 1–19, 2018.
- [12] R. Owen, S. Jack, P. Macnaghten, M. Gorman, E. Fisher, and D. Guston, "A framework for responsible innovation," in *Responsible Innovation*, R. Owen, J. Bessant, and M. Heintz, Eds., John Wiley & Sons, Chichester, UK, pp. 27–50, 2013.
- [13] K. Jarmai and H. Vogel-Pöschl, "Meaningful collaboration for responsible innovation," *Journal of Responsible Innovation*, vol. 7, no. 1, pp. 138–143, 2020.
- [14] D. Arenas, P. Sanchez, M. Murphy, and M. Murphy, "Different paths to collaboration between businesses and civil society and the role of third parties," *Journal of Business Ethics*, vol. 115, no. 4, pp. 723–739, 2013.
- [15] R. Von Schomberg, "Prospects for technology assessment in a framework of responsible research and innovation," in *Technikfolgen Abschätzen Lehren*, M. Dusseldorp and R. Beecroft, Eds., Springer, Wiesbaden, Germany, pp. 39–61, 2012.
- [16] R. Lubberink, V. Blok, J. Van Ophem, and O. Omta, "Lessons for responsible innovation in the business context: a systematic literature review of responsible, social and sustainable innovation practices," *Sustainability*, vol. 9, no. 5, p. 721, 2017.
- [17] A. Pyka, "Avoiding evolutionary inefficiencies in innovation networks," *Prometheus*, vol. 32, no. 3, pp. 265–279, 2014.
- [18] M. P. Schlaile, M. Mueller, M. Schramm, and A. Pyka, "Evolutionary economics, responsible innovation and demand: making a case for the role of consumers," *Philosophy of Management*, vol. 17, no. 1, pp. 7–39, 2018.

- [19] J. Bessant, "Innovation in the twenty-first century," in *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*, R. Owen, J. Bessant, and M. Heintz, Eds., pp. 1–25, John Wiley & Sons, Chichester, UK, 2013.
- [20] H. Zwart, L. Landeweerd, and A. Van Rooij, "Adapt or perish? assessing the recent shift in the European research funding arena from "elsa" to "rri"," *Life Sciences, Society and Policy*, vol. 10, no. 11, pp. 1–19, 2014.
- [21] P. Macnaghten, R. Owen, J. Stilgoe et al., "Responsible innovation across borders: tensions, paradoxes and possibilities," *Journal of Responsible Innovation*, vol. 1, no. 2, pp. 191–199, 2014.
- [22] M. Burget, E. Bardone, and M. Pedaste, "Definitions and conceptual dimensions of responsible research and innovation: a literature review," *Science and Engineering Ethics*, vol. 23, no. 1, pp. 1–19, 2017.
- [23] R. Von Schomberg, *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields*, European Commission Services, Brussels, Belgium, 2011.
- [24] S. R. Davies and M. Horst, "Responsible innovation in the US, UK and Denmark: governance landscapes," in *Responsible Innovation 2*, B.-J. Koops, I. Oosterlaken, H. Romijn, T. Swierstra, and J. V. D. Hoven, Eds., vol. 37, Cham, Switzerland, Springer, 2015.
- [25] V. Blok, L. Hoffmans, and E. F. M. Wubben, "Stakeholder engagement for responsible innovation in the private sector: critical issues and management practices," *Journal on Chain and Network Science*, vol. 15, no. 2, pp. 147–164, 2015.
- [26] L. Mei and J. Chen, "Responsible innovation: origin, attribution and theoretical framework," *Management World*, vol. 8, pp. 39–57, 2015, in Chinese.
- [27] N. M. Dahan, J. P. Doh, and J. D. Raelin, "Pivoting the role of government in the business and society interface: a stakeholder perspective," *Journal of Business Ethics*, vol. 131, no. 3, pp. 665–680, 2015.
- [28] N. Thomas and H. Rogers, "Responsible innovation in supply chains: insights from a car development perspective," in *Proceedings of the 4th International Conference LDIC*, Bremen, Germany, 2016.
- [29] W.-M. Hur, H. Kim, and J. Woo, "How csr leads to corporate brand equity: mediating mechanisms of corporate brand credibility and reputation," *Journal of Business Ethics*, vol. 125, no. 1, pp. 75–86, 2014.
- [30] A. Gurzawska, "Towards responsible and sustainable supply chains—innovation, multi-stakeholder approach and governance," *Philosophy of Management*, 2019.
- [31] I. Bouzguenda, C. Alalouch, and N. Fava, "Examining digital participatory planning: maturity assessment in a small Dutch city," *Journal of Cleaner Production*, vol. 264, Article ID 121706, 2020.
- [32] A. Gurzawska, M. Mäkinen, and P. Brey, "Implementation of responsible research and innovation (rri) practices in industry: providing the right incentives," *Sustainability*, vol. 9, no. 10, p. 1759, 2017.
- [33] L. Samuelson, "Evolution and game theory," *Journal of Economic Perspectives*, vol. 16, no. 2, pp. 47–66, 2002.
- [34] R. Zhao, X. Zhou, J. Han, and C. Liu, "For the sustainable performance of the carbon reduction labeling policies under an evolutionary game simulation," *Technological Forecasting and Social Change*, vol. 112, pp. 262–274, 2016.
- [35] X. Liu, Z. Fang, N. Zhang, K. Liu, and Z. Jingfeng, "An evolutionary game model and its numerical simulation for collaborative innovation of multiple agents in carbon fiber industry in China," *Sustainable Computing: Informatics and Systems*, vol. 24, p. 100350, 2019.
- [36] B. Guo and J. Li, "Research on the evolution of participants collaboration mechanism in ppp model based on computer simulation: based on the old community renovation project," *The Journal of Supercomputing*, vol. 76, no. 4, pp. 2417–2434, 2020.
- [37] Y. Chen, Z. Hu, Q. Liu, and S. Chen, "Evolutionary game analysis of tripartite cooperation strategy under mixed development environment of cascade hydropower stations," *Water Resources Management*, vol. 34, no. 6, pp. 1951–1970, 2020.
- [38] W. Ravesteijn, Y. Liu, and P. Yan, "Responsible innovation in port development: the rotterdam maasvlakte 2 and the dalian dayao bay extension projects," *Water Science and Technology*, vol. 72, no. 5, pp. 665–677, 2015.
- [39] L. Song and W. Ravesteijn, "Responsible port innovation in China: the case of the yangshan port extension project," *International Journal of Critical Infrastructures*, vol. 11, no. 4, pp. 297–315, 2015.
- [40] J. Voeten, J. D. Haan, G. D. Groot, and N. Roome, "Understanding responsible innovation in small producers' clusters in vietnam through actor-network theory," *The European Journal of Development Research*, vol. 27, no. 2, pp. 289–307, 2015.
- [41] W. Ravesteijn, J. He, and C. Chen, "Responsible innovation and stakeholder management in infrastructures: the nansha port railway project," *Ocean & Coastal Management*, vol. 100, pp. 1–9, 2014.
- [42] J. Wainwright, "A dynamical systems approach to bianchi cosmologies: orthogonal models of class a," *Classical and Quantum Gravity*, vol. 6, no. 10, pp. 1409–1431, 1989.
- [43] A. M. Lyapunov and A. Mikhailovich, "The general problem of the stability of motion," *International Journal of Control*, vol. 55, no. 3, pp. 531–534, 1992.
- [44] D. Friedman, "Evolutionary games in economics," *Econometrica*, vol. 59, no. 3, pp. 637–666, 1991.
- [45] D.-H. Kim, D. H. Kim, and D. H. Kimb, "A system dynamics model for a mixed-strategy game between police and driver," *System Dynamics Review*, vol. 13, no. 1, pp. 33–52, 1997.
- [46] J. Sterman, *Business Dynamics: Systems Thinking and Modeling for a Complex World*, McGraw-Hill Education, Boston, MA, USA, 2000.
- [47] D. D. Wu, X. Kefan, L. Hua, Z. Shi, and D. L. Olson, "Modeling technological innovation risks of an entrepreneurial team using system dynamics: an agent-based perspective," *Technological Forecasting and Social Change*, vol. 77, no. 6, pp. 857–869, 2010.
- [48] M. You, S. Li, D. Li, Q. Cao, and F. Xu, "Evolutionary game analysis of coal-mine enterprise internal safety inspection system in China based on system dynamics," *Resources Policy*, vol. 67, p. 101673, 2020.
- [49] Q. Liu, X. Li, and X. Meng, "Effectiveness research on the multi-player evolutionary game of coal-mine safety regulation in China based on system dynamics," *Safety Science*, vol. 111, pp. 224–233, 2019.

Research Article

Integrated Pricing and Distribution Planning for Community Group Purchase of Fresh Agricultural Products

Wenbing Shui and Mengxia Li 

Faculty of Transportation Engineering, Kunming University of Science and Technology, Kunming 650500, China

Correspondence should be addressed to Mengxia Li; limengxia@stu.kust.edu.cn

Received 4 June 2020; Revised 19 June 2020; Accepted 24 June 2020; Published 22 July 2020

Academic Editor: Lu Zhen

Copyright © 2020 Wenbing Shui and Mengxia Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As a new social e-commerce model, community group purchase of fresh agricultural products has been gradually welcomed by the public. However, its development and operation still face homogeneous competition problems. In order to enhance competitive advantages of operators, this paper proposes a collaborative optimization mechanism, including a new pricing model and a new cold chain vehicle route planning model. It aims to ensure the quality of fresh products, reduce logistics costs, and improve enterprise profitability. The model takes into account not only the quality of fresh products and their impact on price and demand but also the impact of quality changes on total distribution costs. A two-layer programming method is applied to realize the collaborative optimization mechanism, and then the upper and lower models are solved by mathematical derivation, proof methods, and optimization procedures, respectively. Finally, the feasibility and effectiveness of the model are verified by combining with specific examples, and the following conclusions are obtained: price, delivery quality, and total profit increase with the increase of potential market demand rate. The lower the refrigeration temperature of the vehicle we choose within a certain range, the higher the quality can be obtained. In order to obtain the highest profit, community group purchase operators can choose a higher distribution temperature on the premise that they can guarantee that the quality of fresh agricultural products can be at an appropriate level.

1. Introduction

Online sales of fresh agricultural products in China are developing rapidly, but the penetration rate is still less than 3%, far lower than that of clothing, electronic products, and so on. For the fact that fresh local products are always with short shelf life and perishability, under the traditional online sales model it has greatly increased the distribution cost of cold chain distribution. The development of fresh agricultural e-commerce has its unique driving force due to the unique nature of fresh agricultural products but also faces many problems that have not occurred in the online sales of other categories of products, such as a higher retail price and uneven product quality. However, the community group purchase has emerged in recent years and has brought hope to deal with these problems.

Community group purchase is a new business model in the context of e-commerce gathering consumers from a real

community with the same needs and forming a scale that meets the bargaining power of the supplier. In this business model, customer can always obtain the corresponding product at a large discount. The head of the group establishes a community WeChat group and publishes the product information of the group purchase in a timely manner. The residents can purchase through an online application program. The advantages of community group purchase are helpful to solve the important problems of high distribution cost and unqualified quality in the process of online sales of fresh agricultural products. At the same time, community group purchase can realize the aggregation of space and time by customer demands, which can maximize the utilization rate of cold chain facilities and equipment and simultaneously increase economies of scale. As a result, the unit cold chain logistics costs can be reduced. In addition, the economies of scale brought about by the space and time aggregation of customer demands provide the internal

motivation for logistics companies to adopt the entire cold chain, which is conducive to reducing losses. However, because the technical threshold of community group purchase is relatively low, it is easy to be copied.

Reasonable pricing is of great practical significance for increasing the attractiveness of fresh agricultural products community group purchase. In order to grab market share, sellers usually adopt certain promotional methods, of which price discounts are one of the most effective means. Reasonable pricing of group purchase is the key to increasing consumer appeal and helping e-commerce enterprises increase their profit. If the price is too high, too many customers with low reservation prices will be lost, which will reduce the attractiveness of community group purchase. But if the price is too low, the profit margin will be small, which is not conducive to the long-term sustainable development of community group purchase.

Proper distribution plan arrangements play an important role in enhancing the competitive advantage of the community group operators. On the one hand, the distribution plan affects the cost of cold chain distribution and the profit of the enterprise and thus directly affects pricing decisions. On the other hand, the effectiveness of distribution plans affects quality of the fresh agricultural products. At the same time, the quality of fresh products has a strong correlation with the demand. Therefore, the reasonable determination of the distribution plan is not only related to the success or failure of the operator to carry out the community group purchase but also related to the sustainable development of the community group purchase model.

Pricing is an important decision-making content for carrying out online community group purchase, and reasonable pricing is the key to the success of fresh agricultural products community group purchase. Zhang et al. [1] analysed the pricing mechanism of online group purchase and compared it with the pricing mechanism for individual sales. The study pointed out that the price change of online group purchase pricing is driven by the number of purchases. The dynamic pricing mechanism originated from foreign researches on online group purchase pricing. Kauffman and Wang [2] are the first scholars to study the dynamic pricing mechanism of online group purchase. The study analysed the changes in the number of orders over time under the dynamic pricing mechanism of online group purchases. Chen et al. [3, 4] used game theory to analyse the supplier's dynamic pricing strategy and the retailer's purchasing decision problem under the circumstance where the supplier and retailer adopted the group purchase sales model and then studied the effects of economies of scale and risk preference of decision-makers on discount prices and quantities under the condition that demand arrivals obey Poisson distribution. Sharif Vaghefi et al. [5] believed that the waiting time of customer is an inherent attribute of group purchase auctions and proposed a dual-market pricing model with the help of game theory concepts. Gu and Cai [6] studied the dynamic group purchase pricing mechanism of suppliers to retailers in a B2B environment and compared the profits of suppliers under traditional fixed

pricing, quantity discounts, and group purchase pricing. Based on the static pricing mechanism, some scholars conducted group purchase pricing strategies or research on specific pricing methods from different perspectives such as different channels, consumer behaviour, supplier capabilities, customer waiting time, immediate supply of goods, group purchase duration, and waiting cost. Tang and Chai [7] designed a network-based catering enterprise group purchase pricing model. They took different consumers into consideration, including loyal customers and random customer groups, as well as two sales channels for individual purchase channel and group purchase channel. Qian and Su. [8] studied the joint decision of online group purchase pricing and advertising investment under the cross-demand mechanism by establishing a decision model that analysed the relationship between advertising investment and pricing.

Cold chain distribution cost optimization is mainly based on the traditional vehicle routing problem, considering the decay characteristics of fresh agricultural product quality and energy consumption to maintain a certain temperature environment. The pure cold chain vehicle routing problem has formed more detailed research directions such as soft time windows and hard time windows, single-objective and multi-objective, multi-temperature codistribution and single temperature distribution, single yard and multiple yards, and single target and multiple targets. Based on the traditional VRP with time window limitation, Hsu et al. [9] considered the random transportation time, capacity consumption, and perishable characteristics of the distribution process caused by traffic congestion and established a random soft time window constrained VRP model and then solved it with the nearest neighbour search algorithm based on time. Osvald and Stirn [10] established a path optimization model with the consideration of quality loss costs, transportation costs, and soft time window constraints with minimum time and used a tabu search algorithm to solve it. Amorim et al. [11] considered the transportation cost, vehicle rental cost, and driver cost, established a cold chain distribution route optimization model with time window constraints for multiple models, and solved it by using a large-scale domain search algorithm. Estrada-Moreno et al. [12] established a multi-yard cold chain distribution route optimization model, which was solved by BRT (biased-randomization techniques) algorithm. Wang et al. [13] established a multi-objective path optimization model with the minimum cost and maximum average freshness with time window limitation and designed a two-stage heuristic algorithm to solve it. But there are not many studies on the optimization of the distribution route, distribution temperature, and vehicle type at the same time. Hsiao et al. took meat [14], fresh fruits, and vegetables [15] as the research objects, considering the fuel cost, carbon emission cost, cooling cost, personnel cost, out-of-stock cost, and vehicle cost and studied the multi-temperature colocated vehicle routing optimization problem. In recent years, metaheuristic algorithms have been widely used to solve different NP-hard problems such as berth allocation and vehicle route problem optimization. Zhen [16] used metaheuristic algorithm to solve a mixed-

integer programming model. In his later research [17], metaheuristic algorithm is used to solve a stochastic programming model and a robust programming model related to tactical BAP problem. Such improved metaheuristic algorithms played important roles in solving cold chain vehicle routing problem in previous studies as we introduced above.

Many scholars have carried out a lot of researches around group purchase pricing and cold chain distribution cost optimization and have achieved rich research results. However, the current research does not meet the needs of group purchase pricing of fresh agricultural products. Firstly, the quality of fresh agricultural products after production decreases with time. The storage and transportation environment at a certain temperature can effectively delay the decline of quality, but the cost of maintaining a specific temperature is very high. Existing researches do not consider the impact of this feature on fresh agricultural product group purchase pricing. Secondly, the community group purchase can realize the time and space gathering of the distribution needs of fresh agricultural products, thereby reducing the unit cold chain distribution costs. Existing cold chain distribution cost optimization models lack the consideration of the impact of the time and space aggregation characteristics of distribution demand. Finally, the cost of cold chain distribution is an important component of the total cost of fresh agricultural products community group purchases and it is a decisive factor. The selection of the distribution path of fresh agricultural products, the refrigerated vehicle type, and the distribution temperature have an important impact on the distribution cost and quality assurance, which in turn affects the price setting. This is the inevitable inherent requirement of collaborative optimization between the community group purchase pricing and the cold chain distribution cost. Existing researches on group purchase pricing rarely consider the optimization of cold chain distribution parameters, while most researches on distribution cost optimization assume that prices are fixed and known input parameters and are not related to the price setting process.

Therefore, in this article, we establish a collaborative optimization mechanism for community group purchase. It is based on the improved community group purchase pricing model and cold chain distribution cost optimization model. This collaborative optimization mechanism combines the characteristics of the short shelf life of fresh agricultural products, high cold chain distribution costs, and the time and space aggregation of community group purchase. It is based on the existing food quality decay theory and also takes online group purchase pricing theory and cold chain distribution cost optimization methods into consideration. In this paper, we take fruit and vegetable as an example to verify the effectiveness and advantages of model, algorithm, and collaborative optimization mechanism. In the model, we quantify the impact of community demand scale, geographic location, refrigerated vehicle fixed costs, variable transportation costs, refrigeration costs, price demand elasticity, quality demand elasticity, and many other factors on collaborative optimization.

The remainder of this paper is organized as follows: Section 2 introduces the descriptions and assumptions of the problems, followed by the community group purchase pricing model, cold chain distribution cost models, and the implementation of collaborative optimization mechanisms. Section 3 takes fresh fruit in an area of Yunnan Province as an example for analysis and presents the experimental settings and results. Finally, Section 4 summarizes the major conclusions of this research and gives several hints for further research.

2. Materials and Methods

2.1. Problem Description and Assumptions. The profitability of community group purchase enterprises is related to quality of the agricultural products, consumer demand, price, and production and distribution costs. Because of the characteristics of fresh agricultural product, quality is an important factor affecting price and demand, and demand is negatively correlated with price. Quality, market potential demand, and price of fresh agricultural product are the important factors affecting the actual demand. Therefore, pricing decisions are directly related to consumers' purchase intentions and enterprises' final profitability. The greater the consumer's demand is, the greater the volume of logistics distribution is. At this time, the economies of scale of distribution will be produced; thereby the unit logistics cost can be reduced. This is the key to the optimization of the cold chain logistics distribution vehicle route planning. In this article, we use the sales price as a decision variable to build an upper-level model.

The design of the distribution plan has a great impact on the quality of fresh agricultural products. Therefore, the lower-level vehicle route problem can be described as the process of planning the distribution route of the cumulative order information after a group purchase by the operator. The cold chain distribution center has different types of refrigerated distribution vehicles that can set different transportation temperatures. Each refrigerated vehicle departs from the distribution center after the assembly is completed and then passes through the demand nodes in turn to deliver the fresh products to the head of the community group. In the distribution process, it is also considered how to reduce the distribution costs, improve the quality of fresh agricultural products, and improve the consumer satisfaction.

We focused on the problem with one central yard in the process of sales and distribution of fresh agricultural products in community group purchase. $N = \{i | i = 0, 1, 2, \dots, L\}$ is a set of nodes including n customer nodes and one central depot which is denoted by "0." The VRP structure can be defined on a direct graph $G = (N, A)$. A is the set of arcs, and (i, j) represents the possibility of traveling from i to j with an associated distance or cost. There are $K = \{k | k = 0, 1, 2, \dots, V\}$, representing vehicles of different types in the central yard. The assumptions used in this paper are given as follows:

- (1) There is only one distribution center, and the distribution center has several vehicles of different

types, which fully meets the distribution volume of the market's potential demand.

- (2) Because different types of fresh agricultural products have different requirements for cold chain logistics, only a single type is selected in this paper.
- (3) The location coordinates of the center yard and the demand nodes are known.
- (4) The total order quantities of each demand node are known, and each node can only be served once by one vehicle.
- (5) The service time at each demand node is known and there is no time window restriction or sequence restriction on service time.
- (6) Once the temperature of the refrigerated vehicle container is determined, it cannot be changed.

2.2. Pricing Model. Some researchers usually use the newsvendor model to describe the relationship between the price and demand for perishable goods, which is one of the classic models on inventory management. The newsvendor problem aims to find the best order quantity that minimizes the expected loss or maximizes the expected profit in some real situations under probabilistic demand. Based on the characteristics of community group purchase in this research, operators choose to conduct unified production or purchase after collecting all of the orders in a group purchase. Thus, the purchase order quantity is consistent with the actual demand and operators will have no inventory and expected loss during the community group purchase. As community group purchase relates to operators, consumers, price, demand, product quality, etc., we build the following pricing model for fresh agricultural products in group purchase, which is widely used as an effective model for agricultural products pricing. This model is based on the price function, demand function, production cost, and cold chain logistics cost and aims to maximize profits and make decisions about different price discounts and delivery quality requirements. The demand for group purchase is a function of market potential demand, price discount, and delivery quality. The objective function (1) stands for the situation where the group purchase businesses hope to win the maximum profit $\pi(p)$. θ denotes the shared revenue proportion. C_d denotes the distribution cost and C_p denotes the production cost. p is the decision variable and it denotes the price. T denotes the duration of group purchase. D_m denotes the demand rate and it is shown in equation (2). D_0 denotes the potential market demands. In this paper, demand is affected not only by price but also by the real-time quality of fresh agricultural products. α_2 and α_3 denote the effect of price on demand coefficient and the effect of quality on demand coefficient, respectively, which are all greater than 0. M_k denotes the quality when the fresh agricultural product reaches the demand nodes. M_E denotes the quality when the consumers give up buying. Here, it means that consumers have certain requirements for the quality level of fresh agricultural products. The price function is shown as equation

(3). α_0 denotes the price constant. α_1 denotes the effect of quality on price coefficient:

$$\text{maximum } \pi(p) = [(1 - \theta)p - C_d - C_p]TD_m, \quad (1)$$

$$\text{subject to } D_m = \alpha_3 \left[\frac{(M_k - M_E)}{M_E} \right] \cdot (D_0 - \alpha_2 p), \quad (2)$$

$$p = \alpha_0 (M_k - M_E)^{\alpha_1}. \quad (3)$$

Proposition 1. *If the quality of fresh agricultural products does not meet the quality required by consumers, the sales price will be 0.*

Proof. We can know from the limit expression $\lim_{M_k \rightarrow M_E} p = 0$ that when the quality delivered to the consumer is infinitely close to the quality of the consumer giving up the purchase of the product, the sales price is infinitely close to 0. This shows that there is no market demand for this quality product at this time. And it also meets the fact that when the quality of fresh agricultural products has deteriorated, it can no longer be sold; that is, the sales price is 0. \square

Proposition 2. *The price of fresh agricultural products gradually increases with the growth of quality.*

Proof. When $M_k > M_E$, we have

$$\frac{d_p}{d_M} = \alpha_0 \alpha_1 (M_k - M_E)^{\alpha_1 - 1} > 0, \quad (4)$$

$$\frac{d_p^2}{d_M^2} = \alpha_0 \alpha_1 (\alpha_1 - 1) (M_k - M_E)^{\alpha_1 - 2} < 0,$$

which means that we can observe from the expression of the derivative function that on $M_k > M_E$ the price of fresh agricultural products gradually increases with the growth of quality. Figure 1 depicts the impact of quality on price. Its growth trend is gradually decreasing, and this trend is also in line with the general trend in the sales price of fresh agricultural products. \square

Proposition 3. *When the quality approaches a level that consumers cannot accept, no matter how other factors change, demand will approach 0.*

Proof. We can know from the limit expression,

$$\lim_{M_k \rightarrow M_E} D_m = 0, \quad (5)$$

that when the quality delivered to the consumer is infinitely close to the quality of the consumer who gave up buying the product, its market demand is close to 0. Figure 2 demonstrates the impact of quality on demand. It is consistent with the actual situation where consumers will not choose to buy products with degraded quality and merchants are not allowed to sell. \square

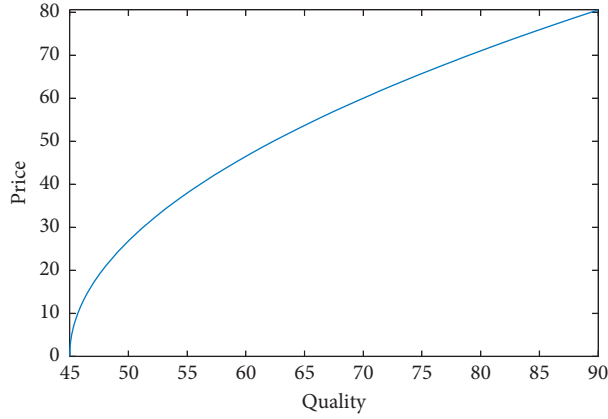


FIGURE 1: Change curve of optimal price with quality.

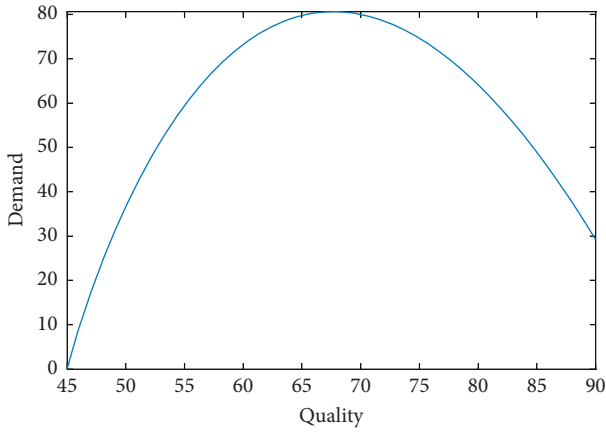


FIGURE 2: Change curve of demand with the quality.

Proposition 4. *The higher the quality of fresh agricultural products is, the more buying consumers it will attract. When the quality is accompanied by high prices, many consumers will be discouraged, resulting in a reduction in demand.*

Proof. Take the first order of derivative D_m with respect to M_k ; we have

$$p^* = \frac{(\alpha_1 + 1)[(1 - \theta)\alpha_0 D_0 + \alpha_0 \alpha_2 (C_d + C_p)]}{2(2\alpha_1 + 1)(1 - \theta)\alpha_0 \alpha_2} + \frac{\sqrt{(\alpha_1 + 1)^2 [(1 - \theta)\alpha_0 - D_0 \alpha_0 \alpha_2 (C_d + C_p)]^2 + 4(1 - \theta)\alpha_0^2 \alpha_1^2 \alpha_2 (C_d + C_p) D_0}}{2(2\alpha_1 + 1)(1 - \theta)\alpha_0 \alpha_2} \quad (10)$$

Proof. See Appendix. □

2.3. Cold Chain Logistics Distribution Cost Model

2.3.1. Objective Function. The objective of this section is to maximize the profit with considering the overall costs, which

$$\frac{d_D(p, M_k)}{d_M} = 0. \quad (6)$$

The quality when the fresh agricultural product reaches the demand nodes is

$$M_k = M_E + e^{\ln D_0 / \alpha_1 - \ln [(\alpha_1 + 1)\alpha_0 \alpha_2] / \alpha_1}. \quad (7)$$

Take the second order of derivative D_m with respect to M_k ; we have $M_k > M_E$. We can find that if demand will grow with quality,

$$M_E < M_k \leq M_E + e^{\ln D_0 / \alpha_1 - \ln [(\alpha_1 + 1)\alpha_0 \alpha_2] / \alpha_1}. \quad (8)$$

Meanwhile, if the demand will decrease with the growth of quality,

$$M_E + e^{\ln D_0 / \alpha_1 - \ln [(\alpha_1 + 1)\alpha_0 \alpha_2] / \alpha_1} < M_k < M_s. \quad (9)$$

This is consistent with the fact that consumers are more inclined to buy fresh agricultural products of better quality, but the price of fresh agricultural products will continue to rise as their quality rises, and excessively high prices will also reduce buying consumers. □

Proposition 5. *When a seller chooses a community group-buying strategy, its unique optimal solution is shown in equation (11):*

include fixed cost of distribution vehicles, variable distribution costs, cargo damage costs, refrigeration costs, labour costs, and sales losses that abandon distribution. Formula (11) is the objective function, which means to obtain the highest profit. q_i denotes the demand need to deliver to customer node i . Binary variable z_i is equal to 1 when the

customer node i is to be serviced. In this formulation, binary variable y_{ij}^k is equal to 1 when the vehicle k visits node j immediately after node i . Binary variable x_i^k is equal to 1 when the vehicle k visits node i . T_k denotes the temperature when the vehicle k is in operation. c_{ij}^k , s_{ij}^k , h_i^k , and f_k are different costs during the cold chain distribution; we will discuss each of the costs as follows:

$$\text{Max p} \sum_{i=1}^L q_i z_i - \sum_{k=1}^V \sum_{i=0}^L \sum_{j=1}^L (c_{ij}^k + s_{ij}^k) y_{ij}^k - \sum_{k=1}^V \sum_{i=1}^L h_i^k x_i^k - \sum_{k=1}^V f_k O_k. \quad (11)$$

Here c_{ij}^k denotes the variable delivery costs from i to j when we choose the vehicle k calculated by equation (12), where a_0 , b_0 , and c_0 are coefficients, and f denotes the unit fuel rate. c_{ij}^k is related to the weight of the vehicle k . w_k and w_i^k shown in equation (13) denote the weight of vehicle k and the weight when vehicle k leaves customer node i , respectively:

$$c_{ij}^k = \frac{a_0 \times (b_0 + c_0 \cdot w_i^k) d_{ij}}{w_i^k} f, \quad \forall i, j \in N, j \neq 0, k \in K, \quad (12)$$

$$w_0^k = w_k + \sum_{i=1}^L q_i x_i^k = w_k + \sum_{i=1}^L \sum_{j=1}^L q_i y_{ij}^k, \quad \forall k \in K, \quad (13)$$

where s_{ij}^k denotes the cost of refrigerant consumption of vehicle k between i and j , and it is calculated by equation (14). δ denotes the cost of unit refrigerant. λ_k denotes the coefficient determined by the vehicle k and it is shown in equation (15). μ denotes the depreciation of the vehicle. ρ denotes the coefficient of the thermal conductivity of the cabin material. In equation (16), M denotes the average surface area of the vehicle. M_W and M_N , respectively, represent the outer and inner surface areas of the vehicle. T_h denotes the environment temperature. Meanwhile T_k denotes the cooling temperature of the vehicle k and it is one of the decision variables:

$$s_{ij}^k = \delta \lambda_k \frac{d_{ij}}{v_{ij}} |T_h - T_k|, \quad \forall i, j \in N, i \neq j, k \in K, \quad (14)$$

$$\lambda_k = (1 + \mu) \times \rho \times M, \quad (15)$$

$$M = \sqrt{M_W M_N}, \quad (16)$$

where h_i^k denotes the refrigerant consumption cost of k car at customer point i , and it is the coefficient of vehicle k calculated by equation (17). η_k is a refrigerant calculation coefficient and it is calculated by equation (18). β denotes the frequency coefficient of door opening [11]. Here we suppose that when the door opening time is between 1 and 5, the value of β is 0.5. V denotes the volume of the vehicle. q_i denotes the demand at node i . v_l denotes the speed of loading and unloading. t_i denotes the loading and unloading time at node i and it is determined by q_i and v_l , shown in equation (19):

$$h_i^k = \delta \eta_k \Delta T t_i, \quad \forall k \in K, i \in N, i \neq 0, \quad (17)$$

$$\eta_k = (0.54V + 3.22) \times \beta, \quad (18)$$

$$t_i = \frac{q_i}{v_l}, \quad \forall i \in N, i \neq 0. \quad (19)$$

In the last term of (11) f_k denotes the fixed cost of the vehicle. O_k is a binary variable denoting whether the vehicle k is chosen for use.

2.3.2. Constraints. Constraint (20) indicates that each demand point that is delivered can only be delivered by one vehicle. Constraint (21) indicates that the vehicle must leave from that point after reaching a certain point. Constraint (22) indicates that if a vehicle is used, it must depart from the depot. Constraint (23) means that if a vehicle is used, it must return to the depot. Constraint (24) means to eliminate the subloop. Constraint (25) is the maximum load constraint of the vehicle. Constraint (26) means the relationship between x and y . Constraint (27) represents the relationship between the distribution temperature and other decision variables:

$$z_i - \sum_{k=1}^V x_i^k = 0, \quad \forall i \in N, i \geq 1 \quad (20)$$

$$\sum_{i=0}^L y_{ip}^k - \sum_{j=0}^L y_{pj}^k = 0, \quad \forall k \in K, p \in N, \quad (21)$$

$$O_k - \sum_{p=1}^L y_{0p}^k = 0, \quad \forall k \in K, \quad (22)$$

$$O_k - \sum_{p=1}^L y_{p0}^k = 0, \quad \forall k \in K, \quad (23)$$

$$u_i - u_j + L y_{ij}^k \leq L - 1, \quad u_i, u_j \geq 0, i, j \in N, i \neq j, j \geq 2, \quad (24)$$

$$\sum_{i=1}^L \sum_{j=1}^L q_i y_{ij}^k \leq G_k O_k, \quad \forall k \in K, \quad (25)$$

$$\text{or } \sum_{i=1}^L q_i x_i^k \leq G_k, \quad \forall k \in K,$$

$$\sum_{i=1}^L q_i x_i^k = \sum_{i=1}^L \sum_{j=1}^L q_i y_{ij}^k, \quad \forall k \in K, \quad (26)$$

$$(1 - 2O_k)(T_h - T_k) \leq 0, \quad \forall k \in V. \quad (27)$$

Constraint (28) is the requirement for delivery quality. M_k is calculated by equation (29). ε_n is the rate constant corresponding to the temperature z_k , calculated by the Arrhenius Equation (see equation (30)). M_0 denotes the initial quality. t_i^k denotes the service time of vehicle k at point

i, u_{ij}^k , denotes the transportation time of the vehicle k between i and j :

$$O_k(M_k - M^*) \geq 0, \quad \forall k \in K, \quad (28)$$

$$M_k = M_0 - \varepsilon_n \left(\sum_{i=1}^L t_i^k x_i^k + \sum_{i=1}^L \sum_{j=1}^L u_{ij}^k y_{ij}^k \right), \quad \forall k \in K, \quad (29)$$

$$k = k_0 e^{-\frac{E_a}{RT}}. \quad (30)$$

2.4. Implementation of the Collaborative Optimization Mechanism. Based on the above models, this paper implements the collaborative optimization mechanism by setting parameter iterations. The delivery quality requirement determined by the pricing model is used as the constraint of the distribution planning model. That is, the quality of products delivered to customers cannot be lower than the delivery quality requirements determined by the pricing model. The price determined by the pricing model is also used as the input parameter in the distribution planning model. The unit distribution cost obtained by the distribution planning model is used as the input parameter of the pricing model. Through iterative calculation, when the unit distribution cost obtained by the distribution planning model and the unit distribution cost of the pricing model tend to be consistent, the group purchase discount pricing strategy and cold chain distribution costs have achieved the collaborative optimization.

The iteration process is shown in Figure 3. At first, we consider an initial unit cold chain logistics cost and calculate the price, quality requirements, and market demand by the upper level. The lower level is used to obtain the distribution plan and the output unit cold chain cost under the constraints of the decision results calculated by the upper level. Then we can compare the unit cold chain logistics cost of the distribution scheme decided by the lower level with the assumed initial cold chain logistics cost. If it is consistent, iterative completion is completed. Otherwise, the unit cold chain logistics cost calculated by the lower level is used to replace the initial unit cold chain logistics cost and does the next round iteration.

3. Results and Discussion

3.1. Parameter Settings and Computation Results. In this section, the purpose price and planning are evaluated through a number of numerical tests. We develop a small-sized test set consisting of 6 orders based on the real setting of the community in Kunming. The parameter settings of the pricing model are shown in Table 1. In the actual operation situation of the community group purchase, generally the operator of the community will extract 10% of the sales. Therefore, the proportion of the operator is set as 0.1 in this paper. The initial unit cost is assumed value, and the optimal value is obtained through continuous iteration of the model. According to the community size, the potential market

demand rate is assumed to be 600. The influence of the potential market demand rate on the objective function and decision variables will be analysed later. The price constant is the average selling price of strawberries. According to curve-fitting, coefficients α_1 , α_2 , and α_3 can make the variation trend of various functions conform to the expected value within the value range. The value of group purchase time refers to the actual community group purchase sales time of 24 hours.

This paper assumes that a community group-buying enterprise sells strawberries, which is a kind of fresh agricultural product that is popular among people in different ages. It is equipped with a distribution center to carry out community group purchase activities in 6 large communities of a city and provide them with logistics and distribution services. Through the website, we obtained the information of latitude and longitude, construction age, total number of households, second-hand housing sources, rental housing sources, and other information of 100 communities in a city. Through the selection of information such as construction age, location of communities, and number of households, we finally determined 6 points as the service points of a community group-buying enterprise in this paper. The order quantity of each community is determined by the total number of households in the community, the occupancy rate combined with the online shopping utilization rate, and the proportion of fresh online shopping users who often buy fruits. Consider a real-world distribution system with the following data: $T_h = 22$, $\delta = 0.72$, $f = 0.7$, and $v_l = 30$. The other parameters of the vehicles are shown in Table 2.

In the case of small scale, we coded the above procedure using Lingo 12.0 and ran it on a PC with 3.20 GHz CPU and 8G memory to validate the efficiency of the solution method and the effectiveness of the proposed model. This section takes fresh fruit in an area of Yunnan Province as an example for analysis and presents the experimental settings and results. The optimal result can be obtained after four iterations and the process is shown in Table 3. At first, we suppose that the initial cold chain cost is 15 and we can calculate the quality and price by the upper level pricing model. Then, we put these two parameters into the lower level of the model and we can get a distribution scheme. After the first iteration, we get a unit cold chain cost which is 7.6025. Then, we use 7.6025 to replace the initial cold chain cost, 15, and make the second iteration until we can get the same value of the unit cold chain cost.

It can be seen from Table 4 that, in this group of calculation examples, the selling price of strawberries is 65.052. When we sold them at this price, if the potential market demand rate was 600, the sales profit could reach 63,440. To complete the delivery of 6 community points in the distribution center, three refrigerated vehicles are needed for the delivery of group purchase order quantity. The distribution path of vehicle 1 is 2-4-5-6, and the refrigeration temperature of the vehicle is about 2°C. Vehicle 2 just services demand node 1, and the refrigeration temperature of this vehicle is about 10°C. The distribution path of vehicle 3 is demand point 3, and the refrigeration temperature of the vehicle is about 14°C. The total delivery cost to complete the

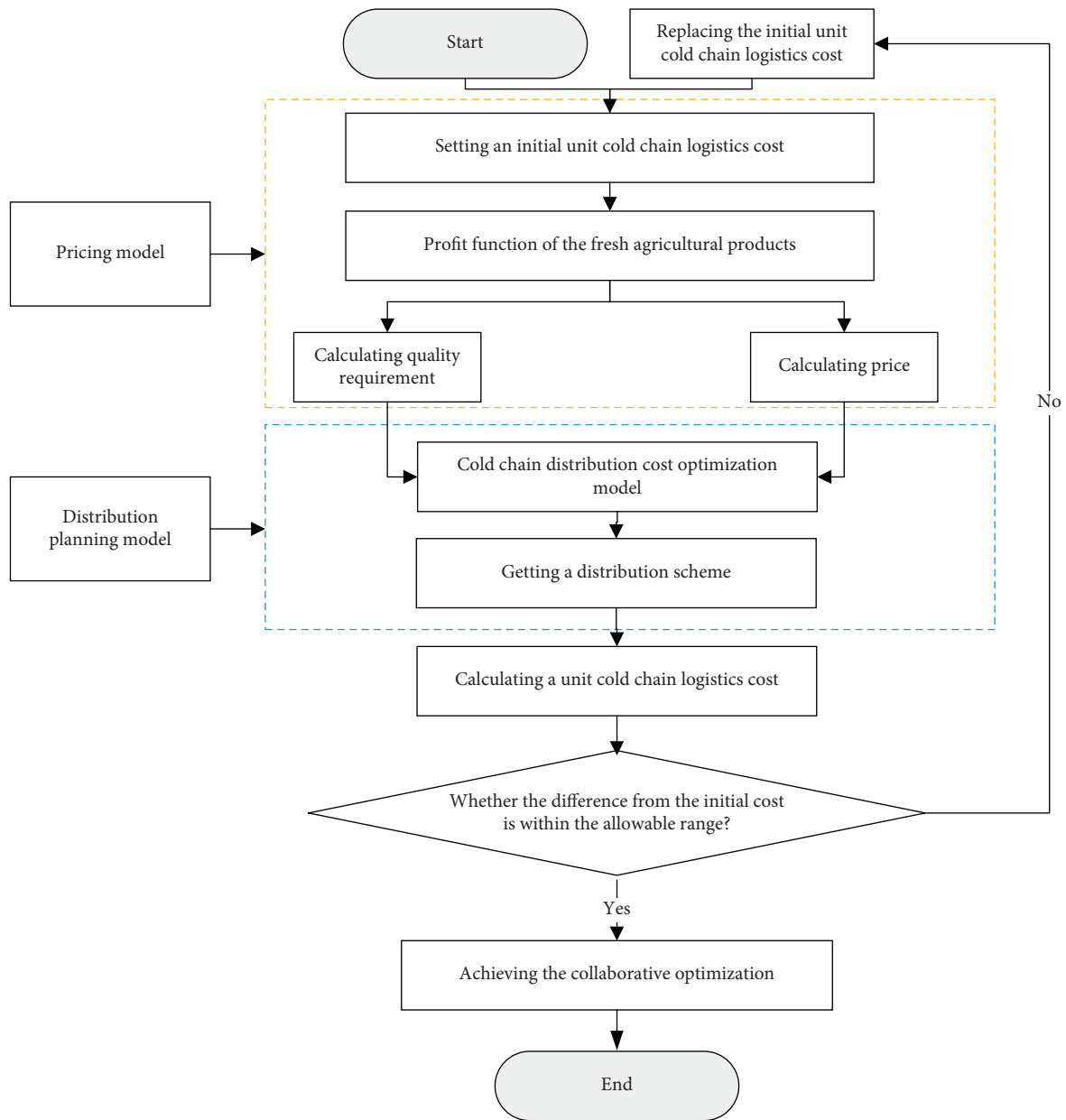


FIGURE 3: Collaborative optimization mechanism of community group purchase pricing model and cold chain distribution cost optimization model.

TABLE 1: Parameter settings of pricing model.

Parameter	Value
θ	0.1
C_d	15
C_p	16
D_0	600
α_0	12
α_1	0.5
α_2	7
α_3	0.8
M_E	45
T	24

actual order quantity of the six communities through the above distribution scheme is 3,366.

In order to illustrate the effect of the parameters on the optimal price and quality, the sensitivity analysis is performed by changing the value of only one parameter at a time and keeping the rest of the parameters at their initial values. These parameters are of significant influences on price and quality in practice but the specific impact is unknowable at present. Furthermore, the analysis of these parameters can reach valuable conclusions; thus, we can provide theoretical and practical reference for fresh agricultural community group purchase enterprises in operation.

TABLE 2: Parameter settings of distribution model.

Parameter	Value		
k	1	2	3
f	268	368	468
G	1485	3500	4830
V	12.68	22.04	23.06
a_0	0.012	1.369	1.749
b_0	0.061	0.061	0.061
c_0	0.014	0.014	0.014
w	2680	4580	4950
μ	0.08	0.08	0.08
M	49.38	68.68	75.68
ρ	0.02	0.02	0.02
η	5.034	7.5608	7.8362
λ	1.067	1.4834	1.6346
M_W	69.63	92.29	106.26
M_N	35.02	51.11	53.9

TABLE 3: Collaborative optimization process.

Upper level			Lower level					M^*	
C_d	M_k	p	Vehicle		Temperature				
			1	2	3	T_1	T_2	T_3	
Iteration 1									
15	75.210	69.956	3-4-5-6	2	1	1.27	12.30	9.75	76.210
The result of the unit cold chain cost is 7.6025. Replace the initial one									
Iteration 2									
7.6025	75.210	65.956	3-4-5-6	1	2	1.27	9.75	12.30	76.210
The result of the unit cold chain cost is 7.5730. Replace the initial one									
Iteration 3									
7.5730	74.388	65.053	2-4-5-6	1	3	1.65	10.31	14.10	75.39
The result of the unit cold chain cost is 7.5649. Replace the initial one									
Iteration 4									
7.5649	74.387	65.052	2-4-5-6	1	3	1.66	10.31	14.10	75.39
The result of the unit cold chain cost is 7.5649. End									

TABLE 4: Collaborative optimization results.

Vehicle	Route	Temperature	M^*	p^*	π	C_d
1	2-4-5-6	1.66				
2	1	10.31	75.39	65.052	63440	3366.414
3	3	14.10				

3.2. Impact of α_1 and α_2 on Price and Quality. As illustrated in Figure 4, by increasing the coefficient α_1 from 0.46 to 0.56 with an interval of 0.02, optimal price and quality will decrease. It implies that the increase in α_1 can influence the impact of quality changes on price. The quality decreases gradually; however, the price changes remain relatively stable.

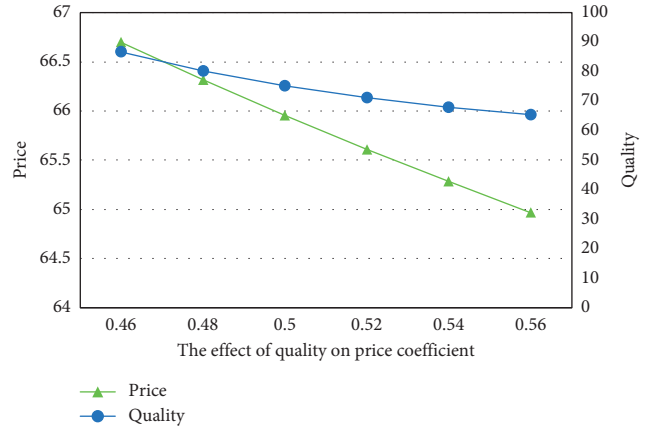


FIGURE 4: Change curves of optimal price and quality with α_1 .

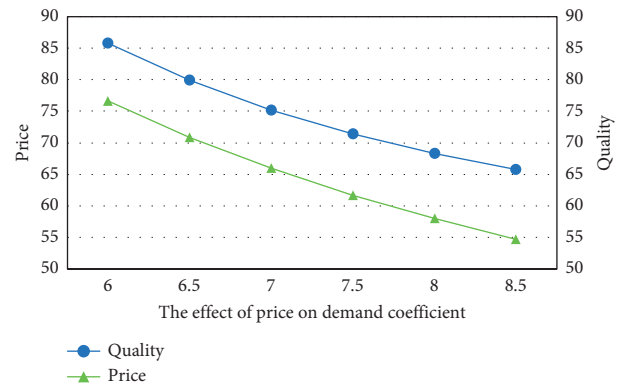


FIGURE 5: Change curves of optimal price and quality with α_2 .

The increase of α_2 means that the customer is more sensitive to the price. It is more likely that the customer will reduce purchase intention with the increase of α_2 . From Figure 5, we can observe that, by increasing α_2 from 6 to 8.5, continuous decreases in optimal price and quality can be achieved. When α_2 is small, the customer may not care about the price, and the operator will sell at a higher price in order to pursue higher profits, which in turn will guarantee a higher quality. However, in order to ensure the purchase intention of the customer, with the growth of α_2 , the operator needs not only to guarantee the corresponding quality but also to provide a lower price.

3.3. Impact of Distribution Costs on Pricing and Quality. In this section, we suppose that the unit distribution cost is increased from 4 to 20 with an interval of 2, consisting of 9 groups of data, and the changes of optimal sales pricing and quality are obtained as shown in Figure 6. It is not difficult to see that, with the increase of distribution cost, the selling price and quality of fresh agricultural products to consumers are also increasing. If the operator is willing to pay higher costs in the process of distribution, using small batches and low temperature or provide fast distribution service, the quality of the fresh agricultural products will be significantly improved, but consumers also have to pay a higher price.

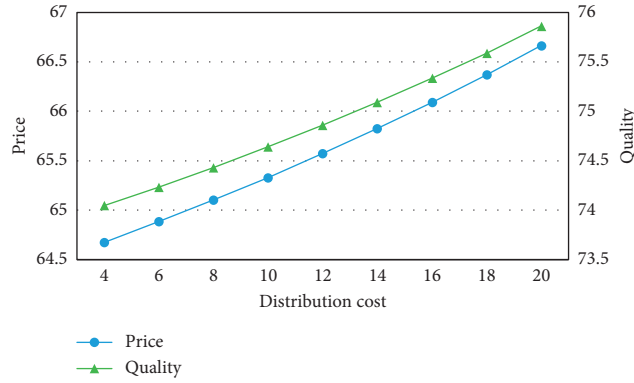


FIGURE 6: Change curves of optimal price and quality with the unit distribution cost.

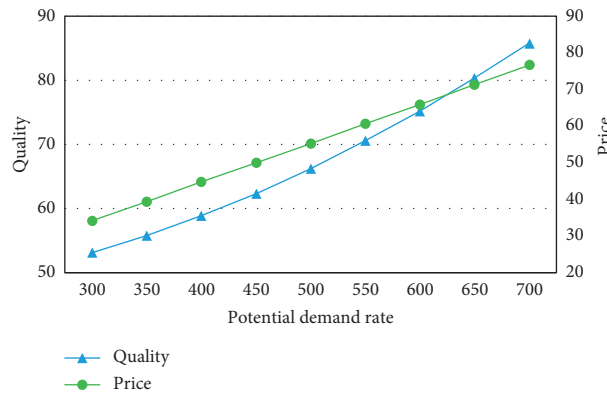


FIGURE 7: Change curves of optimal price and quality with the potential demand rate.

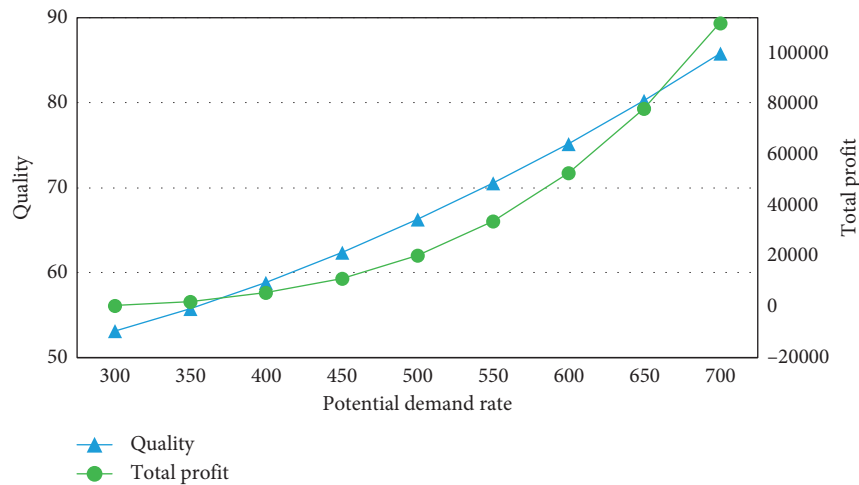


FIGURE 8: Change curves of total profit and quality with the potential demand rate.

3.4. Impact of Potential Rate on Quality, Price, and Total Profit.

In this section, we suppose that the market potential demand rate is increased from 350 to 700 with an interval of 50, which is composed of 9 groups of data. The changes of optimal price and quality are shown in Figure 7, and the changes of total profit and quality are shown in Figure 8. The

quality of fresh agricultural products increases with the increase of market potential demand rate. When the potential demand rate increases, the actual market demand will also increase. When the demand rate is too small, the demand will be far less than the distribution cost. If the vehicle with the minimum load is selected for distribution, the profit

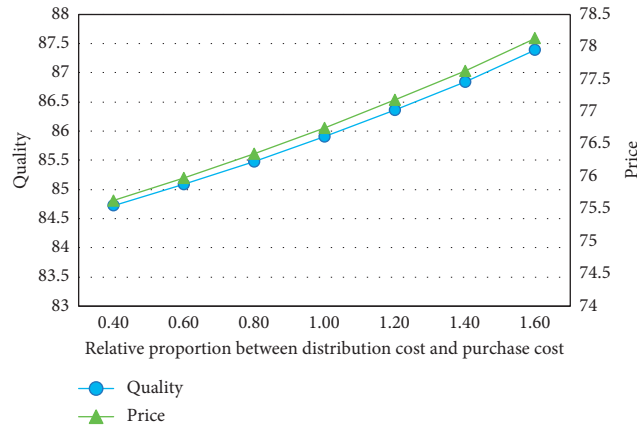


FIGURE 9: Change curves of total price and quality with relative proportion between distribution cost and purchase cost.

will be negative in such situation. When the demand is large enough to meet the maximum vehicle load, the operator will timely arrange multiple batches of distribution, so as to ensure a high level of quality. At the same time, the sales price increases with the increase of the market potential demand rate, which is consistent with the general market rules. Also, it is easy to see that the price trends are consistent with quality, which is mentioned in Proposition 2. Therefore, from the perspective of the operator’s benefits, when the demand increases, the operator should appropriately increase the sales price of fresh agricultural products to obtain the maximum profit under the potential demand during the sales period. At the same time, it can be seen from the above analysis that, in order to meet higher profits, the operators will increase the distribution cost. Therefore, it is suggested that the operators may consider sacrificing part of the profit value to reduce the total logistics cost and optimizing the logistics distribution scheme to improve customer satisfaction when making price decisions, which may result in a high total profit level.

3.5. Impact of Relative Proportion between Distribution Cost and Purchase Cost on Price and Quality. In this section, we suppose that the relative proportion between cold chain distribution cost and purchase cost is increased from 0.4 to 1.6 with an interval of 0.2, which is composed of 7 groups of data. The increase of the relative proportion between distribution cost and purchase cost can be seen from Figure 9. By comparing these two curves, it is implied that the price changes more than quality with the relative proportion between distribution cost and purchase cost. In other words, it is consistent with the real case where online fresh agricultural enterprise should raise its pricing to cope with rising costs and maintain a certain level of profitability.

4. Conclusions

The purpose of this study is to give a collaborative optimization mechanism of community group purchase pricing model and cold chain distribution cost model. Moreover, we use a two-layer mathematical programming method to

search for the optimal solution of this problem with feasibility and effectiveness. The numerical experiment of the model is implemented based on the real-world data in Kunming (China) and the optimal pricing and distribution scheme is obtained through iterations.

Through the parameters impact analysis, the following conclusions are drawn. First of all, the sales price increases with the increase of the market potential demand rate. Second, the quality and price sensitivity of fresh agricultural products have a significant impact on the price and demand and further influence the total profit. During the cold chain transportation services, the quality of fresh agricultural products will drop with the increase of the time and temperature. Third, the results show that the vast majority of cases tend to choose the cooler temperature, not only to guarantee the quality of strawberries but also to keep strawberry’s higher selling price and increase total profits. Of course, sometimes they may choose the high temperature distribution, of which vehicle is with little delivery quantity. This decision result is to satisfy the constraint conditions and objective function after the result of the decision. Under the optimal path solution, even distribution temperature is higher, but the mutual influence among sales price, distribution scheme, and quantity still can guarantee the quality of the fresh products. Finally, the selling price and quality of fresh agricultural products to consumers are also increasing with the increase of distribution cost.

Therefore, it is suggested that the community group purchase enterprise should take the factors including temperature, quality, potential demand rate, and time of duration into consideration to make a more reasonable decision on price and distribution scheme. In the future, we will apply robust optimization algorithms to solve the problem with the situation existing in more communities. We also consider presale strategy used in the community group purchase by dividing the sale into one presale period and one regular sale period.

Appendix

Proof of Proposition 5.

Proposition A.1. *When a seller chooses a community group-buying strategy, its unique optimal solution is*

$$p^* = \frac{(\alpha_1 + 1)[(1 - \theta)\alpha_0 D_0 + \alpha_0 \alpha_2 (C_d + C_p)] + \sqrt{(\alpha_1 + 1)^2 [(1 - \theta)\alpha_0 D_0 - \alpha_0 \alpha_2 (C_d + C_p)]^2 + 4(1 - \theta)\alpha_0^2 \alpha_1^2 \alpha_2 (C_d + C_p) D_0}}{2(2\alpha_1 + 1)(1 - \theta)\alpha_0 \alpha_2}. \quad (\text{A.1})$$

Proof. In order to simplify calculation, we suppose that $M = M_k - M_E$, and

$$\begin{aligned} \frac{d\pi(p)}{dM} &= \left\{ [(1 - \theta)\alpha_0 M^{\alpha_1} - C_d - C_p] T \alpha_3 \frac{M}{M_E} [D_0 - \alpha_2 \alpha_0 M^{\alpha_1}] \right\}, \\ &= \left\{ [(1 - \theta)\alpha_0 M^{\alpha_1} - C_d - C_p] \left[D_0 T \alpha_3 \frac{M}{M_E} - T \alpha_3 \alpha_2 \alpha_0 M^{\alpha_1} \frac{M}{M_E} \right] \right\}, \\ &= \left\{ (1 - \theta) D_0 T \alpha_0 \alpha_3 \frac{1}{M_E} M^{\alpha_1+1} - (C_d + C_p) D_0 T \alpha_3 \frac{M}{M_E} - (1 - \theta) T \alpha_2 \alpha_3 \alpha_0^2 M^{2\alpha_1+1} \frac{1}{M_E} + (C_d + C_p) T \alpha_2 \alpha_3 \alpha_0 M^{\alpha_1+1} \frac{1}{M_E} \right\}, \\ &= \left\{ \alpha_3 T \frac{1}{M_E} [(1 - \theta) D_0 \alpha_0 M^{\alpha_1+1} - (C_d + C_p) D_0 M - (1 - \theta) \alpha_2 \alpha_0^2 M^{2\alpha_1+1} + (C_d + C_p) \alpha_2 \alpha_0 M^{1+\alpha_1}] \right\}, \\ &= \alpha_3 T \frac{1}{M_E} [(\alpha_1 + 1)(1 - \theta) D_0 \alpha_0 M^{\alpha_1} - (C_d + C_p) D_0 - (2\alpha_1 + 1)(1 - \theta) \alpha_2 \alpha_0^2 M^{2\alpha_1} + (\alpha_1 + 1)(C_d + C_p) \alpha_2 \alpha_0 M^{\alpha_1}] \\ &= \alpha_3 T \frac{1}{M_E} \{ -(2\alpha_1 + 1)(1 - \theta) \alpha_2 \alpha_0^2 M^{2\alpha_1} + (\alpha_1 + 1)[(1 - \theta) \alpha_0 D_0 + \alpha_0 \alpha_2 (C_d + C_p)] M^{\alpha_1} - (C_d + C_p) D_0 \}. \end{aligned} \quad (\text{A.2})$$

When $d\pi(p)/dM = 0$,

$$-(2\alpha_1 + 1)(1 - \theta) \alpha_2 \alpha_0^2 M^{2\alpha_1} + (\alpha_1 + 1)[(1 - \theta) \alpha_0 D_0 + \alpha_0 \alpha_2 (C_d + C_p)] M^{\alpha_1} - (C_d + C_p) D_0 = 0. \quad (\text{A.3})$$

Suppose that $M^{\alpha_1} = x$; we can find

$$\begin{aligned}
& -(2\alpha_1 + 1)(1 - \theta)\alpha_2\alpha_0^2x^2 + (\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)]x - (C_d + C_p)D_0 = 0, \\
\Delta &= \{(\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)]\}^2 - 4 \times [-(2\alpha_1 + 1)(1 - \theta)\alpha_2\alpha_0^2] \{-(C_d + C_p)D_0\} \\
&= (\alpha_1 + 1)^2 [(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)]^2 - 4(C_d + C_p)D_0 [(2\alpha_1 + 1)(1 - \theta)\alpha_2\alpha_0^2] \\
&= (\alpha_1 + 1)^2 \left[(1 - \theta)^2\alpha_0^2D_0^2 - 2(1 - \theta)\alpha_0^2\alpha_2(C_d + C_p)D_0 + \alpha_0^2\alpha_2^2(C_d + C_p)^2 + 4(1 - \theta)\alpha_0^2\alpha_2(C_d + C_p)D_0 \right] \\
&\quad - 8(1 - \theta)\alpha_0^2\alpha_1\alpha_2(C_d + C_p)D_0 - 4(1 - \theta)\alpha_0^2\alpha_2(C_d + C_p)D_0 \\
&= (\alpha_1 + 1)^2 [(1 - \theta)\alpha_0D_0 - \alpha_0\alpha_2(C_d + C_p)]^2 + 4(1 - \theta)(\alpha_1 + 1)^2\alpha_0^2\alpha_2(C_d + C_p)D_0 \\
&\quad - 8(1 - \theta)\alpha_0^2\alpha_1\alpha_2(C_d + C_p)D_0 - 4(1 - \theta)\alpha_0^2\alpha_2(C_d + C_p)D_0 \\
&= (\alpha_1 + 1)^2 [(1 - \theta)\alpha_0D_0 - \alpha_0\alpha_2(C_d + C_p)]^2 + 4(1 - \theta)\alpha_0^2\alpha_2(C_d + C_p)D_0 [(\alpha_1 + 1)^2 - 2\alpha_1 - 1] \\
&= (\alpha_1 + 1)^2 [(1 - \theta)\alpha_0D_0 - \alpha_0\alpha_2(C_d + C_p)]^2 + 4(1 - \theta)\alpha_0^2\alpha_1^2\alpha_2(C_d + C_p)D_0 > 0, \\
x_1 &= \frac{(\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)] + \sqrt{(\alpha_1 + 1)^2 [(1 - \theta)\alpha_0D_0 - \alpha_0\alpha_2(C_d + C_p)]^2 + 4(1 - \theta)\alpha_0^2\alpha_1^2\alpha_2(C_d + C_p)D_0}}{2(2\alpha_1 + 1)(1 - \theta)\alpha_0^2\alpha_2}, \\
x_2 &= \frac{(\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)] - \sqrt{(\alpha_1 + 1)^2 [(1 - \theta)\alpha_0D_0 - \alpha_0\alpha_2(C_d + C_p)]^2 + 4(1 - \theta)\alpha_0^2\alpha_1^2\alpha_2(C_d + C_p)D_0}}{2(2\alpha_1 + 1)(1 - \theta)\alpha_0^2\alpha_2}.
\end{aligned} \tag{A.4}$$

M_1 can be obtained as

$$M_1 = e^{\frac{\ln \left[(\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)] + \sqrt{(\alpha_1 + 1)^2 [(1 - \theta)\alpha_0D_0 - \alpha_0\alpha_2(C_d + C_p)]^2 + 4(1 - \theta)\alpha_0^2\alpha_1^2\alpha_2(C_d + C_p)D_0} \right]}{\alpha_1 - \ln [2(2\alpha_1 + 1)(1 - \theta)\alpha_0^2\alpha_2]}}. \tag{A.5}$$

Here, we find that

$$(\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)] < \sqrt{(\alpha_1 + 1)^2 [(1 - \theta)\alpha_0D_0 - \alpha_0\alpha_2(C_d + C_p)]^2 + 4(1 - \theta)\alpha_0^2\alpha_1^2\alpha_2(C_d + C_p)D_0}. \tag{A.6}$$

When $M > 0$, $d\pi(p)/dM = 0$ has only one root. When $0 < M < M_1$, $\partial\pi(p)/\partial M > 0$; $\pi(p)$ is monotonically increasing. The profit increases with the increasing of M :

$$\begin{aligned}
\frac{d^2\pi(p)}{d^2M} &= \left\{ \alpha_3 T \frac{1}{M_E} \left\{ -(2\alpha_1 + 1)(1 - \theta)\alpha_2\alpha_0^2 M^{2\alpha_1} + (\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)] M^{\alpha_1} - (C_d + C_p)D_0 \right\} \right\}, \\
&= \alpha_3 T \frac{1}{M_E} \left\{ -2\alpha_1(2\alpha_1 + 1)(1 - \theta)\alpha_2\alpha_0^2 M^{2\alpha_1 - 1} + \alpha_1(\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)] M^{\alpha_1 - 1} \right\}.
\end{aligned} \tag{A.7}$$

When $d^2\pi(p)/d^2M = 0$,

$$\begin{aligned}
 & -2\alpha_1(2\alpha_1 + 1)(1 - \theta)\alpha_2\alpha_0^2M^{2\alpha_1 - 1} + \alpha_1(\alpha_1 + 1) \\
 & \cdot [(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)]M^{\alpha_1 - 1} = 0, \\
 & 2(2\alpha_1 + 1)(1 - \theta)\alpha_2\alpha_0^2M^{2\alpha_1 - 1} = (\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)]M^{\alpha_1 - 1}, \\
 & \frac{M^{2\alpha_1 - 1}}{M^{\alpha_1 - 1}} = \frac{(\alpha_1 + 1)[(1 - \theta)D_0 + \alpha_2(C_d + C_p)]}{2(2\alpha_1 + 1)(1 - \theta)\alpha_2\alpha_0}, \tag{A.8} \\
 & M^{\alpha_1} = \frac{(\alpha_1 + 1)[(1 - \theta)D_0 + \alpha_2(C_d + C_p)]}{2(2\alpha_1 + 1)(1 - \theta)\alpha_2}, \\
 & M = e^{\ln\{(\alpha_1 + 1)[(1 - \theta)D_0 + \alpha_2(C_d + C_p)]\}/\alpha_1 - \ln[2(2\alpha_1 + 1)(1 - \theta)\alpha_2\alpha_0]/\alpha_1}.
 \end{aligned}$$

We can find that when $M > e^{\ln\{(\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)]\}/\alpha_1 - \ln[2(2\alpha_1 + 1)(1 - \theta)\alpha_2\alpha_0^2]/\alpha_1}$, $\pi(p)$ is a convex function. When $0 < M < e^{\ln\{(\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)]\}/\alpha_1 - \ln[2(2\alpha_1 + 1)(1 - \theta)\alpha_2\alpha_0^2]/\alpha_1}$, $\pi(p)$ is a concave function.

When $0 < M < M_1$, $d\pi(p)/dM > 0$ and $0 < e^{\ln\{(\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)]\}/\alpha_1 - \ln[2(2\alpha_1 + 1)(1 - \theta)\alpha_2\alpha_0^2]/\alpha_1} < M_1$.

We can find that

$$\begin{aligned}
 & \max \pi\left(M = e^{\ln\{(\alpha_1 + 1)(D_0 + \alpha_2(C_d + C_p))\}/\alpha_1 - \ln[2(2\alpha_1 + 1)\alpha_0\alpha_2]/\alpha_1}\right) \\
 & < \max \pi(M = M_1). \tag{A.9}
 \end{aligned}$$

Then, we have

$$\max \pi(M = 0) = 0 < \max \pi(M = M_1). \tag{A.10}$$

We can easily find that when $M > 0$, the maximum profit level is $\max \pi(M = M_1)$. The optimal quality exists and is unique; that is, $M^* = M_1$. The corresponding M_k^* and p^* are shown as follows:

$$\begin{aligned}
 M_k^* & = M_1 + M_E = e^{\ln\left\{(\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)]/\alpha_1 + \sqrt{(\alpha_1 + 1)^2[(1 - \theta)\alpha_0D_0 - \alpha_0\alpha_2(C_d + C_p)]^2 + 4(1 - \theta)\alpha_0^2\alpha_1^2\alpha_2(C_d + C_p)D_0}\right\}/\alpha_1 - \ln[2(2\alpha_1 + 1)(1 - \theta)\alpha_0^2\alpha_2]/\alpha_1} \\
 & \quad + M_E, \\
 p^* & = \frac{(\alpha_1 + 1)[(1 - \theta)\alpha_0D_0 + \alpha_0\alpha_2(C_d + C_p)] + \sqrt{(\alpha_1 + 1)^2[(1 - \theta)\alpha_0D_0 - \alpha_0\alpha_2(C_d + C_p)]^2 + 4(1 - \theta)\alpha_0^2\alpha_1^2\alpha_2(C_d + C_p)D_0}}{2(2\alpha_1 + 1)(1 - \theta)\alpha_0\alpha_2}. \tag{A.11}
 \end{aligned}$$

Thus, we can find the unique optimal community group purchase price for this problem. \square

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

The authors acknowledge the National Natural Science Foundation of China (Grant no. 71462024).

References

- [1] G. Zhang, J. Shang, and P. Yildirim, "Optimal pricing for group buying with network effects," *Omega*, vol. 63, no. 8, pp. 69–82, 2016.
- [2] R. J. Kauffman and B. Wang, "New buyers' arrival under dynamic pricing market microstructure: the case of group-

- buying discounts on the Internet,” *Journal of Management Information Systems*, vol. 18, no. 2, pp. 157–188, 2001.
- [3] J. Chen, Y. Liu, X. Song et al., “A study of Single-vendor and multiple-retailers pricing-ordering strategy under group-buying online auction,” in *Proceedings of the International Conference on Electronic Business—Shaping Business Strategy in A Networked World (ICEB)*, pp. 19–23, Beijing, China, 2004.
- [4] J. Chen, X. Chen, and X. Song, “Comparison of the group-buying auction and the fixed pricing mechanism,” *Decision Support Systems*, vol. 43, no. 2, pp. 445–459, 2007.
- [5] M. Sharif Vaghefi, M. Sharif Vaghefi, N. Beheshti et al., “A pricing model for group-buying auction based on customers’ waiting-time,” *Marketing Letters*, vol. 25, no. 4, pp. 425–434, 2014.
- [6] W. Gu and X. Cai, “Application of group buying pricing in two—echelon distribution systems,” in *Proceedings of the 13th International Conference on Service Systems and Service Management (ICSSSM)*, IEEE, Kunming, China, pp. 1–6, August 2016.
- [7] F. C. Tang and K. P. Chi, “Research on pricing strategy of online group buying in two-sided network,” *Chinese Journal of Management Science*, vol. 21, no. 3, pp. 185–192, 2013.
- [8] D. Qian and H. Su, “Joint decision of pricing and advertising investment for online group-buying under intersecting demand regimes,” *Journal of Applied Sciences*, vol. 13, no. 22, pp. 4949–4954, 2013.
- [9] C.-I. Hsu, S.-F. Hung, and H.-C. Li, “Vehicle routing problem with time-windows for perishable food delivery,” *Journal of Food Engineering*, vol. 80, no. 2, pp. 465–475, 2007.
- [10] A. Osvald and L. Z. Stirn, “A vehicle routing algorithm for the distribution of fresh vegetables and similar perishable food,” *Journal of Food Engineering*, vol. 85, no. 2, pp. 285–295, 2008.
- [11] P. Amorim, S. N. Parragh, F. Sperandio, and B. Almada-Lobo, “A rich vehicle routing problem dealing with perishable food: a case study,” *Top*, vol. 22, no. 2, pp. 489–508, 2014.
- [12] A. Estrada-Moreno, C. Fikar, A. Angel, Juan et al., “A biased-randomized algorithm for redistribution of perishable food inventories in supermarket chains,” *International Transactions in Operational Research*, vol. 26, no. 6, pp. 2077–2095, 2019.
- [13] X. Wang, M. Wang, J. Ruan, and H. Zhan, “The multi-objective optimization for perishable food distribution route considering temporal-spatial distance,” *Procedia Computer Science*, vol. 96, no. 1, pp. 1211–1220, 2016.
- [14] Y.-H. Hsiao, M.-C. Chen, and C.-L. Chin, “Distribution planning for perishable foods in cold chains with quality concerns: formulation and solution procedure,” *Trends in Food Science & Technology*, vol. 61, no. 1, pp. 80–93, 2017.
- [15] Y.-H. Hsiao, M.-C. Chen, K.-Y. Lu, and C.-L. Chin, “Last-mile distribution planning for fruit-and-vegetable cold chains,” *The International Journal of Logistics Management*, vol. 29, no. 3, pp. 862–886, 2018.
- [16] L. Zhen, “Modeling of yard congestion and optimization of yard template in container ports,” *Transportation Research Part B*, vol. 90, pp. 83–104, 2016.
- [17] L. Zhen, “Tactical berth allocation under uncertainty,” *European Journal of Operational Research*, vol. 247, no. 3, pp. 928–944, 2015.

Research Article

Study on Coordination and Optimization of Contract Farming Supply Chain Based on Uncertain Conditions

Xinquan Liu ¹, Xiaojing Shen ¹ and Ming You²

¹School of Logistics Management and Engineering, Nanning Normal University, Nanning, Guangxi 530001, China

²Office of Human Resources, Nanning Normal University, Nanning, Guangxi 530001, China

Correspondence should be addressed to Xiaojing Shen; 614243786@qq.com

Received 28 May 2020; Revised 19 June 2020; Accepted 29 June 2020; Published 20 July 2020

Academic Editor: Lu Zhen

Copyright © 2020 Xinquan Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As the transaction subject of contract farming, agricultural products are featured with a long production cycle and a short sales cycle, just like other perishable commodities. In the process of executing a contract, both the company and the farmer are running the risk of great uncertainty. This paper studies the coordination of agricultural supply chain in terms of the uncertainty of agricultural output and market demand. First of all, the random output volatility factor and the market demand volatility factor as two random variables are used to represent the uncertainty of the agricultural output and market demand, and revenue functions are set up, respectively, for the company and the farmer with the objective of maximizing expected returns. The theoretical derivation of these revenue functions proves that there is an optimal targeted yield in a centralized decision-making supply chain system and a single optimal solution that maximizes farmers' revenue can be obtained in a decentralized one, but the centralized decision-making supply chain is superior to the decentralized and uncoordinated counterpart for overall benefit. Secondly, a revenue-sharing-plus-margin contract mechanism is proposed to coordinate income distribution between the two parties of the supply chain through the revenue-sharing coefficient and margin. Thirdly, calculation examples are given and solved by MATLAB based on the assumption that both the agricultural output volatility factor and the market demand volatility factor are uniformly distributed, and the theory and result are then verified consistently. Finally, the numerical analysis of the coordination mechanism of the revenue-sharing coefficient and the margin on both sides of the supply chain provides an optimal value range so that Pareto improvement on the company's and the farmers' income can be achieved.

1. Introduction

With the deepening reform of the agricultural product circulation system, the order-based agriculture supply chain has emerged and become prevailing in China since the 1990s. It is worth mentioning that the company-plus-smallholder type has also gained its popularity as a basic business mode, which refers to an order contract signed by the company as the processing and sales party and the farmer as the producer before the production of agricultural products, where the two parties agree on the type, quantity, quality, purchase price, time, and place of agricultural products to be traded. Farmers grow crops and then both parties conduct spot transactions of the produce at the end of the production cycle in accordance with the order contract.

The order contract between companies and farmers represents the market demand which functions as a guide for farmers to planned farming in avoidance of blind production, achieving a dynamic connection between the production of smallholder farmers and market, which explains the reason why the order-based agriculture supply chain has been developing.

However, there are problems coming with the rise of the order-based agricultural supply chain, one of which is that the high default rate is highly detrimental to the stability and growth of this supply chain. Hobbs and Young [1] believed the main reason for this problem lied in the necessity of reducing transaction costs after a probe into the vertical coordination in the supply chain based on the transaction cost theory. Martinez [2] also utilized the

theory to make an analysis on the influence factors of transaction cost which affect the order contract business mode of agricultural by-products and found that this mode was directly affected by asset specificity, product quality, and farmers' efforts. Guo and Jolly [3] found that the legal mechanism played a smaller role in improving contract performance than nonformal means on the basis of an investigation of order-based agriculture in Zhejiang Province, China. In terms of companies with dual-source raw materials and multiproduct combinations in the beef industry, Boyabatli et al. [4] assumed that long-term and short-term contract mechanisms could be adopted to improve their own effectiveness under specific circumstances. Based on the analysis of potato farmers' preference for given contract types in Ethiopia, Abebe et al. [5] found that the uncertainty of the import market rather than the export market was more likely to sway farmers' decision whether to be engaged in the business, and farmers were inclined to do business with buyers specialized in seeds, raw materials, and technical support from government and nongovernmental organizations in an attempt to reduce risks. Liu [6] made an analysis on the existing problems of sales contracts of in China led by the incomplete contract theory. They pointed out that the indeterminacy of contracts as well as the opportunism behavior tendency of both parties, namely, the incompleteness of contract caused by information asymmetry, is the root reason for the low-performance rate of sales contracts, and the market risks are proposed to be shouldered by both parties. Liu and Qi [7] applied transaction cost theory to make a quantitative analysis of the factors affecting the option of contracts between enterprises and farmers and proposed that the final contract solution should be bounded and acceptable to both parties so as to improve the stability and performance rate of contracts. Wu et al. [8] discussed the payment costs incurred by farmers and leading enterprises in different game behaviors under the conditions of good, medium, and bad market price fluctuations, proving that the increase of liquidated damages is a means to guarantee the stability of contracts. Fang [9] investigated the conditions for managing default risk under the order-plus-futures mode by establishing the market equilibrium model before and after the introduction of futures trading, which was thought to be conducive to stabilizing the spot price and improve the performance rate of contracts. Based on 381 pseudocopies of order contracts from Shaanxi Province, the research team of Xi'an Branch of the People's Bank of China [10] analyzed the factors affecting the contract performance rate under the direction of logistic and made a proposal that the contract performance rate could be raised by means of cultivating the order subject, standardizing the contract text and innovating financial products. Yang et al. [11] analyzed the default risk under different modes of contract farming, and the results showed that a higher conversion rate of agricultural products processing and default penalty coefficient contributed to increasing the performance rate of contracts.

As the transaction subject of contract farming, agricultural products are featured with a long production

cycle and a short sales cycle, just like other perishable commodities. In the process of executing a contract, companies and farmers are both taking a great risk. The company will face market demand uncertainty if order contracts with farmers are signed before the production of agricultural products, while farmers will have to confront output uncertainty due to the adverse impact of environmental factors such as a terrible climate or pest disasters. Considering the weather problems encountered during the production of seasonal products, Chen and Yano [12–14] based on the measures of uncertainty put forward risk compensation to coordinate the contractual relationship between companies and farmers so as to ensure the stability of contract farming. Wang et al. [15] found through analysis that smallholders are expected to have independent management rights given their small size and poor resistance to risks, and they will only fulfill contracts if they are satisfied with existing transactions. On the premise of a uniform distribution of the fluctuation factors of yield and demand, Zhao and Wu [16] analyzed how to achieve the supply chain coordination of individual producers and retailers of produce and proposed that revenue-sharing contracts would be a solution. Ling et al. [17] put forward a solution that ensures a long-term cooperation between farmers and supermarkets under random yield and demand against the background of China's farmer-to-supermarket mode. Liu and Xie [18] came up with a risk-sharing contract to cope with the coordination between manufacturers and retailers in the agriculture supply chain based on the uncertainty of yield and market demand. Sun and Xu [19] brought up a two-tier supply chain model of agricultural products consisting of an individual supplier and retailer in a perfect competition market. And after studying option contracts in terms of stochastic supply and demand, they found better coordination of option contracts with the increased randomness of suppliers. Taking the company-plus-farmer order-based agriculture as the research object and the wholesale price of the traditional contract as the reference object, Qin and Teng [20] studied comprehensively the uncertain market demand faced by companies, the uncertain yield faced by farmers, and the fluctuations of wholesale price confronted by both parties and conducted an analysis of the performance of the supply chain after the implementation of the "Guaranteed Minimum Purchase and Following the Market" policy that was considered beneficial for farmers. Tong [21] analyzed the decision-making behavior of companies and farmers in terms of the price mechanism of traditional order-based agriculture under the conditions of random purchase prices and uncertain market demand and proposed a contract mechanism featured with revenue sharing plus two-way subsidy plus franchise fee to promote the coordination of order agriculture supply chain.

Referring to the above-mentioned relevant researches on the coordination of supply chain with uncertain yield and demand and the ideas of literature [18], this paper undertook a research in the case of company-farmer order agriculture into the production decision of the two parties of the supply

chain targeted at a maximized expected return and put forward a revenue-sharing-plus-margin coordination strategy and gave examples to further analyze the strategy in the premise of uncertain yield and market demand.

2. Assumptions and Parameters of the Model

Based upon the Newsvendor model, a company-plus-farmer contract farming supply chain was established in this paper characterized by one company, one smallholder, and one single production cycle. According to this model, the company predicts the market demand for the next sales cycle based on the past sales of agricultural products before the production season and signs an order contract with farmers, clarifying the quantity, quality, and purchase price of agricultural products. Farmers then determine the targeted yield based on the quantity and the purchase price in the contract. After the end of the production cycle of produce, purchase and selling of the produce will be carried out by both parties of the order contract.

2.1. Assumptions

- (1) It is a supply chain of one company plus one farmer. The information is symmetric between the company and the farmer.
- (2) There is only one agricultural product with one production-sale cycle in the supply chain.
- (3) The members of the supply chain are risk neutral.
- (4) The processing, storage, and transportation costs of agricultural products are not taken into account.
- (5) Take into account the company's coping strategy towards the insufficient supply in the contract farming supply chain but not the stock-out cost of farmers.

2.2. Parameters

- (1) q : the quantity of agricultural products confirmed in the order contract between the company and the farmer.
- (2) Q : the targeted yield of certain agricultural product.
- (3) x : the random fluctuation factor of the yield; $x \in [a_1, a_2]$, the probability density function is $f(\cdot)$, and $E(x) = \mu$ (the expectation of x). When the targeted yield is Q , the actual yield will be Qx .
- (4) c : the unit production cost of the agricultural product which is related to the targeted yield.
- (5) w : the purchase price for the produce.
- (6) h : the unit recovery value of the remaining produce that exceeds the agreed amount.
- (7) De : market demand, related to sales price and random fluctuation factor; $De = \alpha p^{-k} y$, α and $-k$ are constants, y is the random fluctuation factor of market demand, $y \in [b_1, b_2]$, and the probability density function is $g(\cdot)$. Let $D = \alpha p^{-k}$; that is, $De = Dy$.

- (8) p : sales price for the produce.
- (9) h_1 : the unit recovery value of the unsold produce.
- (10) s : unit out-of-stock cost that the company fails to meet market demand.
- (11) Without loss of generality, $h < c < w < h_1 < s < p$.

3. Production Decision Model and Analysis of Contract Farming Supply Chain

3.1. *Optimal Production Decision of Supply Chain under Centralized Decision-Making.* Under centralized decision-making, farmers and companies are a whole aiming at maximizing the overall return of the supply chain. In this case, the agricultural products produced by farmers are sold by the company. So here, the total revenue function of the supply chain is

$$I_0 = p \min\{Dy, Qx\} + h_1 (Qx - Dy)^+ - s (Dy - Qx)^+ - cQ, \quad (1)$$

$$\begin{aligned} E(I_0) = & p \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} Qx f(x) g(y) dx dy \\ & + p \int_{b_1}^{b_2} \int_{(Dy/Q)}^{a_2} Dy f(x) g(y) dx dy \\ & + h_1 \int_{b_1}^{b_2} \int_{(Dy/Q)}^{a_2} (Qx - Dy) f(x) g(y) dx dy \\ & - s \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} (Dy - Qx) f(x) g(y) dx dy - cQ. \end{aligned} \quad (2)$$

According to equation (2), the first and second partial derivatives of $E(I_0)$ with respect to Q can be obtained:

$$\begin{aligned} \frac{\partial E(I_0)}{\partial Q} = & p \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} x f(x) g(y) dx dy \\ & + h_1 \int_{b_1}^{b_2} \int_{(Dy/Q)}^{a_2} x f(x) g(y) dx dy \\ & + s \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} x f(x) g(y) dx dy - c, \end{aligned} \quad (3)$$

$$\begin{aligned} \frac{\partial^2 E(I_0)}{\partial Q^2} = & p \int_{b_1}^{b_2} \left(-\frac{Dy}{Q^2} \right) \frac{Dy}{Q} f\left(\frac{Dy}{Q}\right) g(y) dy \\ & - h_1 \int_{b_1}^{b_2} \left(-\frac{Dy}{Q^2} \right) \frac{Dy}{Q} f\left(\frac{Dy}{Q}\right) g(y) dy \\ & + s \int_{b_1}^{b_2} \left(-\frac{Dy}{Q^2} \right) \frac{Dy}{Q} f\left(\frac{Dy}{Q}\right) g(y) dy \\ = & -\frac{p + s - h_1}{Q^3} \int_{b_1}^{b_2} (Dy)^2 f\left(\frac{Dy}{Q}\right) g(y) dy. \end{aligned} \quad (4)$$

$(\partial^2 E(I_0)/\partial Q^2) < 0$ is constant, $E(I_0)$ is the concave function of Q , and here Q^* comes that makes $\max E(I_0)$. Let $(\partial E(I_0)/\partial Q) = 0$, and there is

$$(p + s) \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Qh)} xf(x)g(y)dx dy + h_1 \int_{b_1}^{b_2} \int_{(Dy/Qh)}^{a_2} xf(x)g(y)dx dy = c. \quad (5)$$

When c increases, Q' decreases; that is, when the unit production cost of produce increases, the targeted yield will decrease, which is reasonable.

3.2. Decentralized Production Decisions by Both Parties of the Supply Chain. An order contract will be signed by the company and the farmer to specify the amount and the purchase price of agricultural products that are agreed to be produced. Farmers decide the targeted yield of produce and put it into production based on the contract. The farmer then supplies agricultural products to the company after one production cycle and the company sells them to the market. The reverse induction is used for the following analysis. Firstly, we try to analyze the optimum condition for the decision variables to maximize the revenue of the farmer, to clear the relationship between Q and q . Secondly, based on the optimum condition, we analyze the decision-making of the company to confirm Q . Lastly, q is determined according to the relationship between the two parameters described above.

3.2.1. Farmers' Production Decisions under Decentralized Decision-Making. The revenue function under decentralized decision-making is

$$I_1 = w \min\{q, Qx\} + h(Qx - q)^+ - cQ. \quad (6)$$

In this equation, the first part is the amount paid for farmers by companies purchasing the produce; the second part is the recovery residual value of agricultural products that exceeds the contracted amount; and the third part is the production cost of farmers.

The expected revenue of farmers is

$$E(I_1) = w \int_{(q/Q)}^{a_2} qf(x)dx + w \int_{a_1}^{(q/Q)} Qxf(x)dx + h \int_{(q/Q)}^{a_2} (Qx - q)f(x)dx - cQ. \quad (7)$$

According to equation (7), the first partial derivatives of $E(I_1)$ with respect to Q and q can be obtained:

$$\frac{\partial E(I_1)}{\partial Q} = w \int_{a_1}^{(q/Q)} xf(x)dx + h \int_{(q/Q)}^{a_2} xf(x)dx - c, \quad (8)$$

$$\begin{aligned} \frac{\partial E(I_1)}{\partial q} &= w \int_{a_1}^{(q/Q)} f(x)dx - h \int_{(q/Q)}^{a_2} f(x)dx \\ &= (w - h) \int_{(q/Q)}^{a_2} f(x)dx. \end{aligned} \quad (9)$$

The Hesse Matrix of equation (7) is

$$\begin{aligned} H(Q, q) &= \begin{vmatrix} \frac{\partial^2 E(I_1)}{\partial Q^2} & \frac{\partial^2 E(I_1)}{\partial Q \partial q} \\ \frac{\partial^2 E(I_1)}{\partial q \partial Q} & \frac{\partial^2 E(I_1)}{\partial q^2} \end{vmatrix} \\ &= \begin{vmatrix} -\frac{(w-h)q^2}{Q^3} f\left(\frac{q}{Q}\right) & -\frac{(w-h)q}{Q^2} f\left(\frac{q}{Q}\right) \\ -\frac{(w-h)q}{Q^2} f\left(\frac{q}{Q}\right) & -\frac{w-h}{Q} f\left(\frac{q}{Q}\right) \end{vmatrix}. \end{aligned} \quad (10)$$

Let $A_1 = (\partial^2 E(I_1)/\partial Q^2)$, $A_2 = (\partial^2 E(I_1)/\partial Q \partial q)$, $A_3 = (\partial^2 E(I_1)/\partial q \partial Q)$, and $A_4 = (\partial^2 E(I_1)/\partial q^2)$, where $A_1 < 0$ and $A_1 A_4 - A_2 A_3 = ((2(w-h)^2 t q^2)/Q^2) f(q/Q) > 0$. Therefore, $H(Q, q)$ is negative definite matrix and $E(I_1)$ is strictly concave function and has a unique maximum point (Q^*, q^*) .

Let equations (8) and (9) be equal to 0, and the simultaneous equations can obtain Q^* (targeted yield) and q^* (contracted quantity), satisfying the following equation:

$$\begin{cases} \int_{a_1}^{(q^*/Q^*)} xf(x)dx = \frac{c - \mu h}{w - h}, \\ (w - h) \int_{(q^*/Q^*)}^{a_2} xf(x)dx = 0. \end{cases} \quad (11)$$

Since the equation $(w - h) > 0$ is constant and $(w - h) \int_{(q^*/Q^*)}^{a_2} xf(x)dx = 0$, we can have $(q^*/Q^*) \geq a_2$; $q^* \geq a_2 Q^*$.

In conclusion, to maximize the expected revenue $E(I_1)$, equation (11) has to be satisfied and $q^* \geq a_2 Q^*$.

3.2.2. Company's Order Decision under Decentralized Decision-Making. The revenue function under decentralized decision-making for the company is

$$\begin{aligned} I_2 &= p \min\{q, Qx, Dy\} - w \min\{q, Qx\} \\ &\quad - s(Dy - \min\{q, Qx\})^+ + h_1 (\min\{q, Qx\} - Dy)^+. \end{aligned} \quad (12)$$

To maximize $E(I_1)$, there must be $q^* \geq a_2 Q^*$ and $x \in [a_1, a_2]$, so we can make $q^* \geq Q^* x$ constant.

In this case, the revenue function can be transformed into

$$I_2 = p \min\{Qx, Dy\} - wQx - s(Dy - Qx)^+ + h_1 (Qx - Dy)^+. \quad (13)$$

The expected revenue function is

$$\begin{aligned}
 E(I_2) &= p \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} Qx f(x)g(y) dx dy \\
 &+ p \int_{b_1}^{b_2} \int_{(Dy/Q)}^{a_2} Dy f(x)g(y) dx dy - w\mu Q \\
 &- s \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} (Dy - Qx) f(x)g(y) dx dy \\
 &+ h_1 \int_{b_1}^{b_2} \int_{(Dy/Q)}^{a_2} (Qx - Dy) f(x)g(y) dx dy.
 \end{aligned} \tag{14}$$

According to equation (14), the first and second partial derivatives of $E(I_2)$ with respect to Q are obtained:

$$\begin{aligned}
 \frac{\partial E(I_2)}{\partial Q} &= (p + s) \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} x f(x)g(y) dx dy - w\mu \\
 &+ h_1 \int_{b_1}^{b_2} \int_{(Dy/Q)}^{a_2} x f(x)g(y) dx dy,
 \end{aligned} \tag{15}$$

$$\begin{aligned}
 \frac{\partial^2 E(I_2)}{\partial Q^2} &= (p + s) \int_{b_1}^{b_2} \frac{Dy}{Q} \times \left(-\frac{Dy}{Q^2} \right) \times f\left(\frac{Dy}{Q}\right) \times g(y) dy \\
 &- h_1 \int_{b_1}^{b_2} \frac{Dy}{Q} \times \left(-\frac{Dy}{Q^2} \right) \times f\left(\frac{Dy}{Q}\right) \times g(y) dy \\
 &= -\frac{p + s - h_1}{Q^3} \int_{b_1}^{b_2} (Dy)^2 \times f\left(\frac{Dy}{Q}\right) \times g(y) dy.
 \end{aligned} \tag{16}$$

Since $(p + s - h_1) > 0$ is constant, $(\partial^2 E(I_2)/\partial Q^2) < 0$; the expected revenue $E(I_2)$ is a concave function of Q , so there is Q^0 that makes $\max E(I_2)$.

Let $(\partial E(I_2)/\partial Q) = 0$; we got

$$\begin{aligned}
 (p + s) \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q^0)} x f(x)g(y) dx dy \\
 + h_1 \int_{b_1}^{b_2} \int_{(Dy/Q^0)}^{a_2} x f(x)g(y) dx dy = w\mu.
 \end{aligned} \tag{17}$$

From equation (17), we got $Q^0 = Q^*$, and from equation (11), we got q^* . Obviously, under decentralized decision-making, there is a unique Nash equilibrium (Q^*, q^*) .

Corollary 1. *In a supply chain with uncertain yield and demand, the overall expected return under centralized decision-making is no less than that of decentralized decision-making.*

Proof. When $q^* \geq a_2 Q^*$, that is, the maximized revenue for the farmer under decentralized decision-making, since $x \in [a_1, a_2]$, $q^* \geq Qx$ and the total expected revenue of the company and farmers is

$$\begin{aligned}
 E[I_1(Q)] + E[I_2(Q)] &= p \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} Qx f(x)g(y) dx dy \\
 &+ p \int_{b_1}^{b_2} \int_{(Dy/Q)}^{a_2} Dy f(x)g(y) dx dy \\
 &- s \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} (Dy - Qx) f(x)g(y) dx dy \\
 &+ h_1 \int_{b_1}^{b_2} \int_{(Dy/Q)}^{a_2} (Qx - Dy) f(x)g(y) dx dy - cQ.
 \end{aligned} \tag{18}$$

The expected total revenue of the supply chain under the centralized decision is

$$\begin{aligned}
 E[I_0(Q)] &= p \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} Qx f(x)g(y) dx dy \\
 &+ p \int_{b_1}^{b_2} \int_{(Dy/Q)}^{a_2} Dy f(x)g(y) dx dy \\
 &+ h_1 \int_{b_1}^{b_2} \int_{(Dy/Q)}^{a_2} (Qx - Dy) f(x)g(y) dx dy \\
 &- s \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} (Dy - Qx) f(x)g(y) dx dy - cQ.
 \end{aligned} \tag{19}$$

Now $\max E[I_1(Q)] + E[I_2(Q)] = E[I_0(Q)]$, so $E(I_1) + E(I_2) \leq E(I_0)$ is constant. \square

3.3. Supply Chain Coordination Production Decision Based on Revenue Sharing plus Margin. In the case of decentralized decision, there must be $q^* \geq Q^*x$, so there can be $E(I_1)$; namely, the expected revenue of the farmer can be maximized. The revenue function of farmers does not consider the stock-out risk, which means that the actual output of farm products is less than the agreed quantity. As a result, the company will be unable to meet the market demand because of insufficient supply and confronted with high out-of-stock costs. Besides, under decentralized decision-making, both farmers and companies will aim to maximize their expected income, which will generate a double marginal effect and reduce the overall benefit of the supply chain.

Therefore, this paper proposes a supply chain coordination strategy based on revenue sharing plus margin, so as to restrain both parties of the contract, reduce the default risk of both parties in the company-plus-farmer mode of order agriculture, and improve the stability of the supply chain. The coordination mechanism of revenue-sharing-plus-margin supply chain is as follows:

- (1) The company purchases the whole output of agricultural products at the purchase price w in the order contract.

- (2) Farmers enjoy a part of the company's sales revenue, and the income sharing coefficient is r , thereby reducing the loss of farmers considering the possibility that the purchase price of agricultural products in the order contract is lower than the sales price.
- (3) The technical guidance or financial support from the company is needed in the process of production. Therefore, farmers are supposed to pay a certain amount of margin M to the company as compensation for farmers' wrongdoing.

Under the coordination mechanism of revenue-sharing-plus-margin supply chain, the revenue function of the farmer is

$$I_1 = wQx + rp \min\{Dy, Qx\} - cQ - M. \quad (20)$$

The revenue function of the company is

$$I_2 = (1-r)p \min\{Dy, Qx\} + h_1(Qx - Dy)^+ + M - s(Dy - Qx)^+ - wQx. \quad (21)$$

And the total revenue function of the supply chain is

$$I_0 = p \min\{Dy, Qx\} + h_1(Qx - Dy)^+ - s(Dy - Qx)^+ - cQ. \quad (22)$$

At this point, the total revenue function of the supply chain is consistent with that of the centralized decision-making supply chain; that is, the total revenue of the risk-sharing supply chain achieves the same effect as that of centralized decision-making supply chain. One has

$$E(I_1) = wQ\mu + rp \int_{b_1}^{b_2} \int_{(Dy/Q)}^{a_2} Dy f(x)g(y) dx dy + rp \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} Qx f(x)g(y) dx dy - cQ - M, \quad (23)$$

$$E(I_2) = (1-r)p \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} Qx f(x)g(y) dx dy + (1-r)p \int_{b_1}^{b_2} \int_{(Dy/Q)}^{a_2} Dy f(x)g(y) dx dy + h_1 \int_{b_1}^{b_2} \int_{(Dy/Q)}^{a_2} (Qx - Dy) f(x)g(y) dx dy + M - s \int_{b_1}^{b_2} \int_{a_1}^{(Dy/Q)} (Dy - Qx) f(x)g(y) dx dy - wQ\mu. \quad (24)$$

According to equation (5), Q' is obtained through equation $E(I_0)$, and with equation (23) as well as equation (24), $E[I_1(Q')]$ and $E[I_2(Q')]$ can be obtained.

Since $(\partial E(I_1)/\partial r) > 0$, $(\partial E(I_1)/\partial M) < 0$, $(\partial E(I_2)/\partial r) < 0$, and $(\partial E(I_2)/\partial M) > 0$, in other words, the expected revenue of the farmer is the increasing function of r (the revenue-sharing coefficient) and the decreasing function of M (the margin). The expected revenue of the company is a decreasing function of r and an increasing function of M .

The expected total return of the revenue-sharing-plus-margin supply chain levels up to that of the centralized decision-making supply chain, and the overall return of the supply chain fulfills Pareto improvement. In addition, under revenue-sharing-plus-margin coordination, the profit coordination between the company and the farmer is mainly achieved by adjusting r (the income sharing coefficient) and M (the margin), and during the course of coordination, the total return under decentralized decision-making mechanism also requires Pareto improvement in order to maintain the stability of company-plus-farmer supply chain.

4. Example Analysis

Based on the above deduction and analysis, the example design and result analysis are carried out on the assumption that the random output factor and the market demand fluctuation factor are uniformly distributed. Take the contract farming supply chain consisting of one single agricultural product, one company, and one farmer as an example, and the parameters of the calculation example are shown in Table 1.

MATLAB is used for programming calculation, and the result is shown in Table 2.

By comparing the numerical results of the example, it can be concluded that the expected revenue of the supply chain under the centralized decision is higher than that of the one under the decentralized decision, realizing the Pareto improvement and also proving the feasibility of revenue-sharing-plus-margin supply chain coordination.

4.1. Influence of Changes of r (Revenue-Sharing Coefficient) on the Coordination of the Supply Chain. Make an analysis of the influence of the changes of r (revenue-sharing coefficient) on the revenue coordination and adjust the value of r . The results are shown in Table 3 and Figure 1.

It can also be seen from the results of revenue coordination that the expected income of farmers increases with the increase of r (revenue-sharing coefficient), while the expected income of the company decreases on the other hand. Moreover, the expected income of farmers achieves Pareto improvement under the coordinated production decision of the supply chain. When $r \in [0, 0.020743]$, the expected income of the company is expected to realize Pareto optimization.

4.2. Influence of the Change of M (Margin) on the Coordination of Supply Chain. Make an analysis of the influence of the changes of M (the margin) on the revenue coordination and adjust the value of M . The results are shown in Table 4 and Figure 2.

It can also be seen from the results of revenue coordination that the expected income of farmers decreases with the increase of M (the margin), while the expected income of the company increases on the other hand. Moreover, the expected income of farmers achieves Pareto improvement under the coordinated production decision of the supply

TABLE 1: The basic data of the example.

Parameter	p	w	c	h	h_1	s	D	r	M	x	y
Value	12	6	4	2	7	10	100	0.0005	50	$x \sim U(0.5, 1.1)$	$y \sim U(0.8, 1.3)$

TABLE 2: The basic result of the example.

	Farmer's decision		Company's decision		The expected total revenue $E(I_0)$
	Q	$E(I_1)$	q	$E(I_2)$	
Decentralized decision-making	245.752	185.747	242.038	694.751	880.498
Centralized decision-making	303.441	198.659	—	713.346	912.005

TABLE 3: The influence of the changes of r on the coordination of the supply chain.

r	Expected revenue of the farmer $E(I_1)$	Expected revenue of the company $E(I_2)$
0*	192.753	719.252
0.005	198.659	713.346
0.01	204.565	707.44
0.015	210.471	701.534
0.02	216.377	695.628
0.020743*	217.254	694.751
0.025	222.283	689.722

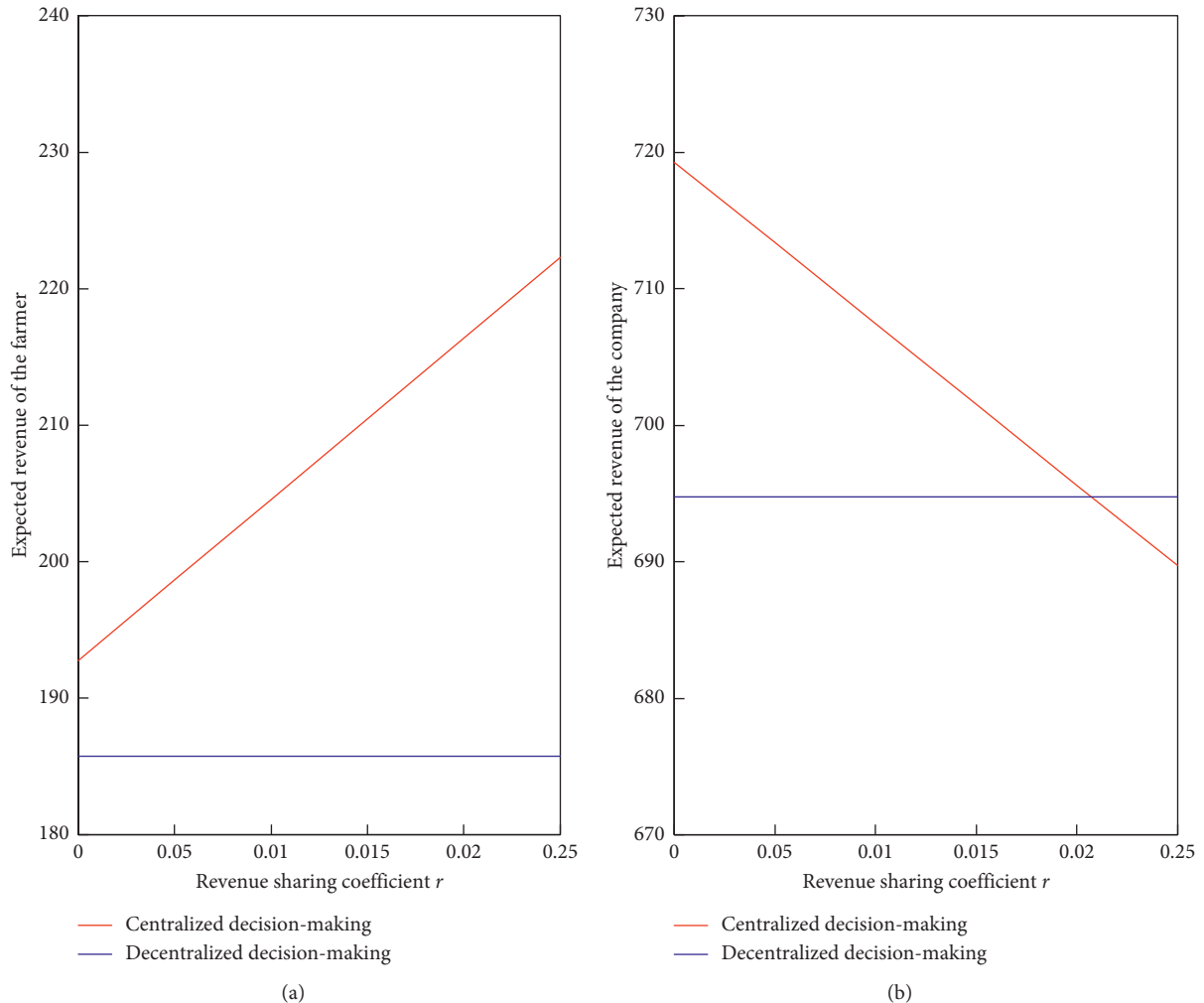


FIGURE 1: The influence of the changes of r on the coordination of the supply chain.

TABLE 4: The influence of the changes of M (the margin) on the coordination of supply chain.

M	Expected revenue of the farmer, $E(I_1)$	Expected revenue of the company, $E(I_2)$
20	228.659	683.346
30	218.659	693.346
31.405*	217.254	694.751
40	208.659	703.346
50	198.659	713.346
60	188.659	723.346
62.912*	185.747	726.258
70	178.659	733.346

*The critical value of the parameter M when the expected revenue of the farmer and the company reaches the basic value in Table 2.

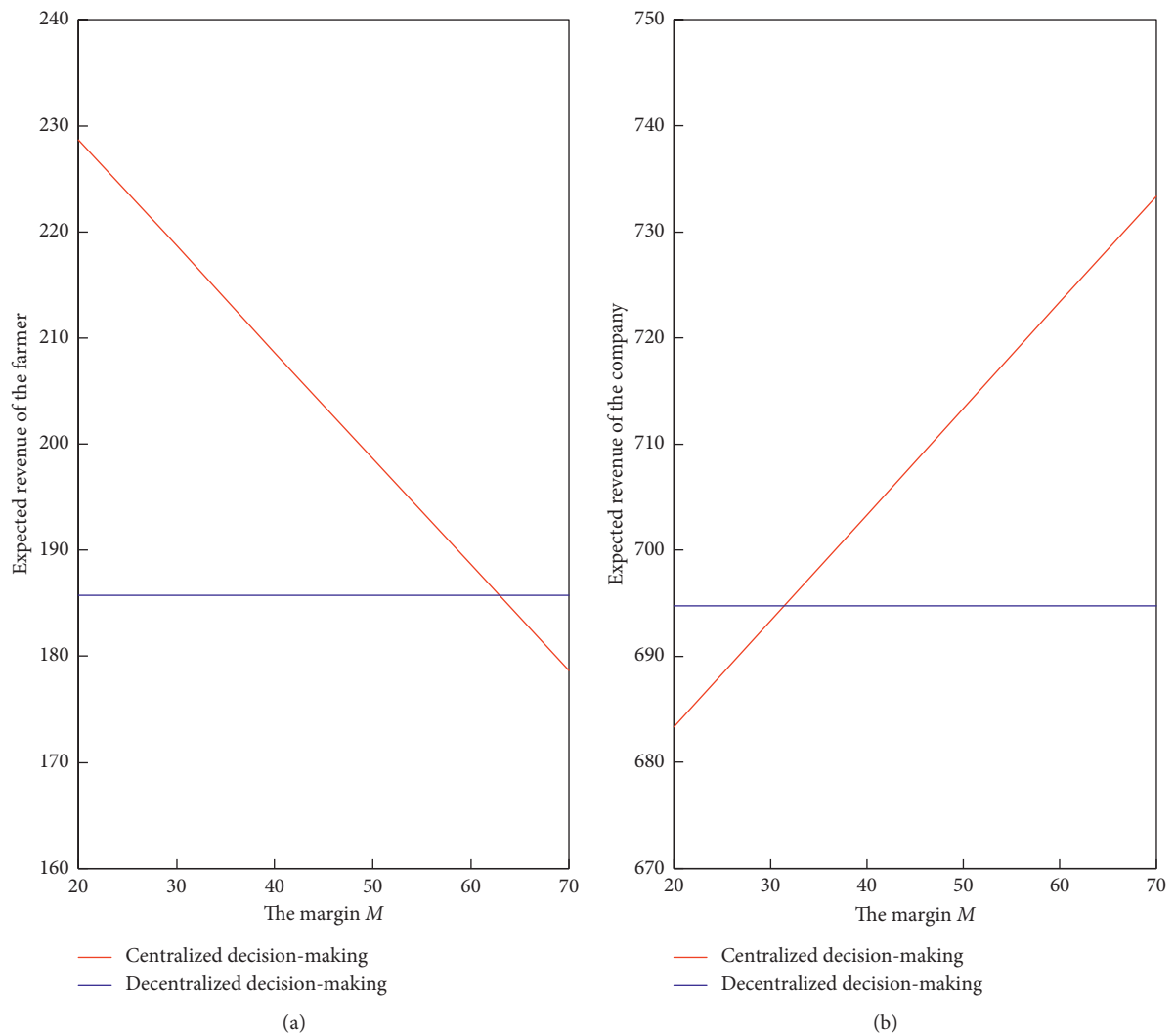


FIGURE 2: The influence of the changes of M (the margin) on the coordination of supply chain.

chain. When $M \in [31.405, 62.912]$, the expected income of both parties is expected to realize Pareto optimization.

Combined with the above analysis, the coordination of the revenue distribution between the company and the farmer can be optimized by adjusting the values of r (income sharing coefficient) and M (the margin) under the coordination mechanism of revenue sharing plus margin in the supply

chain. In order to maintain the stability of the supply chain, make sure $r \in [0, 0.020743]$ as well as $M \in [31.405, 62.912]$, so as to achieve Pareto optimization of both expected revenues. Therefore, it is advisable for companies and farmers to adjust r and M within the range to achieve centralized supply chain overall revenue and Pareto improvement of supply chain coordination under centralized decision.

5. Conclusions

Based upon the Newsvendor model, a company-plus-farmer contract farming supply chain was established in this paper characterized by one company, one smallholder farmer, and one single production cycle. In the case of decentralized decision-making coordination, the expected revenue functions of farmers and of companies were constructed, respectively, and the theoretical analysis showed that there was an optimal range of order quantity and targeted yield to maximize the expected income of farmers. Taking into account both decentralized coordination and centralized production decision-making, the revenue-sharing-plus-margin coordination mechanism of both parties under the centralized decision-making was put forward and proved to be conducive to Pareto improvement in overall revenue of the supply chain, based on theories and calculation examples. And the Pareto optimization of the expected returns of the company and the smallholder farmer in the supply chain coordination of revenue sharing plus margin was discussed by means of examples.

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 article.

Acknowledgments


The research was funded by the Basic Competence Improvement Project for Young and Middle-Aged Teachers in Higher Education of Guangxi (no. 2018KY0341) and the National Natural Science Foundation of China (no. 71462005).

References

- [1] J. E. Hobbs and L. M. Young, "Increasing vertical linkages in agrifood supply chains: a conceptual model and some preliminary evidence," *Canadian Journal of Agricultural Economics/Revue Canadienne D'Agroeconomie*, vol. 47, no. 4, p. 478, 1999.
- [2] S. W. Martinez, "Vertical coordination of marketing systems: lessons from the poultry, egg, and pork industries," Agricultural Economics Report 34051, United States Department of Agriculture, Economic Research Service, Washington, DC, USA., 2002.
- [3] H. Guo and R. W. Jolly, "Contractual arrangements and enforcement in transition agriculture: theory and evidence from China," *Food Policy*, vol. 33, no. 6, pp. 570–575, 2008.
- [4] O. Boyabtli, P. R. Kleindorfer, and S. R. Koontz, "Integrating long-term and short-term contracting in beef supply chains," *Management Science*, vol. 57, no. 10, pp. 1771–1787, 2011.
- [5] G. K. Abebe, J. Bijman, R. Kemp, O. Omta, and A. Tsegaye, "Contract farming configuration: smallholders' preferences for contract design attributes," *Food Policy*, vol. 40, pp. 14–24, 2013.
- [6] F. Liu, "Incomplete contract and performance obstacle—a case study of order agriculture," *Economic Research*, vol. 49, pp. 22–30, 2003.
- [7] J. Liu and C. Qi, "Research on the influencing factors of company-plus-farmer contract: a framework of transaction cost analysis," *Economic Latitude*, vol. 25, pp. 106–109, 2009.
- [8] Z. Wu, X. Huo, and J. Liu, "The stability of farmer-plus-leading enterprise contract," *Journal of Hebei University of Technology (Social Sciences)*, vol. 10, pp. 55–57, 2010.
- [9] Y. Fang, "Empirical analysis of the management of contract farming default risk by futures market," *Agricultural Technology Economy*, vol. 30, pp. 97–104, 2011.
- [10] Research Group, Xi'an Branch, People's Bank of China, "Logistic empirical study on influencing factors of order agriculture contract performance-based on 381 copies of order contracts from Shaanxi Province," *Contemporary Economic Science*, vol. 36, pp. 118–128, 2014.
- [11] H. Yang, D. Zhang, and H. Wang, "Research on profit sharing in contract farming-based on the analysis of produce price fluctuation and default risk," *Theory and Practice of Price*, vol. 10, pp. 38–42, 2019.
- [12] F. Y. Chen and C. A. Yano, "Improving supply chain performance and managing risk under weather-related demand uncertainty," *Management Science*, vol. 56, no. 8, pp. 1380–1397, 2010.
- [13] L. Zhen, "Modeling of yard congestion and optimization of yard template in container ports," *Transportation Research Part B: Methodological*, vol. 90, pp. 83–104, 2016.
- [14] L. Zhen, "Tactical berth allocation under uncertainty," *European Journal of Operational Research*, vol. 247, pp. 928–944, 2015.
- [15] Q. Wang, J. J. Li, W. T. Ross Jr., and C. W. Craighead, "The interplay of drivers and deterrents of opportunism in buyer-supplier relationships," *Journal of the Academy of Marketing Science*, vol. 41, no. 1, pp. 111–131, 2013.
- [16] X. Zhao and F. Wu, "Research on the revenue sharing contract of agricultural products supply chain coordination under random output and demand," *Management Science of China*, vol. 17, pp. 88–95, 2009.
- [17] L. Ling, Z. Hu, and X. Guo, "Analysis and coordination of farmer-to-supermarket mode under random output and demand," *System Engineering*, vol. 29, pp. 36–40, 2011.
- [18] P. Liu and T. Xie, "Coordination of agricultural products supply chain under random output and demand," *Logistics Technology*, vol. 32, pp. 381–384, 2013.
- [19] G. Sun and L. Xu, "Research on option contracts of produce supply chain under random supply and demand," *Journal of Management Engineering*, vol. 28, pp. 201–210, 2014.
- [20] K. Qin and L. Teng, "Research on order agriculture supply chain under uncertain conditions," *Economic Problems*, vol. 2, pp. 111–116, 2016.
- [21] X. Tong, "Research on order agriculture supply chain coordination under uncertain market demand," *Technology Economics and Management Research*, vol. 4, pp. 67–71, 2018.

Research Article

Heuristic Algorithms for MapReduce Scheduling Problem with Open-Map Task and Series-Reduce Tasks

Feifeng Zheng,¹ Zhaojie Wang,¹ Yinfeng Xu,¹ and Ming Liu^{1,2} 

¹Glorious Sun School of Business & Management, Donghua University, Shanghai, China

²School of Economics & Management, Tongji University, Shanghai, China

Correspondence should be addressed to Ming Liu; mingliu@tongji.edu.cn

Received 20 May 2020; Revised 14 June 2020; Accepted 23 June 2020; Published 15 July 2020

Academic Editor: Lu Zhen

Copyright © 2020 Feifeng Zheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Based on the classical MapReduce concept, we propose an extended MapReduce scheduling model. In the extended MapReduce scheduling problem, we assumed that each job contains an *open-map task* (the *map task* can be divided into multiple unparallel operations) and *series-reduce tasks* (each *reduce task* consists of only one operation). Different from the classical MapReduce scheduling problem, we also assume that all the operations cannot be processed in parallel, and the machine settings are unrelated machines. For solving the extended MapReduce scheduling problem, we establish a mixed-integer programming model with the minimum makespan as the objective function. We then propose a genetic algorithm, a simulated annealing algorithm, and an *L-F* algorithm to solve this problem. Numerical experiments show that *L-F* algorithm has better performance in solving this problem.

1. Introduction

For meeting the unpredictable demand, manufacturing companies often acquire more manufacturing equipment. Scheduling jobs to be processed on these manufacturing equipment increase the complexity of business operations [1]. Job scheduling is mainly used to assist decision makers in deciding the processing sequence of jobs and the allocation of machines. Many scholars have studied the classical scheduling problems, such as single-machine scheduling problem and parallel-machine scheduling problem. It is worth noting that the combination of MapReduce job and machine scheduling has attracted more and more scholars' attention in recent years. In fact, MapReduce system is a popular method of big data processing [2]. MapReduce generally consists of one map task and one reduce task, and the reduce task is strictly behind the map task. Normally, the map task can be arbitrarily split into multiple map operations, and all the map operations can be processed in parallel on multiple parallel machines, while the reduce tasks are unsplitable; thus, the reduce tasks can only be processed by one machine. In the classical MapReduce scheduling problem, each job consists of the map tasks and the reduce

tasks, and the reduce task is strictly behind the map task [3]. Normally, the reduce task contains only one operation.

Many current studies on MapReduce scheduling problem assume that the available machines are identical parallel machines, i.e., for the same operation, the processing speed of machines is the same. In actual production and scheduling environment, production models of MapReduce-like systems are widely used. For example, the manufacturing and manual assembly of parts and components of some valuable watches can be regarded as the production mode of MapReduce. Another example, after finishing the basic production of a high-grade cheongsam, it often requires manual embroidery by multiple processors. In the process of embroidery, many embroidered patterns are often not processed in sequence. After the embroidery is finished, it needs to be ironed by one of the processors and then be packed. From embroidery to packaging, this cheongsam can be regarded as the extended MapReduce-like job. Undoubtedly, MapReduce scheduling has a strong practical significance.

However, in the actual situation, there is a kind of jobs similar to MapReduce job. It may be easier for us to describe their features in terms of MapReduce terminology. Thus, we

will mention such jobs as *extended MapReduce jobs*. Specifically, each extended MapReduce job consists of one *open-map task* and several *series-reduce tasks*. Different from the classical MapReduce job, the extended MapReduce job cannot be processed in parallel. The operations of open-map task of the extended MapReduce job can be processed by any machine without strict sequential processing sequence (like operations of an open shop job), but cannot be processed in parallel. This is similar to open shop, so we call it *open-map task*. We say each open-map task of a job can be divided into multiple map operations. The expanded MapReduce job also consists of several *series-reduce tasks*. In this paper, we mainly study the expanded MapReduce scheduling problem.

We consider an extended MapReduce scheduling problem with the open-map task and series-reduce tasks, motivated by their practical fixed partition applications. There are three main differences between our research and the classical MapReduce scheduling problem:

- (1) We assume that the parallel machines are unrelated, while the machines are identical in the classical MapReduce model
- (2) In the extended MapReduce scheduling problem, map operations (i.e., a partition of the open-map task) cannot be processed in parallel. In the extended MapReduce scheduling problem, the division of an open-map task, i.e., map operations, is known a priori
- (3) We extend the classical MapReduce scheduling problem by assuming that there are several series-reduce tasks during processing.

In the extended MapReduce scheduling problem, each job consists of two kinds of tasks: one open-map task and several series-reduce tasks. There is no strict requirement of processing sequence for the map operations (like operations of one job in open shop), but there is strict requirement of processing sequence for the reduce operations (each operation corresponds to one series-reduce task). Moreover, it is obvious that the extended MapReduce scheduling problem is an NP-hard problem, because it is an extension of the parallel machine makespan minimization problem.

To address the extended MapReduce scheduling problem, we establish a mixed-integer programming model with the minimum makespan as the objective function. We then solve it by CPLEX, a genetic algorithm, and a simulated annealing algorithm, respectively. Besides, we design a heuristic algorithm which we called *L-F* algorithm to improve the solution efficiency. We compare *L-F* algorithm with the genetic algorithm and the simulated annealing algorithm to verify its performance.

The studies of MapReduce scheduling problem often assume that all the jobs are processed on identical parallel machines [4, 5]. First, in the case of online study based on this assumption, [6] considers a two-stage classical flexible flow shop problem based on MapReduce system, and give an online 6-approximation and an online $(1 + \epsilon)$ -speed $O(1/\epsilon^4)$ -competitive algorithm. Assuming that the reduce task is

unknown before the map task is completed, [7] considers the preemptive and nonpreemptive processing scenarios and prove that no matter what the processing scenario is the competition ratio of any algorithm is at least $2 - 1/n$. For improving system utilization for batch jobs, [8] proposes a modified MapReduce architecture that allows data to be pipelined between operators. Jiang et al. [9] propose an online algorithm which can optimize their MapReduce problem, where the map operations are fractional and the reduce operations are preemptive. The paper [10] studies an online MapReduce scheduling problem on two uniform machines and discusses the competitive ratio with preemption or nonpreemption, respectively.

Second, in the online scenario, [11] solves an online job scheduling problem with flexible deadline bounds via a receding horizon control algorithm. The study in [12] expands the MapReduce tasks by allowing user specify a job's deadline and try to make the job finished before the deadline. Li et al. [13] study a MapReduce periodical batch scheduling problem with makespan minimization and propose three heuristic algorithms to solve the problem. Fotakis et al. [14] propose an approximation algorithm for the nonpreemptive MapReduce scheduling. They also verify the performance of the algorithm through numerical experiments. Most studies in the field of MapReduce scheduling have only focused on the improvement of algorithms for the classical MapReduce scheduling problem. However, few researches expand the classical MapReduce scheduling problem according to its practical applications.

In this work, we consider an extended MapReduce scheduling problem with the open-map task and several series-reduce tasks. In this problem, we assume that the open-map task of each job is splittable (with fixed partition pattern), and all the series-reduce tasks are unsplittable in processing. Specifically, the operations of the open-map task of one job are given beforehand. These operations must be performed in sequence, but without precedence relation. We establish a mixed-integer programming model with the minimum makespan as the objective function then adopt CPLEX to solve it. We design a heuristic algorithm called *L-F* algorithm for this problem and compare it with the baseline heuristic algorithms. Finally, we found that *L-F* algorithm has advantages through numerical experiments.

The main contributions of this work are threefold:

- (1) Different from the classical MapReduce scheduling problem, we consider series-reduce tasks (in this work, we are mainly considering two series-reduce tasks). The map operations cannot be processed in parallel.
- (2) We consider the processing scenario in which each job is processed on a set of unrelated parallel machines with different processing speeds.
- (3) We establish a mixed-integer programming model; then, we adopt CPLEX and two classical heuristic algorithms to solve this problem. Besides, we design an *L-F* algorithm to solve this problem more efficiently.

The remainder of this paper is organized as follows. Section 2 mathematically models the problem and solves the problem via CPLEX. In order to solve this problem in large scale, Section 3 designs heuristic algorithms (a genetic algorithm, a simulated annealing algorithm, and an L - F algorithm). Section 4 compares the designed L - F algorithm with genetic algorithm and simulated annealing algorithm. Section 5 concludes the paper and gives future research directions.

2. Problem Statement and Mathematical Model

There are n jobs to be processed on the m unrelated machines. Each job contains k operations. There are one open-map task and several series-reduce tasks. The operations 1 to $(k-r)$ represent the map operations of the open-map task of each job. The operations $(k-r+1)$ to k represent reduce operations of the series-reduce tasks of the job. The specific characteristics of a job with two series-reduce tasks are shown in Figure 1. The processing time of operation u ($1 \leq u \leq k$) for job j by machine i is denoted by P_i^{ju} . In addition, the processing of any two (no matter map or reduce) operations of a job cannot be overlapped. That is to say, any job cannot be processed on more than one machine at the same time. The other processing rules are as follows. We aim to produce a processing schedule so as to minimize the makespan.

- (1) In the open-map task of any job, there is no strict requirement for the processing sequence of the given map operations.
- (2) For any job, only after completing map operations, the reduce operations can start to be processed.
- (3) The reduce operations can be processed by any available machine, and the reduce operations of a job are processed in a strict sequence. After the previous reduce operation is finished, the next one can be processed.

The problem under consideration is based on the assumptions that setup time of machine and transportation time of job between operations are not considered. In this work, we propose scheduling methods to minimize the makespan for this problem. Without loss of generality, we assume that in a reasonable schedule, there may be idle time between the processing times of any two consecutive operations on a machine. Below, we propose a mixed-integer linear programming model for this problem. For convenience, we first explain the basic symbols and decision variables and then build the model.

2.1. Basic Notations

2.1.1. Indices

- i : index of machines
- j : index of jobs
- u : index of all operations of a job
- h : index of positions on a machine
- v : index of reduce operations of a job
- s : index of processing orders of operations of a job

2.1.2. Input Parameters

- \mathcal{M} : set of machines, and $\mathcal{M} = \{1, 2, \dots, m\}$
- J : set of jobs, and $J = \{1, 2, \dots, n\}$
- r : the number of reduce operations
- U : set of operations, and $U = \{1, 2, \dots, k-r, k-r+1, \dots, k\}$
- H : set of positions on a machine, and $H = \{1, 2, \dots, p\}$
- M : a sufficiently large positive integer
- P_i^{ju} : the processing time of operation u of job j when processed by machine i

2.1.3. Decision Variables

- x_{ih}^{ju} : a binary variable equal to 1 if operation u of job j begins to be processed by machine i at position h , 0 otherwise.
- z_i^{ju} : a binary variable equal to 1 if operation u of job j is processed by machine i , 0 otherwise.
- y_s^{ju} : a binary variable equal to 1 if operation u in job j is processed in the order of s , 0 otherwise. Note that if $u \in [r+1, k]$, s must be equal to u , because they correspond to series-reduce tasks.
- C^{ju} : the completion time of the operation u of job j .
- C_{\max} : the maximum completion time of operation u of job j .
- S^{ju} : the time when operation u of job j began to be processed.
- P^{ju} : the processing time of operation u of job j .

2.2. Mathematical Model. The objective is to minimize the makespan, as shown in the following formula:

$$\min f = C_{\max}, \quad (1)$$

subject to

$$\sum_{j \in J} \sum_{u \in U} x_{ih}^{ju} = 1, \quad i \in \mathcal{M}, h \in H, \quad (2)$$

$$\sum_{h \in H} x_{ih}^{ju} = z_i^{ju} 1, \quad i \in \mathcal{M}, j \in J, u \in U, \quad (3)$$

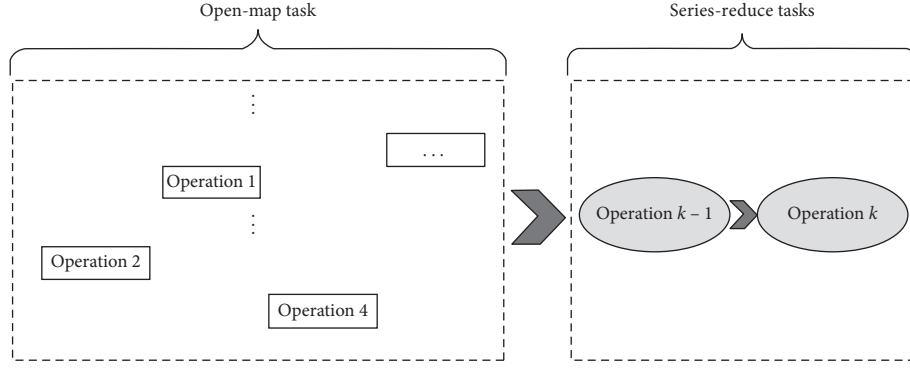


FIGURE 1: The specific process characteristics of the extended MapReduce job (the case that $r=2$).

$$\sum_{i \in \mathcal{M}} z_i^{ju} = 1, \quad j \in J, u \in U, \quad (4)$$

$$\sum_{s \in U} y_s^{ju} = 1, \quad j \in J, u \in U, \quad (5)$$

$$\sum_{u \in U} y_s^{ju} = 1, \quad j \in J, u \in U, \quad (6)$$

$$S^{ju} - C^{ju'} \geq 0 - M * (2 - y_s^{ju} - y_s^{ju'}), \quad j \in J, u, u', s, s' \in U, \text{ and } u' \neq u, s \geq s', \quad (7)$$

$$P^{ju} = \sum_{i \in \mathcal{M}} P_i^{ju} * z_i^{ju}, \quad j \in J, u \in U, \quad (8)$$

$$S^{ju} \geq 0 - M * (1 - x_1^{ju}), \quad j \in J, u \in U, \quad (9)$$

$$S^{ju} - C^{ju'} \geq 0 - M * \left(2 - \sum_{i \in \mathcal{M}} x_{ih}^{ju} - \sum_{i \in \mathcal{M}} x_{ih}^{ju'} \right), \quad j, j' \in J, u, u' \in U, \text{ and } j \neq j' \text{ or } u \neq u', h, h' \in H \text{ and } h > h', \quad (10)$$

$$S^{ju} \geq \sum_{j' \in J} \sum_{u' \in U} \sum_{h' \in [1, h-1]} P_i^{j'u'} * x_{ih}^{j'u} - M * (1 - x_{ih}^{ju}), \quad (11)$$

$$j \in J, u \in U, h \in H \text{ and } j \neq j' \text{ or } u \neq u',$$

$$C^{ju} = S^{ju} + P^{ju}, \quad j \in J, u \in U, \quad (12)$$

$$C_{\max} \geq C^{jk}, \quad j \in J, \quad (13)$$

$$S^{jv} \geq C^{jv'}, \quad j \in J, v \in \{k-r+1, k-r+2, \dots, k-1\}, v' \in \{v+1, v+2, \dots, k\}, \quad (14)$$

$$S^{jv} \geq C^{ju}, \quad j \in J, u \in \{1, 2, \dots, k-r\}, v \in \{k-r+1, k-r+2, \dots, k\}, \quad (15)$$

$$x_{ih}^{ju}, z_i^{ju}, y_s^{ju} \in \{0, 1\}, \quad i \in \mathcal{M}, j \in J, u, s \in U, h \in H, \quad (16)$$

$$C^{ju}, C_{\max}, P^{ju}, S^{ju}, \quad j \in J, u \in U. \quad (17)$$

Constraints (2)–(4) guarantee that each operation can only be processed by at most one machine at a time, and any operation must be completed. Constraints (5)–(7) guarantee that any operation cannot be processed in parallel. Constraint (8) calculates the actual processing time of operation u of job j . Constraints (9)–(11) calculate the start processing time of any operation and describe the relationship between the start processing time of two operations on the same machine. Constraints (12)–(13) calculate the completion time of any operation and the makespan. Constraints (14)–(15) calculate the start processing time of the reduce operation. Constraints (16)–(17) limit the ranges of the variables.

3. Heuristic Algorithms

As the problem is NP-hard, we propose a genetic algorithm and a simulated annealing algorithm to solve it. Since the job with two reduce tasks are common in practice, we solve the extended MapReduce scheduling problem with two reduce tasks via CPLEX and heuristic algorithms. Otherwise, a new heuristic algorithm, i.e., L - F algorithm, is proposed to solve this problem more efficiently.

3.1. Genetic Algorithm. Genetic algorithms are computational models to simulate the natural selection and genetic mechanism of Darwin's biological evolution theory. It is a method to search the optimal solution by simulating the natural evolution process. Genetic algorithms have been widely used in scheduling problems [15, 16] and other NP-complete problems [17]. In genetic algorithms, each chromosome represents a feasible solution to the problem studied. In this section, for solving this problem, we generate a certain number of initial feasible solutions, i.e., initial population, and then through cyclic selection of the parents, we get offspring through the process of crossover and mutation of the parents' chromosomes, constantly optimizing to find the optimal solution.

3.1.1. Chromosome Representation. In this study, each chromosome can be divided into two parts on average. All the elements of the first part represent all the operations of all the jobs, which are called operation genes. All the elements of the second part represent the number of machines, which are called machine genes. For example, in Figure 2, the first element of the chromosome is 21, representing operation 1 of job 2, and the tenth element (the first machine gene) is 1, representing machine 1. The operation genes correspond to the machine genes one by one. For example, in Figure 2, the first operation gene is 21, and the first machine gene is 1, which means that operation 1 of job 2 is processed by machine 1.

In addition, on each chromosome, if both operations belong to a job, the start processing time in the subsequent operation is later than the completion time of the preceding operation. Similarly, on each chromosome, if two operations are assigned to the same machine, the start time of the

subsequent operation is later than the completion time of the previous operation.

3.1.2. Initialization of Chromosomes. We randomly generate chromosomes according to the number of jobs, the number of machines, and the processing rules. For example, in Figure 2, in the first part of chromosomes, operation 2 of job 2 (22) and operation 1 of job 1 (11) correspond to machine 2. However, operation 2 of job 2 is ahead of operation 1 of job 1. That is to say, the starting time of operation 1 of job 1 must be after the completion time of operation 2 of job 2.

3.1.3. Fitness Function. In the genetic algorithm, fitness determines the probability that a chromosome is selected as a parent chromosome. We assume that NIND is the number of individuals in the population, j is the j th chromosome in the population, and $obj(j)$ is the corresponding objective value of chromosome j in the population. Since this study aims to minimize the objective function, we define the fitness value of chromosome j as $1/obj(j)$.

3.1.4. Crossover and Mutation. In genetic algorithm, after mutation and crossover, the offspring cannot only retain the high-quality genes stained by the parent chromosomes, but also make the new chromosome population more diverse. However, the processes of crossover and mutation are easy to generate infeasible solution. For example, in this study, if we cross or mutate the operation genes of the chromosomes, this will easily lead to some operations being reprocessed and some operations not being processed, which is obviously inconsistent with our intention. Adopting a correcting algorithm to modify unfeasible offspring is very time-consuming; hence, it is preferable to design operators such that precedence constraints are not violated [18].

In this problem, any operation can be processed by any machine. Therefore, in order to avoid infeasible solutions, we only cross the machine genes. Since there is no crossover of the operations genes in the chromosomes, the order of operation genes in the chromosome is still feasible. It is obvious that each of the obtained offspring still represents a feasible solution. From the aspect of chromosome mutation, we randomly select a chromosome in a population. When the generated random number is less than the mutation probability, based on empirical data (the following six steps are executed with a probability of 0.1 to 1, respectively; we find that the effect is better when the probability is 0.3), we divide the genetic algorithm into the following six steps to carry out chromosome mutation based on empirical data:

Step 1. Randomly select two machine genes on the chromosome with the probability of 0.3, and then generate them again.

Step 2. Randomly select two jobs with the probability of 0.3, and then exchange all the operation genes of the two jobs in chromosome. The genes of the corresponding machine parts are then randomly generated.

If so, accept the new feasible solution; otherwise, refuse to accept the new feasible solution.

3.3. *L-F Algorithm.* For solving the extended MapReduce scheduling problem, an algorithm based on Least Processing Time first (LPT) rule and First Come, First Served (FCFS) rule is proposed, which is called an *L-F* algorithm. We divide all the extended MapReduce jobs' processing process into three stages. The first stage is dedicated to the open-map task, the second stage is for the first series-reduce task, and the third stage is for the second series-reduce.

When an operation is processed by a machine which takes the longest processing time to process this operation in the machine set, we call it processed in the worst case. For example, if the processing time of an operation is 10 on machine 1 and 5 on machine 2, then the processing time in the worst case is 10. In order that a job with a longer processing time in the worst case can be processed on a faster machine, in the first stage, all the map operations of all jobs are processed in the sequence of processing time under the worst case from the longest to the shortest. When starting to process each operation, we choose the machine with the minimum completion time for the operation. For the operations which belong to the second and third stages of the process, we sort them according to the principle of first come, first served. Among all the jobs, open-map task of a job is the first to be completed, and the first series-reduce task of the job is the first to start being processed. Similarly, the first series-reduce task of a job is the first to be completed, and the second series-reduce task of the job is the first to start being processed.

The specific steps of *L-F* algorithm are as follows:

Step 1. According to the processing time in the worst case, the map operations are sorted in a nonincreasing order. Each operation is processed on the machine that can complete it in the shortest completion time.

Step 2. After all the map operations of a job is completed, the first reduce operation 10 of the job starts to be processed on the available machine that can complete it in the shortest completion time.

Step 3. After the first reduce operation of a job is completed, the second reduce operation starts to be processed on the available machine that can complete it in the shortest completion time.

4. Numerical Experiments

In this section, in order to compare the performance of the *L-F* algorithm, the genetic algorithm, and the simulated annealing algorithm, we carry out three numerical experiments in different scales by MATLAB R2014b on a PC with Intel Core i5. This paper focuses on solving the extended MapReduce scheduling problem with two reduce tasks; thus, in numerical experiments, only two reduce tasks are included in each job.

In small-scale instances, we adopt CPLEX to solve the mathematical model. CPLEX is one of the important tools for solving mixed linear programming problems [23]. In the

genetic algorithm, based on parameter of experiments, we assume that the initial population number is 100, the number of iterations is 200, the crossover rate is 0.3, and the mutation rate is 0.7. In the simulated annealing algorithm, we assume that the initial temperature is 1000, the minimum temperature is 0.001, the cooling rate is 0.1, and the number of iterations at each temperature is 100. In addition, we assume that the solution value obtained by one heuristic algorithm is obj , and the optimal solution obtained by CPLEX is obj^* ; then, there is $gap = (obj - obj^*)/obj^*$.

4.1. *Small-Scale Instances.* We first carry out numerical experiments on small-scale instances. In each instance, we first randomly generate 10 sets of processing time data for operations of all jobs from 1 to 5. We then calculate the average of the objective values, the gap value, and the running time of CPLEX, the genetic algorithm, the simulated annealing algorithm, and *L-F* algorithm; the results are shown in Table 1. Besides, we also randomly generate 10 sets of processing time data for operations of all jobs from 1 to 10 for each instance. Similarly, we then calculate the average of the objective values, the gap values, and the running times of CPLEX and the three heuristic algorithms; the results are shown in Table 2.

From the small-scale instances, it can be seen that when the range of the processing time of operations is enlarged from (1, 5) to (1, 10), the running time of CPLEX increases significantly. While the running time of the genetic algorithm, the simulated annealing algorithm, and *L-F* algorithm does not change significantly. From the aspect of gap value, when the range of the processing time of operations is enlarged, the gap values of the genetic algorithm and simulated annealing algorithm decrease slightly. However, the gap value of *L-F* algorithm increases slightly with the increase of processing time of all the operations.

Besides, Tables 1 and 2 report that the running times of all the three heuristic algorithms are significantly less than CPLEX. *L-F* algorithm is slightly better than the genetic algorithm and the simulated annealing algorithm in solution quality and computational time. In addition, the genetic algorithm is better than the simulated annealing algorithm in terms of running time. The gap value of the genetic algorithm is slightly smaller than that of the simulated annealing algorithm. Specifically, when the number of machines is 2, the genetic algorithm and the simulated annealing algorithm are better than *L-F* algorithm in solution quality. In general, gap values of the genetic algorithm and the simulated annealing algorithm are proportional to the number of machines and jobs. However, with the increase of the number of machines and jobs, the gap value of *L-F* algorithm decreases in general.

4.2. *Large-Scale Instances.* In this section, for comparing the performance of the genetic algorithm, the simulated annealing algorithm, and *L-F* algorithm, we carry out numerical experiments on large scale. In this numerical experiment, we also randomly generate 10 sets of data in a range of 1 to 10 of processing time of operations of all jobs under the different scale of machines and jobs. Due to the fact that

TABLE 1: Experimental results for small-scale instances of the processing time of operations range from 1 to 5.

(m, n, k)	CPLEX		GA			SA			L-F		
	obj*	Time	obj	Time	gap	obj	Time	gap	obj	Time	gap
(2, 3, 3)	10.90	0.24	11.40	1.75	0.05	11.50	2.75	0.06	12.40	0.00	0.14
(2, 3, 4)	14.40	5.24	15.20	2.18	0.06	15.30	3.17	0.06	17.00	0.00	0.18
(2, 3, 5)	16.00	45.95	16.90	2.53	0.06	17.60	3.63	0.10	18.40	0.00	0.15
(2, 4, 3)	13.50	7.14	13.80	2.07	0.02	13.90	3.12	0.03	15.00	0.00	0.11
(2, 4, 4)	16.70	545.07	18.10	2.61	0.08	18.40	3.82	0.10	19.70	0.00	0.18
(2, 5, 3)	16.70	32.14	17.30	2.42	0.04	17.10	3.67	0.02	17.70	0.00	0.06
(3, 4, 3)	8.00	1.07	9.50	2.07	0.19	9.90	3.22	0.24	9.40	0.00	0.18
(3, 4, 4)	10.60	112.79	12.80	2.61	0.21	13.60	3.80	0.28	13.30	0.00	0.25
(3, 5, 3)	9.40	25.75	11.30	2.41	0.20	11.30	3.61	0.20	11.10	0.00	0.18
(3, 6, 3)	11.60	7231.70	13.30	2.74	0.15	13.90	4.02	0.20	13.90	0.00	0.20
(4, 5, 3)	7.30	1.35	9.40	2.43	0.29	9.90	3.58	0.36	8.20	0.00	0.12
(4, 5, 4)	9.30	243.26	13.40	3.13	0.44	14.10	4.46	0.52	12.60	0.00	0.35
(4, 6, 3)	8.30	49.19	10.90	2.82	0.31	11.80	4.04	0.42	9.50	0.00	0.14
(5, 5, 4)	7.50	12.23	11.20	3.13	0.49	12.60	4.40	0.68	9.40	0.00	0.25
(5, 6, 3)	6.20	2.52	8.90	2.78	0.44	10.50	3.99	0.69	7.60	0.00	0.23
(5, 6, 4)	8.20	177.18	12.80	3.51	0.56	14.60	4.95	0.78	10.60	0.00	0.29
(5, 7, 3)	6.90	251.84	11.10	3.16	0.61	11.80	4.19	0.71	8.40	0.00	0.22
(8, 10, 3)	5.70	304.38	11.00	4.16	0.93	12.00	5.71	1.11	6.80	0.00	0.19
Average	10.40	502.83	12.68	2.70	0.28	13.32	3.90	0.36	12.28	0.00	0.19

Note. gap = (obj - obj*/obj*).

TABLE 2: Experimental results for small-scale instances of the processing time of operations range from 1 to 10.

(m, n, k)	CPLEX		GA			SA			L-F		
	obj*	Time	obj	Time	gap	obj	Time	gap	obj	Time	gap
(2, 3, 3)	19.20	0.24	20.30	1.76	0.06	20.70	2.74	0.08	21.80	0.00	0.14
(2, 3, 4)	24.00	5.93	26.40	2.19	0.10	25.90	3.16	0.12	27.40	0.00	0.14
(2, 3, 5)	34.40	553.70	36.50	2.55	0.06	37.80	3.74	0.10	39.80	0.00	0.16
(2, 4, 3)	21.50	7.93	22.80	2.09	0.06	22.90	3.08	0.07	24.80	0.00	0.15
(2, 4, 4)	30.90	8047.90	32.60	2.63	0.06	34.40	3.77	0.11	37.20	0.00	0.20
(2, 5, 3)	30.90	2910.40	32.00	2.44	0.04	32.00	3.60	0.04	36.30	0.00	0.17
(3, 4, 3)	15.20	1.20	28.20	2.11	0.20	18.60	3.15	0.22	17.30	0.00	0.14
(3, 4, 4)	19.00	32.20	23.70	2.65	0.25	25.10	3.80	0.32	23.30	0.00	0.23
(3, 5, 3)	16.60	251.24	19.70	2.45	0.19	20.80	3.60	0.25	20.00	0.00	0.20
(3, 6, 3)	19.30	2181.60	22.80	2.80	0.18	24.00	4.15	0.24	22.80	0.00	0.18
(4, 5, 3)	13.10	1.12	15.50	2.44	0.18	17.10	3.60	0.31	13.70	0.00	0.05
(4, 5, 4)	15.10	209.78	22.40	3.10	0.48	24.70	4.12	0.64	18.90	0.00	0.25
(4, 6, 3)	14.10	35.06	18.00	2.80	0.28	19.50	4.07	0.38	16.40	0.00	0.23
(5, 5, 4)	12.00	9.94	19.40	3.09	0.62	21.90	4.48	0.83	14.50	0.00	0.21
(5, 6, 3)	11.60	3.03	16.00	2.76	0.38	17.90	4.03	0.54	12.90	0.00	0.11
(5, 6, 4)	11.80	22.69	18.20	3.12	0.54	20.20	4.43	0.71	13.70	0.00	0.16
(5, 7, 3)	10.60	77.17	17.20	3.13	0.62	20.50	4.52	0.93	13.10	0.00	0.24
(8, 10, 3)	9.10	101.94	18.70	4.13	1.05	21.40	5.81	1.35	10.70	0.00	0.18
Average	18.11	802.95	22.24	2.68	0.30	23.69	3.89	0.40	21.37	0.00	0.17

Note. gap = (obj - obj*/obj*).

CPLEX takes a lot of time to solve the problem, we calculate gap value by replacing the exact objective value with the minimum objective value obtained by the three heuristic algorithms. In Table 3, we further expand the scale of jobs and machines to compare the three heuristic algorithms.

From Table 3, we can see that L-F algorithm has better performance than the genetic algorithm and the simulated annealing algorithm in large-scale instances. The gap value of the genetic algorithm and the simulated annealing algorithm is proportional to the number of operations.

Otherwise, when the number of jobs and operations is constant, the gap values of the genetic algorithm and the simulated annealing algorithm are proportional to the number of machines.

Table 3 also reports that when the number of machines increases from 4 to 6, the gap values of the genetic algorithm and the simulated annealing algorithm increase significantly. However, when the number of operations of each job and machines is constant, with the increase of the number of jobs, the gap value of L-F algorithm decreases, while the gap

TABLE 3: Experimental results for large-scale instances.

(m, n, k)	GA			SA			$L-F$		
	obj	Time	gap	obj	Time	gap	obj	Time	gap
(2, 30, 3)	178.70*	12.17	0.00	183.90	15.96	0.03	191.50	0.00	0.07
(2, 30, 5)	323.90*	21.53	0.00	332.10	27.28	0.03	333.50	0.01	0.03
(2, 30, 10)	683.40	55.08	0.12	696.80	65.45	0.14	611.60*	0.01	0.00
(2, 50, 3)	312.80*	21.03	0.05	318.30	27.29	0.02	327.00	0.01	0.05
(2, 50, 5)	563.20*	47.70	0.00	572.80	50.83	0.02	565.40	0.01	0.00
(2, 50, 10)	1204.10	117.87	0.17	1193.80	138.27	0.16	1029.90*	0.01	0.00
(2, 100, 3)	671.30	51.13	0.02	662.70	64.93	0.01	658.00*	0.01	0.00
(2, 100, 5)	1181.30	111.44	0.07	1168.30	135.79	0.06	1103.70*	0.01	0.00
(2, 100, 10)	2536.50	376.14	0.21	2480.60	435.82	0.18	2102.60*	0.03	0.00
(4, 30, 3)	87.80	11.85	0.19	95.10	15.66	0.29	73.70*	0.00	0.00
(4, 30, 5)	164.80	21.72	0.28	173.60	27.71	0.35	128.50*	0.01	0.00
(4, 30, 10)	368.70	55.72	0.49	378.10	66.39	0.53	247.20*	0.01	0.00
(4, 50, 3)	151.30	21.28	0.31	157.80	27.58	0.37	115.20*	0.01	0.00
(4, 50, 5)	280.60	41.30	0.39	291.90	51.65	0.45	201.50*	0.01	0.00
(4, 50, 10)	652.60	116.99	0.57	638.80	137.93	0.60	398.90*	0.02	0.00
(4, 100, 3)	338.70	53.83	0.47	333.50	68.64	0.44	231.10*	0.01	0.00
(4, 100, 5)	606.30	116.50	0.47	604.50	141.69	0.47	411.80*	0.02	0.00
(4, 100, 10)	1329.60	376.39	0.58	1301.20	433.62	0.62	821.20*	0.05	0.00
(6, 30, 3)	62.30	21.29	0.64	68.60	16.10	0.81	37.90*	0.01	0.00
(6, 30, 5)	114.50	22.41	0.77	125.70	28.28	0.94	64.70*	0.01	0.00
(6, 30, 10)	265.60	57.08	0.71	274.70	67.99	0.77	155.30*	0.01	0.00
(6, 50, 3)	109.80	21.16	0.70	113.80	27.39	0.76	64.50*	0.01	0.00
(6, 50, 5)	198.90	42.64	0.81	207.60	53.06	0.89	109.70*	0.01	0.00
(6, 50, 10)	447.00	118.65	0.94	457.50	139.43	0.99	230.40*	0.03	0.00
(6, 100, 3)	232.80	53.58	0.89	233.60	68.20	0.89	123.40*	0.01	0.00
(6, 100, 5)	416.30	114.81	0.88	421.50	139.36	0.90	221.80*	0.03	0.00
(6, 100, 10)	923.20	363.43	1.01	907.60	417.32	0.98	458.20*	0.06	0.00
Average	532.56	89.89	0.43	533.16	107.02	0.47	408.08	0.02	0.01

Note. gap = (obj - obj*)/obj*.

value of the genetic algorithm and the simulated annealing algorithm does not change significantly.

5. Conclusion

In this paper, we study an extended MapReduce scheduling model with one open-map task and several series-reduce tasks. Different from the classical MapReduce scheduling problem, we assume that the map operations cannot be processed in parallel and the available machines are unrelated machines. We then propose a genetic algorithm and a simulated annealing algorithm and design $L-F$ algorithm to solve the problem. Finally, compared with the genetic algorithm and the simulated annealing algorithm, $L-F$ algorithm has better performance in large-scale instances.

We propose two future research directions for studying this problem. First, when different operations or jobs are processed on the same machine, there may be setup time. Future research on this problem can take setup time into account. Second, considering preemptive processing, this may be more realistic in some actual production scenarios.

Data Availability

The parameter setting data involved are available, and the data involved in the numerical experiments are all obtained from the solver and also included within the article.

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 (Grant nos. 71771048, 71832001, 71531011, and 71571134) and the Fundamental Research Funds for the Central Universities (Grant no. 2232018H-07).

References

- [1] A. Vital-Soto, A. Azab, and M. F. Baki, "Mathematical modeling and a hybridized bacterial foraging optimization algorithm for the flexible job-shop scheduling problem with sequencing flexibility," *Journal of Manufacturing Systems*, vol. 54, pp. 74–93, 2020.
- [2] J. Dean and S. Ghemawat, "MapReduce," *Communications of the ACM*, vol. 53, no. 1, pp. 72–77, 2010.
- [3] Y. Ji, L. Tong, T. He, K.-W. Lee, and L. Zhang, "Improving multi-job MapReduce scheduling in an opportunistic environment," in *Proceedings of the 2013 IEEE Sixth International Conference on Cloud Computing*, pp. 9–16, Santa Clara, CA, USA, June 2013.
- [4] H. Chang, M. Kodialam, R. R. Kompella, T. V. Lakshman, M. Lee, and S. Mukherjee, "Scheduling in MapReduce-like

- systems for fast completion time,” in *Proceedings of the IEEE INFOCOM*, pp. 3074–3082, Shanghai, China, April 2011.
- [5] J. Huang, F. Zheng, Y. Xu, and M. Liu, “Online MapReduce processing on two identical parallel machines,” *Journal of Combinatorial Optimization*, vol. 35, no. 1, pp. 216–223, 2017.
- [6] B. Moseley, A. Dasgupta, R. Kumar, and T. Sarlóš, “On scheduling in map-reduce and flow-shops,” in *Proceedings of the 23rd ACM Symposium on Parallelism in Algorithms and Architectures—SPAA ’11*, pp. 289–298, San Jose, CA, USA, June 2011.
- [7] T. Luo, Y. Zhu, W. Wu, Y. Xu, and D.-Z. Du, “Online makespan minimization in MapReduce-like systems with complex reduce tasks,” *Optimization Letters*, vol. 11, no. 2, pp. 271–277, 2015.
- [8] T. Condie, N. Conway, P. Alvaro, and J. M. Hellerstein, “MapReduce online,” in *Proceedings of the USENIX Conference on Networked Systems Design & Implementation*, San Jose, CA, USA, April 2010.
- [9] Y. Jiang, W. Zhou, and P. Zhou, “An optimal preemptive algorithm for online MapReduce scheduling on two parallel machines,” *Asia-Pacific Journal of Operational Research*, vol. 35, no. 3, pp. 1–11, 2018.
- [10] Y. Jiang, P. Zhou, T. C. E. Cheng, and M. Ji, “Optimal online algorithms for MapReduce scheduling on two uniform machines,” *Optimization Letters*, vol. 13, no. 7, pp. 1663–1676, 2019.
- [11] D. Cheng, X. Zhou, Y. Xu, L. Liu, and C. Jiang, “Deadline-aware MapReduce job scheduling with dynamic resource availability,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 30, no. 4, pp. 814–826, 2019.
- [12] Z. Tang, J. Zhou, K. Li, and R. Li, “MTSD: a task scheduling algorithm for MapReduce base on deadline constraints,” in *Proceedings of the 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & Ph. D. Forum*, Shanghai, China, May 2012.
- [13] X. Li, T. Jiang, and R. Ruiz, “Heuristics for periodical batch job scheduling in a MapReduce computing framework,” *Information Sciences*, vol. 326, no. 1, pp. 119–133, 2016.
- [14] D. Fotakis, I. Milis, O. Papadigenopoulos, E. Zampetakis, G. Zois, and Z. S. Georgios, “Scheduling MapReduce jobs and data shuffle on unrelated processors,” *Experimental Algorithms*, vol. 9125, pp. 137–150, 2015.
- [15] U. Aickelin and K. A. Dowsland, “An indirect Genetic Algorithm for a nurse-scheduling problem,” *Computers & Operations Research*, vol. 31, no. 5, pp. 761–778, 2004.
- [16] Z. Zakaria and S. Petrovic, “Genetic algorithms for match-up rescheduling of the flexible manufacturing systems,” *Computers & Industrial Engineering*, vol. 62, no. 2, pp. 670–686, 2012.
- [17] A. Hassani and J. Treijs, “An overview of standard and parallel genetic algorithms,” in *Proceedings of the IDT Workshop on Interesting Results in Computer Science and Engineering*, pp. 1–7, Västerås, Sweden, 1975.
- [18] F. Pezzella, G. Morganti, and G. Ciaschetti, “A genetic algorithm for the flexible job-shop scheduling problem,” *Computers & Operations Research*, vol. 35, no. 10, pp. 3202–3212, 2008.
- [19] J. Ji, M. Mika, R. Rafał, W. Grzegorz, and W. Jan, “Simulated annealing for multi-mode resource-constrained project scheduling,” *Annals of Operations Research*, vol. 102, no. 1–4, pp. 137–155, 2001.
- [20] M. Kolonko, “Some new results on simulated annealing applied to the job shop scheduling problem,” *European Journal of Operational Research*, vol. 113, no. 1, pp. 123–136, 1999.
- [21] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by simulated annealing,” *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
- [22] D. Bertsimas and J. Tsitsiklis, “Simulated annealing,” *Statistical Science*, vol. 8, no. 1, pp. 10–15, 1993.
- [23] Y. Shinano and T. Fujie, “ParaLEX: a parallel extension for the CPLEX mixed integer optimizer,” in *Recent Advances in Parallel Virtual Machine and Message Passing Interface*, pp. 97–106, Springer, Berlin, Germany, 2007.

Research Article

A DE-LS Metaheuristic Algorithm for Hybrid Flow-Shop Scheduling Problem considering Multiple Requirements of Customers

Yingjia Sun ¹ and Xin Qi²

¹University of Science and Technology of China, Hefei 230009, China

²Baoshan District People's Government, Shanghai 201999, China

Correspondence should be addressed to Yingjia Sun; sunyingjia@joyintech.com

Received 14 April 2020; Revised 5 May 2020; Accepted 3 June 2020; Published 15 July 2020

Academic Editor: Lu Zhen

Copyright © 2020 Yingjia Sun and Xin Qi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we address a hybrid flow-shop scheduling problem with the objective of minimizing the makespan and the cost of delay. The concerned problem considers the diversity of the customers' requirements, which influences the procedures of the productions and increases the complexity of the problem. The features of the problem are inspired by the real-world situations, and the problem is formulated as a mixed-integer programming model in the paper. In order to tackle the concerned problem, a hybrid metaheuristic algorithm with Differential Evolution (DE) and Local Search (LS) (denoted by DE-LS) has been proposed in the paper. The differential evolution is a state-of-the-art metaheuristic algorithm which can solve complex optimization problem in an efficient way and has been applied in many fields, especially in flow-shop scheduling problem. Moreover, the study not only combines the DE and LS, but also modifies the mutation process and provides the novel initialization process and correction strategy of the approach. The proposed DE-LS has been compared with four variants of algorithms in order to justify the improvements of the proposed algorithm. Experimental results show that the superiority and robustness of the proposed algorithm have been verified.

1. Introduction

Flow shop involves a large amount of jobs and machines, and the jobs need to be processed in machines, which formulates a series of stages in the factories. Flow-shop scheduling problem (FSP) [1, 2] has been widely investigated by the researchers, which aims to find the optimal schedule in the production procedures. The FSP, with reference to many constraints of resources, the number of jobs and machines, is a complex combination optimization problem [3, 4]. The research on FSP can be traced back to 1954 [5], in which an exact algorithm has been proposed to deal with the proposed three-machine flow-shop scheduling problem. Then, many scholars have devoted enormous effort and time to the research of FSP. For example, Riahi and Kazemi [6] have proposed a no-wait flow-shop scheduling problem,

which is a variant of FSP, considering the makespan and flow time. In the paper, the authors have provided the mathematical formulation and a hybrid metaheuristic algorithm combined ant colony optimization and simulated annealing algorithm. Marichelvam et al. [7] address a flow-shop scheduling problem with minimization of the makespan and total flow time, which is a biobjective flow-shop scheduling problem. In order to solve the proposed problem, a hybrid monkey search algorithm based on a subpopulation has been studied.

It is worth mentioning that hybrid flow-shop scheduling problem (HFSP) is one type of FSP, which need to schedule the jobs in machines and complete the assignments of machines [8–10]. In HFSP, there are parallel machines in a stage, which has higher requirements for the schedule. Because redundancy and shortage of machines both can reduce

the efficiency of the production and increase the cost. In addition, hybrid flow-shop scheduling problem has been proved the NP-hard [11, 12]. For instance, Yu et al. [13] propose a HFSP to minimize the total tardiness in consideration of several practical assumptions. Moreover, a genetic algorithm has been utilized to tackle the proposed problem, which involves a novel decoding approach to get the objective. Pan et al. [14] provide nine algorithms to solve the proposed HFSP, and the objective of the concerned problem is to minimize the makespan. Zhang et al. [12] study the HSFP with a lot streaming, and the objective is to minimize the total flow time. A mathematical formulation is provided for the studied problem, and an effective modified birds optimization algorithm has been applied to obtain the results. In this paper, a hybrid flow-shop scheduling problem has been studied, which considers many stages with parallel machines and multiple requirements of customers. Different requirements of customers indicate that the manufacturer needs to produce different types of productions in a horizon. The diversity of productions reflects on the procedures in the manufacturer. For example, two productions need to be processed in different stages. The first one has two procedures and the other one has three procedures. The objective of the proposed problem is to minimize the makespan and the cost of delay. Thus, the schedule should not only consider the flow time but also take into account the time requirements of the customers, which is inspired by real-world situations.

Due to the NP-hard of the HFSP, in many prior studies, metaheuristic algorithms have been applied to tackle the HFSP [6, 11, 13], such as greedy algorithm, simulated annealing algorithm, genetic algorithm, etc. Metaheuristic algorithms can obtain the solution of complexity optimization problem in available time, which have been developed in many fields, i.e., manufacturing industry [15–17], airline industry [18–21], energy sources [22–25], etc. In addition, Differential Evolution (DE), firstly proposed by Storn and Price [26], is a powerful evolutionary algorithm for global optimization. Like most evolutionary algorithms, DE is a population-based algorithm involving three main operations [25, 27, 28], i.e., mutation operations, crossover operation, and selection operation. These operations update the solution by the specific mechanism and help the algorithm to find the global optimal solution. Recently, DE algorithm has been widely studied by many researchers and has shown good performance in solving many optimization problems, especially in flow-shop scheduling problem. For example, Zhang and Xing [29] have utilized DE to solve a distributed limited-buffer flow-shop scheduling problem, with minimization of the makespan. In the study, two heuristics and two differential evolution metaheuristics have been proposed for various situations. Zhou et al. [30] presented a hybrid differential evolution named DE-eEDA, which combined DE with an estimation of distribution algorithm. The authors have applied the proposed DE-eEDA to tackle the reentrant hybrid flow-shop scheduling problem with the objective of minimizing the total weighted completion time. In this paper, a hybrid differential evolution with local search algorithm (denoted by DE-LS) has been

proposed to solve the studied problem, which has modified the mutation process and provides the novel initialization process and correction strategy of the approach. Moreover, the local search has been utilized in the specific situation which can improve the solution by several operators, i.e., 2-opt and swap. Moreover, the comparison between previous studies works and this study is provided in Table 1, which compares the studies from three aspects, i.e., types of problems, objectives, and approaches.

The contributions of our paper can be organized as follows: (1) A hybrid flow-shop scheduling problem with consideration of multiple requirements of customers has been considered, with reference to diverse productions and different procedures. (2) The novel mixed-integer programming model has been provided in the paper to formulate the studied problem. (3) The hybrid metaheuristic algorithm with Differential Evolution and Local Search (DE-LS) has been proposed to tackle the proposed problem, which involves the modified mutation process, initialization process, correction strategy, etc. (4) The proposed DE-LS has been compared with four variant algorithms, and the robustness and superiority of DE-LS have been verified.

The remaining components of this paper can be summarized as follows. The problem description and mathematical formulation are provided in Section 2. In Section 3, the proposed hybrid algorithm DE-LS has been illustrated detailedly, with the encoding strategy and correction strategy, initialization, etc. The results of computational experiments can be seen in Section 4. The conclusion of the paper is presented in Section 5.

2. Problem Description and Mathematical Formulation

The flow-shop scheduling problem has been proved as NP-hard [35]. Moreover, the hybrid flow-shop scheduling problem with two machines in one stage has also been proved NP-hard [36]. In this paper, a hybrid flow-shop scheduling problem with consideration of several machines in one working processing and more than one stage is being studied, which is NP-hard. This study focuses on the multiple demands of customers in the hybrid flow-shop scheduling problem, which is based on the real-world situations. Generally, the manufacturers will revive many bookings from different customers, which have different requirements about the productions. For example, some customers need the productions adding one or more processes based on the basic productions. On the contrary, some of them may need the productions omitting one or more processes based on the basic productions. The differences between customers directly influence the efficiency and the working patterns of the manufacturers. Thus, the paper considers the differential demands of the customers. Moreover, the paper aims to improve the efficiency of the manufacturers and reduce the delay cost. The notations used in the paper can be found in Table 2.

To better understand the proposed problem, an example of the processes of the manufacturing and the delivering is shown in Figure 1. Firstly, the suppliers provide the original

TABLE 1: Comparison between previous studies and this study.

References	Types of problems	Objectives	Approaches
Pan et al. [1]	Lot-streaming FSP	Minimize the total weighted earliness and tardiness penalties	Discrete artificial bee colony
Riahi and Kazemi [6]	No-wait FSP	Minimize the makespan	A hybrid ant colony optimization and simulated annealing algorithm
Marichelvam et al. [7]	FSP	Minimize the makespan and total flow time	A hybrid monkey search algorithm based on subpopulation
Ruize and Stützle [31]	Permutation FSP	Minimize the makespan	Iterated greedy algorithm
Xu et al. [32]	Permutation FSP	Minimize the makespan and total weighted tardiness	Iterated local search
Lei and Zheng [11]	HFSP	Minimize the total tardiness, maximum tardiness, and makespan	A novel neighborhood search with global exchange
Zhang et al. [12]	HFSP	Minimize the total flow time	Effective modified migrating birds optimization
Yu et al. [13]	HFSP	Minimize the total tardiness	Genetic algorithm
Pan et al. [14]	HFSP	Minimize the makespan	Nine effective metaheuristics
Tasgetiren et al. [33]	Blocking FSP	Minimize the makespan	Iterated greedy algorithm
Yagmahan and Yenisey [34]	FSP	Minimize the makespan and total flow time	Ant colony system algorithm
Zhang and Xing [29]	FSP	Minimize the makespan	Differential evolution
Zhou et al. [30]	HFSP	Minimize the total weighted completion time	A hybrid differential evolution with an estimation of distribution algorithm
This study	HFSP	Minimize the makespan and cost of delay	A hybrid differential evolution and local search

TABLE 2: Notations in the paper.

N	The total number of jobs
n	The index of jobs
C	The total number of customers
c	The index of customers
M	The total number of machines
m	The index of machines
J	The total number of working processes
j	The index of working processes
k	The number of production types
c_k	The demands number of customer c for production with type k
D_c	The latest delivery time of customer c
Q_{jm}	1, if the machine m in the j working process; 0, otherwise
P_{kjm}	The processing time of production with type k in the j working process in machine m
B_{njm}	The begin time of job n in the j working process in machine m
C_{njm}	The completion time of job n in the j working process in machine m
Max_m	The maximum number of jobs can be assigned to machine m in the same working process
ρ_c	The penalty cost of customer c if the latest delivery time is not satisfied

materials, and then the following processes can be started. Due to the different requirements of the customers, the jobs need to be processed with the specific sequences of the machines. For example, job 2 and job 3 need to be processed in the first process, while job 1 and job N do not need that. Jobs with different processing sequences can manufacture the productions with multiple types, which aims to satisfy the differential demands of customers.

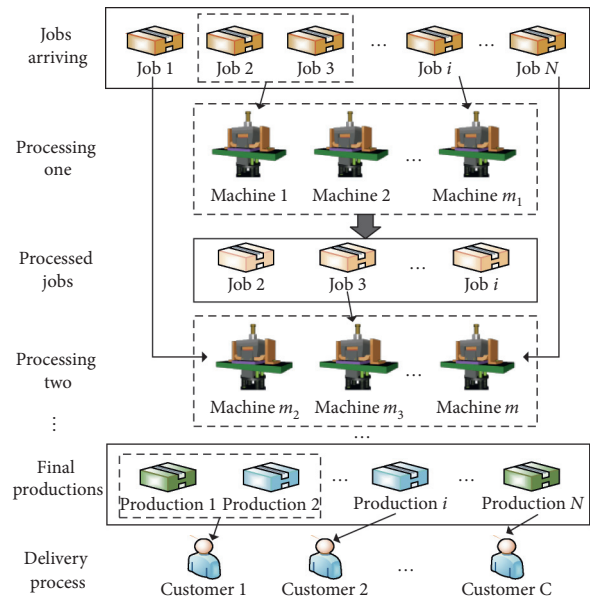


FIGURE 1: An example of the manufacturing and delivery.

In order to study the proposed problem, the following assumptions have been made in this paper:

- (a) The job will not be interrupted when it starts
- (b) The production with the same type is homogeneous
- (c) The machine can process the jobs list with different types of jobs
- (d) The arrival times of all jobs are known and the same.

Decision variables are as follows:

$$\begin{aligned}
x_{njm} &= \begin{cases} 1, & \text{if the job } n \text{ is processed with machine } m \text{ in } j \text{ working process,} \\ 0, & \text{otherwise,} \end{cases} \\
y_{nk} &= \begin{cases} 1, & \text{if the job } n \text{ will be processed to production with type } k, \\ 0, & \text{otherwise,} \end{cases} \\
z_{nc} &= \begin{cases} 1, & \text{if the job } n \text{ is assigned to customer } c, \\ 0, & \text{otherwise,} \end{cases} \\
\varphi_{imm} &= \begin{cases} 1, & \text{if the job } i \text{ is processed before the job } n \text{ in machine } m, \\ 0, & \text{otherwise.} \end{cases}
\end{aligned} \tag{1}$$

The MIP formulation is

subject to

$$\text{Min } C_{\max} + \sum_{c=1}^C \left(\sum_{n=1}^N \sum_{m=1}^M \sum_{j=1}^J \max\{C_{njm} - D_c, 0\} z_{nc} \right) \rho_c, \tag{2}$$

$$C_{\max} = \max_{n \in \{1, 2, \dots, N\}} C_{njm}, \quad \forall j \in \{1, 2, \dots, J\}, \forall m \in \{1, 2, \dots, M\}, \tag{3}$$

$$\sum_{n=1}^N y_{nk} z_{nc} \geq c_k, \quad \forall c \in \{1, 2, \dots, C\}, \forall k \in \{1, 2, \dots, K\}, \tag{4}$$

$$\sum_{n=1}^N x_{njm} \leq \text{Max}_m, \quad \forall j \in \{1, 2, \dots, J\}, \forall m \in \{1, 2, \dots, M\}, \tag{5}$$

$$\sum_{m=1}^M x_{njm} \leq 1, \quad \forall n \in \{1, 2, \dots, N\}, \forall j \in \{1, 2, \dots, J\}, \tag{6}$$

$$\sum_{k=1}^K y_{nk} \leq 1, \quad \forall n \in \{1, 2, \dots, N\}, \tag{7}$$

$$\sum_{c=1}^C z_{nc} \leq 1, \quad \forall n \in \{1, 2, \dots, N\}, \tag{8}$$

$$C_{njm} = B_{njm} + \sum_{k=1}^K P_{kjm} y_{nk}, \quad \forall n \in \{1, 2, \dots, N\}, \forall j \in \{1, 2, \dots, J\}, \forall m \in \{1, 2, \dots, M\}, \tag{9}$$

$$C_{njm} \leq C_{\max}, \quad \forall n \in \{1, 2, \dots, N\}, \forall j \in \{1, 2, \dots, J\}, \forall m \in \{1, 2, \dots, M\}, \tag{10}$$

$$C_{ijm} \leq C_{njm} \varphi_{imm}, \quad \forall i, n \in \{1, 2, \dots, N\}, \forall j \in \{1, 2, \dots, J\}, \forall m \in \{1, 2, \dots, M\}, \tag{11}$$

$$x_{njm}, y_{nk}, z_{nc}, \varphi_{imm} \in \{0, 1\}, \quad \forall i, n \in \{1, 2, \dots, N\}, \forall j \in \{1, 2, \dots, J\}, \forall m \in \{1, 2, \dots, M\}, \forall k \in \{1, 2, \dots, K\}, \forall c \in \{1, 2, \dots, C\}. \tag{12}$$

Objective (2) involves two phases, which are C_{\max} and the penalty part. The definition of C_{\max} is provided in (3). Equation (3) means that C_{\max} is the maximum makespan in all jobs. The penalty part means that it will have an external cost if the production is delayed to each customer. Constraint (4) denotes that the demands of each customer will be

satisfied. Constraint (5) means that the number of jobs assigned to one machine in a working process will not exceed the limitation of the machine. Constraint (6) limits that one job will not be scheduled to more than one machine in each working process. Constraint (7) denotes that only one type can be assigned to a job, and constraint (8) means one job

only can be assigned to one customer. Constraint (9) is the definition of the completion time, which can be calculated by the begin time of the job and the processing time. Constraint (10) means that the completion time of any job is less than C_{\max} . Constraint (11) defines that if two jobs are processed in one machine, the completion time of the latter one is larger than the former one. Four decision variables are defined in constraint (12).

3. The Proposed Hybrid Approach DE-LS

In this section, a hybrid metaheuristic algorithm DE-LS has been proposed, which involves a specific encoding strategy, correction strategy, initialization, etc., for the studied problem. In addition, the mutation strategy of the DE has been modified in order to improve the solution. The framework of the proposed DE-LS has also been provided in the section, which describes all processes and operation logic of the proposed algorithm.

3.1. Encoding Strategy and Correction Strategy. The difference of the productions lies in the processes in the manufacturing. For example, there are four kinds of the productions and three processes in the manufacturing totally. Production A needs to be processed by the first two processes, and production B is processed by the last two processes. The example of the relationship between productions and processes is provided in Figure 2.

Therefore, we define the procedures according to the different processes, and each job corresponds to a procedure. The processes in each procedure is defined before. Firstly, each job must be assigned to one kind of procedure, which will indicate one type of production. Thus, the number of each procedure is limited by the requirements of the customers. We provide an example of solution presentation in Figure 3. In the figure, there are seven jobs, and each job has a specific encoding number, which indicates the assigned procedure of the job. For example, the encoding number of job 3 and job 5 is 3, which means that these two jobs will be processed with procedure 3. The last line of the figure is the customers' requirements, which is known as the limitation of the encoding.

The solutions in this paper are integer and have a specific range. However, noninteger and out-of-range solutions will be generated during the processes of solutions computation. A correction strategy of solution is provided in the paper. Firstly, the range of the solution should be determined, which is related to the number of procedures. For example, there are four procedures totally, and each element in the solutions is integer in $[1, 4]$. Then, correct each element in the solution using round-off method. Moreover, the out-of-range elements are corrected according to the bound of the solution. Finally, count the number of procedures in the solutions and compare them with the customers' requirements. A simple example of the correction strategy is provided in Figure 4. The first line in the picture is the computed solution which is noninteger and did not meet the range of solution. After rounding the solution, a new

solution is generated which is integer and met the range of solution. However, another limitation of the solution is the requirements of customers. The number of procedure 4 in the solution is more than the requirements of customers and the number of procedures 2 and 3 is less than the requirements of customers. Hence, randomly select two jobs with procedure 4 (i.e., job 4 and job 7) and change the procedure of them into 2 and 3, respectively. For a better understanding, we provide the pseudocode of the correction strategy in Algorithm 1.

3.2. Initialization. The initial solution should be generated before the solution changing, which is the available solution of the proposed problem. The initialization of the classical DE generated the initial solution using the minimum and maximum values of the solution. In this paper, the first step of the initialization is also to generate random variables for each solution, which are integer and between the range of the solution. Then, the limitation of the customers' requirements should be considered. The correction strategy can be used to make the solution available. The details of the initialization can be found in Algorithm 2.

3.3. Details of the Proposed DE-LS. The details of the proposed DE-LS can be seen in this section, which involves the mutation strategy, crossover, selection, and local search. The first three phases are the part of the classical DE, and the local search in the paper is to improve the searched solution.

3.3.1. Selective Mutation Strategy. Generally, many mutation operators are widely used in the DE, such as rand/1, rand/2, best/1, etc. The solution can be changed after using the mutation operations, and we call the new solution mutant solution, $V_i = \{v_{i1}, v_{i2}, \dots, v_{ij}, \dots, v_{iD}\}$. Some popular mutation operators can be found as follows:

$$\frac{\text{Rand}}{1}: V_i = X_{r_1} + F \times (X_{r_2} - X_{r_3}), \quad (13)$$

$$\begin{aligned} \frac{\text{Rand}}{2}: V_i &= X_{r_1} + F \times (X_{r_2} - X_{r_3}), V_i \\ &= X_{r_1} + F \times (X_{r_2} - X_{r_3}) + F \times (X_{r_4} - X_{r_5}), \end{aligned} \quad (14)$$

$$\frac{\text{Best}}{1}: V_i = X_{\text{best}} + F \times (X_{r_1} - X_{r_2}), \quad (15)$$

$$\frac{\text{Best}}{2}: V_i = X_{\text{best}} + F \times (X_{r_1} - X_{r_2}) + F \times (X_{r_3} - X_{r_4}), \quad (16)$$

$$\begin{aligned} \frac{\text{Current - to - best}}{1}: V_i &= X_i + F \times (X_{\text{best}} - X_i) \\ &+ F \times (X_{r_1} - X_{r_2}), \end{aligned} \quad (17)$$

where the integer variables $r_1, r_2, r_3, r_4, r_5, i$ are from $[1, NP]$. The variable i is determined, and other five variables are randomly generated. Moreover, they satisfy the inequality $r_1 \neq r_2 \neq r_3 \neq r_4 \neq r_5 \neq i$. X_{best} is the current best solution. The variable F is randomly generated between 0 and 1.

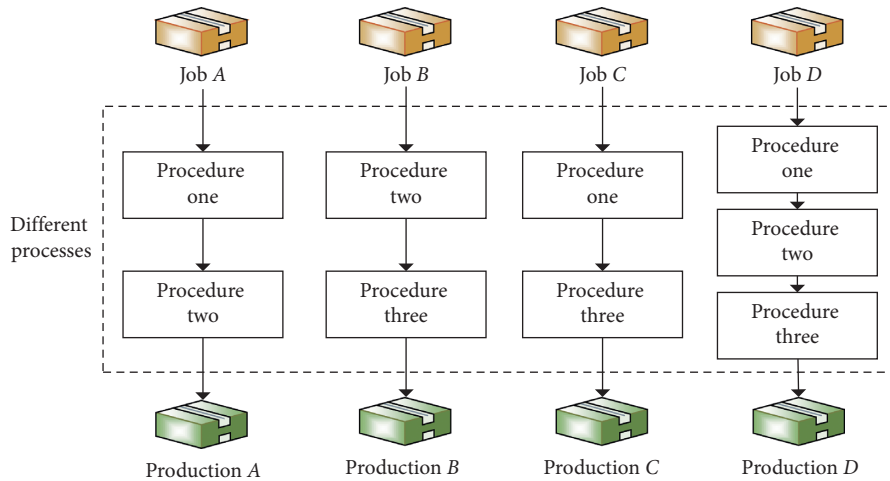
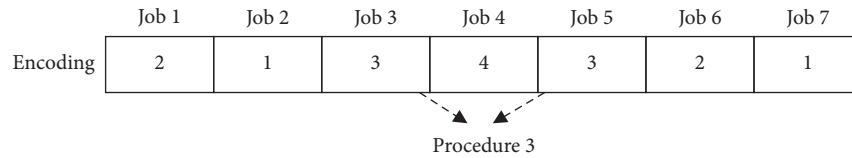
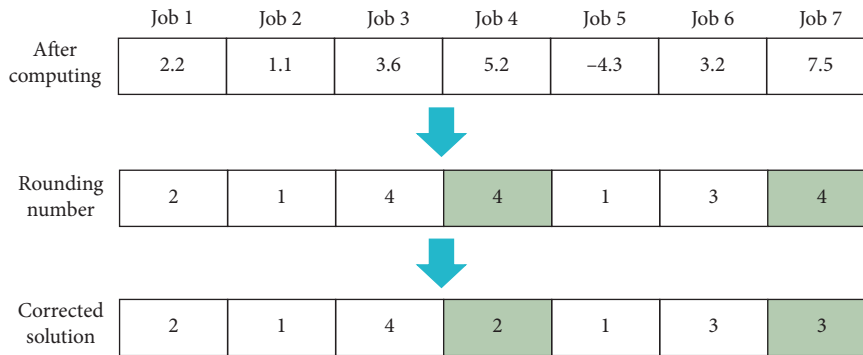


FIGURE 2: An example of the relationship between productions and processes.



Customers' requirements: production A: 2, production B: 2, production C: 2, and production D: 1

FIGURE 3: An example of solution presentation.



Customers' requirements: production A: 2, production B: 2, production C: 3, and production D: 1

FIGURE 4: An example of correction strategy.

In this paper, selective mutation strategy is proposed to update the solutions. For instance, if the best solution is updated in the last iteration, the current-to-best/1 operation will be applied in the mutation process. Otherwise, a random mechanism will be utilized to select the mutation operator. The details of the proposed selective mutation strategy are described as follows:

- (1) If $\text{rand}(0, 1) < CR$ or $j == j_{\text{rand}}$ then
- (2) If X_{best} is updated then
- (3) Apply current-to-best/1: $V_i = X_i + F \times (X_{\text{best}} - X_i) + F \times (X_{r_1} - X_{r_2})$
- (4) Else $s = \text{rand}(0, 1)$
- (5) If $s < S_1$ then

- (6) Apply rand/1: $V_i = X_{r_1} + F \times (X_{r_2} - X_{r_3})$
- (7) Else
- (8) Apply best/1: $V_i = X_{\text{best}} + F \times (X_{r_1} - X_{r_2})$
- (9) End if
- (10) End if
- (11) End if

where the parameters CR and S_1 are random numbers between 0 and 1. The integer parameter j_{rand} is randomly generated between 1 and D .

3.3.2. *Crossover*. The crossover operation in the DE can generate a new solution vector, which is always called trial

Initialization: the solution x , the number of jobs J , the lower-bound of the solution $Lower$, the upper-bound of the solution $Upper$, the list of customers' requirements CR , procedures count list $count$.

```

(1) While  $j \leq J$  then
(2)   if  $x[j] < Lower$  or  $x[j] > Upper$  then
(3)      $x[j] = \text{random}(Lower, Upper)$ 
(4)   End if
(5)    $x[j] = \text{round}(x[j])$ 
(6)    $j = j + 1$ 
(7) End while
(8) While  $j \leq J$  then
(9)    $count[x[j]] = count[x[j]] + 1$ 
(10)   $j = j + 1$ 
(11) End while
(12) *The following processes is used to correct the limitation of customers' requirements
(13) While  $i \leq \text{length}(count)$  then
(14)   if  $count[i] < CR[i]$  then
(15)     For position  $tims = 1: (CR[i] - count[i])$ 
(16)       index = False
(17)       While  $k \leq \text{length}(count)$  and  $in\ de\ x = \text{False}$  then
(18)         if  $count[k] > CR[k]$  then
(19)           For position  $j = 1: J$ 
(20)             if  $x[j] == k$  then
(21)                $x[j] = i$ 
(22)                $count[i] = count[i] + 1$ 
(23)                $count[k] = count[k] - 1$ 
(24)               index = True
(25)             break
(26)           End if
(27)         End for
(28)       Else  $k = k + 1$ 
(29)     End while
(30)   End for
(31) End if
(32)    $i = i + 1$ 
(33) End while

```

ALGORITHM 1: The pseudocode of the correction strategy.

Initialization: the population size NP , the dimension of the solution D , the lower-bound of the solution $X^L = \{x_1^L, x_2^L, \dots, x_D^L\}$, the upper-bound of the solution $X^U = \{x_1^U, x_2^U, \dots, x_D^U\}$, the empty initial solution $X = \{X_1, X_2, \dots, X_i, \dots, X_{NP}\}$, and $X_i = \{x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{iD}\}$.

```

(1) For position  $i = 1: NP$ 
(2)   For position  $j = 1: D$ 
(3)      $x_{ij} = x_j^L + \text{rand}(0, 1) \times (x_j^U - x_j^L)$ 
(4)      $x_{ij} = \text{round}(x_{ij})$ 
(5)   End for
(6) End for
(7) Using the processes of correcting the limitation of customers' requirements which can be found in the Algorithm 1.

```

ALGORITHM 2: The processes of initialization.

vector. Each element in trial vector is selected from the corresponding element which may from solution before mutation or mutated solution. In this paper, binomial crossover operation is applied, and the details can be represented as follows:

$$u_{ij} = \begin{cases} v_{ij}, & \text{if } \text{rand}(0, 1) \leq CR \text{ or } j == j_{\text{rand}}, \\ x_{ij}, & \text{otherwise,} \end{cases} \quad (18)$$

where CR is a random number between 0 and 1, and the integer number j_{rand} is selected from 1 to D . The index $i \in \{1, 2, \dots, NP\}$ means the individual of the solution. The equation indicates that when the random number is less than CR and $j == j_{\text{rand}}$, the element of mutated solution is selected for the trial solution. Otherwise, the element of solution before being mutated is selected for trial solution.

3.3.3. Selection. The selection is the last process of DE in an iteration. The selection mechanism in the paper is the same as that in most studies, which remains the better solution after the solution changing. The selection regulation can be found as follows:

$$X_i = \begin{cases} U_i, & \text{if } \text{obj}(U_i) \leq \text{obj}(X_i), \\ X_i, & \text{otherwise,} \end{cases} \quad (19)$$

where $\text{obj}(U_i)$ and $\text{obj}(X_i)$ are the objective value of the solution U_i and X_i , respectively. The equation means that the solution with lower objective value can survive to the next generation.

3.3.4. Local Search. Local search (LS) has been widely applied to solve complex optimization problems, which can get the optimal solution in an available time. Moreover, local search is to find neighborhoods of the selected solution by using specific neighborhood structure operators, such as, 2-opt, exchange, swap, etc. In order to improve the efficiency of the proposed approach, local search has been used in this paper. The proposed local search is utilized after the selection process, which search the neighborhoods by the improved solutions. In addition, two neighborhood structure operations have been applied in the paper, i.e., 2-opt and swap. These two operations are very popular and simple. In Figure 5, the changing mechanisms of 2-opt and swap operations have been illustrated. For example, two jobs need to be selected before the operations. In the example, job 2 and job 5 have been selected for the 2-opt operation, and job 1 and job 7 have been selected for the swap operation. Then, the 2-opt operation is to convert all jobs between the selected two jobs as seen in the figure. The swap operation is to exchange the value of the selected two jobs. We provide the pseudocode of the proposed local search, which can be found in Algorithm 3.

3.3.5. The Proposed DE-LS. In this section, the proposed DE-LS approach is described in Algorithm 4, which involves the basic processes of DE, i.e., mutation, crossover, and selection, and a proposed local search is also embedded.

4. Computational Experiments

In this section, the performance of the proposed algorithm will be evaluated by 18 randomly generated instances. The computational experiments are set with the number of jobs $n \in \{20, 30, 40, 50, 60, 70, 80, 90, 100\}$. We classify the number of jobs into two scales, i.e., the small scale and large scale.

The small scale has number of jobs more than 50; otherwise, when the number of jobs is more than 50, it is the large scale. Moreover, there are two scales for the number of machines for each number of jobs, and we define each instance according to the number of jobs and machine. For example, the first instance is defined as J30-M3, which means the problem setting for this instance is 30 jobs and 3 machines. The processing time of production with each type in each working process is generated from a discrete uniform distribution. All experiments are implemented in Python and run on a PC with 8 GB of RAM memory, 64-bit operating system, and Intel® Core™ i7-6600U CPU @ 2.60 GHz 2.81 GHz.

The study proposes a DE-LS to tackle the concerned hybrid flow-shop scheduling problem, which combined the metaheuristic differential evolution algorithm with local search in order to improve the quality of the found solution and solve the study problem efficiently. Moreover, A selective mutation strategy of differential evolution has been provided, such as mutation strategy. Therefore, four algorithms have been compared with the proposed DE-LS, i.e., without LS, best/1, current-to-best/1, and rand/1, which are the variants of the proposed DE-LS. In other words, the without LS algorithm is the proposed DE-LS while without using local search. The other three algorithms, best/1, current-to-best/1, and rand/1, are different in their mutation processes and do not have the selected mechanism; for example, the best/1 algorithm only uses the best/1 mutation operation in the mutation process. In addition, each instance has been executed 10 times for avoiding the random results. The mutant factor F is randomly generated from $(F', 1)$, and the parameter F' is set as 0.4. The crossover control parameter CR is randomly generated from $(CR', 1)$, and the parameter CR' is set as 0.6. The population size NP is set as 10, and the dimensions of solution D is equal to the number of jobs. The local search selection is limited by the parameter S' , which is set as 0.5.

The maximum number of iterations for each instance is 200, which is the terminal condition for each algorithm. The objective (denoted by Obj) of each algorithm in each instance is the average objective value of 10 runtimes. The minimum value (denoted by Min) is the minimum objective that the algorithm could find in 10 runtimes. The relative percentage deviation (denoted by RPD) is used as the performance measure in the paper which is widely used to evaluate the performance of algorithms [31, 33]. The RPD can be calculated by the given formulation as follows:

$$\text{RPD}_i = \frac{Z(\text{Objective}_i) - Z(\text{Minimum})}{Z(\text{Minimum})} * 100, \quad (20)$$

where $Z(\text{Objective}_i)$ is the average objective value of the algorithm i . $Z(\text{Minimum})$ denotes the minimum objective has been found in all the compared algorithms for the same instance. RPD_i is the RPD value of the specific algorithm i . Moreover, the standard deviation (denoted by SD) is also applied to measure the robustness of the compared algorithm, which is calculated as follows:

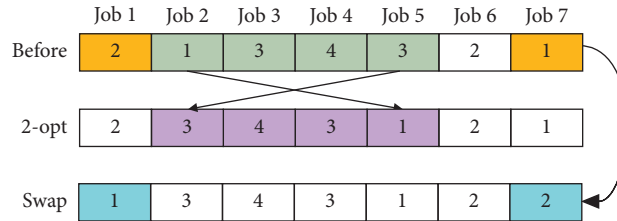


FIGURE 5: An example of 2-opt and swap operations.

Input: solution $X_i = \{x_1, x_2, \dots, x_j, \dots, x_D\}$, the objective value $\text{obj}(X_i)$, maximum number of iterations T , pre-defined index $\text{index} \in [0, 1]$.

- (1) While $t < T$ then
- (2) $\text{opt} = \text{rand}(0, 1)$
- (3) $j_1 = \text{randint}(1, D - 1)$
- (4) $j_2 = \text{randint}(r_1, D)$
- (5) If $\text{opt} < \text{index}$ then
- (6) Get the new solution X'_i by using 2-opt operation, and the selected jobs are j_1 and j_2
- (7) Else
- (8) Get the new solution X'_i by using swap operation, and the selected jobs are j_1 and j_2
- (9) End if
- (10) Calculate the objective value $\text{obj}(X'_i)$
- (11) If $\text{obj}(X_i) > \text{obj}(X'_i)$ then
- (12) Update $X_i = X'_i$
- (13) End if
- (14) $t = t + 1$
- (15) End while

ALGORITHM 3: The proposed local search.

Initialization: the population size NP , the dimension of the solution D , the lower-bound of the solution $X^L = \{x_1^L, x_2^L, \dots, x_D^L\}$, the upper-bound of the solution $X^U = \{x_1^U, x_2^U, \dots, x_D^U\}$, the empty initial solution $X = \{X_1, X_2, \dots, X_i, \dots, X_{NP}\}$, and $X_i = \{x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{iD}\}$, maximum number of iterations Iter^{\max} , F' , CR' , and S' are the pre-defined parameters.

- (1) Apply the initialization process to get the initial solution X and the objective value $\text{obj}(X)$
- (2) Find the current best solution X_{best} and $\text{obj}(X_{\text{best}})$
- (3) While $\text{iter} < \text{Iter}^{\max}$ then
- (4) For position $i = 1: NP$ then
- (5) $F = \text{rand}(F', 1)$
- (6) $CR = \text{rand}(CR', 1)$
- (7) For position $j = 1: D$ then
- (8) *The processes of mutation
- (9) If $\text{rand}(0, 1) < CR$ or $j == j_{\text{rand}}$ then
- (10) Generate $r_1 \neq r_2 \neq r_3 \in \{1, 2, \dots, NP\}$
- (11) If X_{best} is updated then
- (12) $u_{ij} = x_{ij} + F \times (x_{\text{best},j} - x_{ij}) + F \times (x_{r_1,j} - x_{r_2,j})$
- (13) Else $s = \text{rand}(0, 1)$
- (14) If $s < S'$ then
- (15) $u_{ij} = x_{r_1,j} + F \times (x_{r_2,j} - x_{r_3,j})$
- (16) Else
- (17) $u_{ij} = x_{\text{best},j} + F \times (x_{r_1,j} - x_{r_2,j})$
- (18) End if
- (19) End if
- (20) Else
- (21) $u_{ij} = x_{ij}$
- (22) End if
- (23) End for
- (24) Using correction strategy
- (25) Get the objective value $\text{obj}(U_i)$

ALGORITHM 4: Continued.

```

(26)   If  $\text{obj}(U_i) \leq \text{obj}(X_i)$  then
(27)      $X_i = U_i$ 
(28)   Else
(29)     Apply the proposed local search which can be found in Algorithm 3 to update the solution  $X_i$ 
(30)   End if
(31) End for
(32) Search the new best solution and update the solution  $X_{\text{best}}$  and  $\text{obj}(X_{\text{best}})$ 
(33)  $\text{iter} = \text{iter} + 1$ 
(34) End while

```

ALGORITHM 4: The proposed DE-LS.

TABLE 3: Computational results about objective.

Instances	Types	DE-LS		Without LS		Best/1		Current-to-best/1		Rand/1	
		Obj	Min	Obj	Min	Obj	Min	Obj	Min	Obj	Min
1	J20-M3	104.68	104.68	105.25	104.68	105.44	104.68	105.15	104.68	104.79	104.68
2	J20-M4	79.60	79.60	80.47	79.60	81.17	79.60	80.11	79.60	79.90	79.60
3	J30-M3	134.63	134.46	137.30	134.99	137.90	134.99	136.51	134.46	136.14	134.98
4	J30-M4	120.55	120.51	121.53	120.53	123.46	122.37	121.71	120.56	120.88	120.51
5	J40-M3	196.20	195.89	198.10	196.51	201.96	198.00	199.40	196.59	198.28	196.70
6	J40-M4	155.23	154.96	157.66	155.62	158.81	156.54	156.12	155.28	155.89	154.96
7	J50-M3	229.65	229.08	230.94	229.05	232.62	230.61	231.07	229.25	230.16	229.21
8	J50-M4	207.54	206.65	209.93	207.63	213.70	210.03	208.81	207.32	208.56	207.11
9	J60-M3	283.95	282.97	286.76	284.36	289.78	286.25	285.62	284.11	286.50	283.53
10	J60-M4	255.86	254.68	257.96	255.77	265.17	261.83	261.88	256.21	260.33	257.66
11	J70-M3	324.68	321.56	329.79	325.87	334.77	328.82	328.82	325.60	332.79	330.23
12	J70-M4	337.72	336.66	343.81	340.52	351.49	345.41	346.60	343.06	345.64	342.46
13	J80-M3	414.80	411.70	423.99	418.48	431.32	424.06	423.16	419.02	429.11	420.77
14	J80-M4	362.54	359.00	367.54	364.58	370.72	365.28	369.02	361.06	372.58	364.22
15	J90-M3	399.48	398.22	400.91	399.24	403.39	399.32	400.82	398.44	405.90	402.07
16	J90-M4	366.55	364.21	373.06	367.38	379.74	374.52	375.61	372.49	384.60	377.17
17	J100-M3	432.39	430.50	434.83	433.20	439.23	437.19	435.94	430.04	443.80	438.62
18	J100-M4	463.80	460.39	470.86	465.35	477.93	467.42	471.54	463.94	477.54	471.25
Average		270.55	269.21	273.93	271.30	277.70	273.72	274.33	271.21	276.30	273.10

$$SD_i = \sqrt{\frac{\sum_{r=1}^{\text{Runtimes}} (Z(\text{Objective}_i) - Z(\text{Minimum}))^2}{\text{Runtimes}}}, \quad (21)$$

where $Z(\text{Objective}_i)$ means the objective value of algorithm i in r th runtimes, and then $Z(\text{Minimum})$ denotes the minimum objective has been found in all the compared algorithms for the same instance. The standard deviation value of algorithm i is denoted by SD_i , which measures the deviation between the found optimal solution and the calculated solutions of algorithm i in all runtimes. The smaller the value of SD, the more robust the algorithm.

In Table 3, the average objective value and minimum objective value of each compared algorithm for each instance have been provided. We can find that the proposed DE-LS has found the optimal solution in all instances. Moreover, the average objectives of DE-LS in all instances are also better than the other four compared algorithms. In addition, we can find that when the number of jobs is 20 whatever the number of machines is, 3 or 4, the minimum

values of all compared algorithms are the same. Moreover, the average objectives value of the proposed DE-LS is equal to the minimum value found, which means that the proposed algorithm can find the minimum value in all runtimes. Although other compared algorithms can find the minimum value, the robustness of other compared algorithms is worse than the proposed DE-LS. When the number of jobs is 30 and 40, and the number of machines is 4, the minimum objective value of Rand/1 algorithm is the same as that of the proposed algorithm. When the number of jobs is 30 and the number of machines is 4, the current-to-best/1 algorithm has found the minimum objective value as same as that of the proposed algorithm. With the increasing number of jobs, the deviation between the proposed DE-LS and the other four compared algorithms becomes obvious.

The RPD values and SD values of each algorithm for all instances have been provided in Table 4. It is easy to find that RPD and SD values of the proposed DE-LS are smaller than those of the other compared algorithms in all instances. In addition, the RPD values of the proposed DE-LS in all instances are no more than 1, and the SD values of the

TABLE 4: Computational results about RPD and SD.

Instances	Types	DE-LS		Without LS		Best/1		Current-to-best/1		Rand/1	
		RPD	SD	RPD	SD	RPD	SD	RPD	SD	RPD	SD
1	J20-M3	0.00	0.00	0.54	0.77	0.72	1.17	0.45	1.04	0.10	0.21
2	J20-M4	0.00	0.00	1.09	1.13	1.97	1.98	0.63	0.66	0.37	0.44
3	J30-M3	0.12	0.25	2.11	3.27	2.56	4.22	1.52	2.41	1.24	2.31
4	J30-M4	0.03	0.10	0.84	1.24	2.44	3.12	0.99	1.41	0.31	0.57
5	J40-M3	0.16	0.49	1.13	2.50	3.10	6.71	1.79	3.95	1.22	2.63
6	J40-M4	0.17	0.46	1.74	3.07	2.48	4.40	0.75	1.50	0.60	1.23
7	J50-M3	0.26	0.73	0.82	2.38	1.56	4.04	0.88	2.27	0.49	1.32
8	J50-M4	0.43	1.13	1.59	3.53	3.41	7.38	1.05	2.52	0.92	2.26
9	J60-M3	0.35	1.10	1.34	4.11	2.40	7.11	0.93	3.03	1.25	3.94
10	J60-M4	0.46	1.52	1.29	3.71	4.12	10.96	2.82	8.86	2.22	5.88
11	J70-M3	0.97	3.74	2.56	8.58	4.11	13.58	2.26	7.55	3.49	11.32
12	J70-M4	0.31	1.46	2.12	7.37	4.40	15.05	2.95	10.21	2.67	9.24
13	J80-M3	0.75	3.42	2.99	13.08	4.77	20.11	2.78	12.00	4.23	17.87
14	J80-M4	0.99	4.61	2.38	8.76	3.27	12.39	2.79	10.89	3.78	14.46
15	J90-M3	0.31	1.43	0.67	3.15	1.30	5.99	0.65	3.41	1.93	8.10
16	J90-M4	0.64	3.10	2.43	9.43	4.26	15.82	3.13	11.79	5.60	20.73
17	J100-M3	0.55	2.61	1.11	5.04	2.14	9.41	1.37	7.00	3.20	14.33
18	J100-M4	0.74	4.20	2.27	10.89	3.81	18.06	2.42	11.73	3.72	17.64
Average		0.40	1.69	1.61	5.11	2.93	8.97	1.68	5.68	2.07	7.47

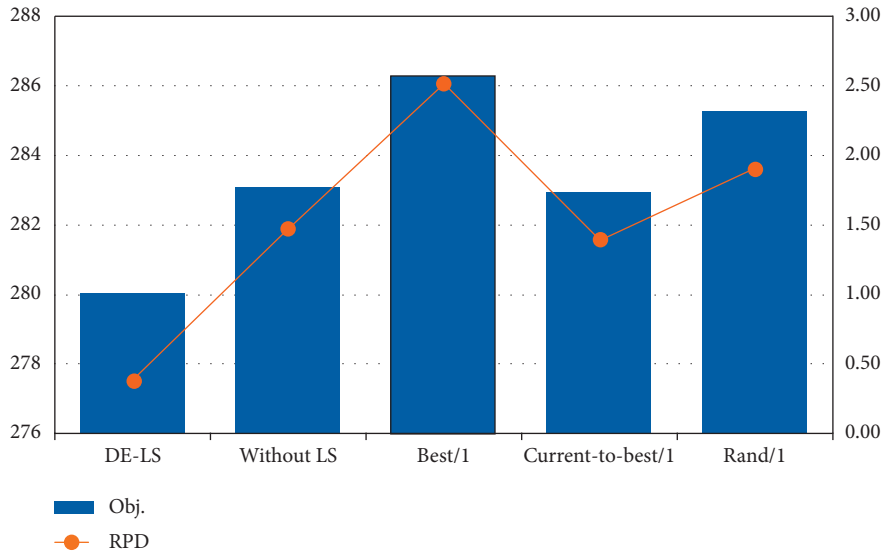


FIGURE 6: Two-dimensional diagram when the number of machines equals 3.

proposed DE-LS are smaller than 5, which indicate that the proposed DE-LS has stronger robustness than the other compared algorithms. It is worth mentioning that when the number of jobs is 20, the RPD and SD values of the proposed DE-LS are equal to zero, which means that DE-LS has found the minimum value in all runtimes in these two instances. When the number of jobs is no more than 60, the RPD values of DE-LS are no more than 0.5 and the SD values of DE-LS are no more than 1.6, which indicates the proposed DE-LS has shown more superiority in the small-scale problem than in the large-scale problem.

Moreover, two-dimensional comparison plots have been provided in Figures 6 and 7 in order to make a deeper analysis

of the experimental results. We select the average objective value and RPD as the variables of the combination-comparison plots. Furthermore, according to the number of machines, we classify all instances into two groups, i.e., the number of machines equal to 3 and the number of machines equal to 4. There is a two-dimensional comparison plot for each group, respectively. The objective value in the diagram for each algorithm is the further average objective, which is calculated by the average objective of the algorithm for each instance in the specific group. Similarly, the RPD value in the diagram for each algorithm is the average RPD, which is calculated by the RPD value of the algorithm for each instance in the specific group. These two combination plots show that

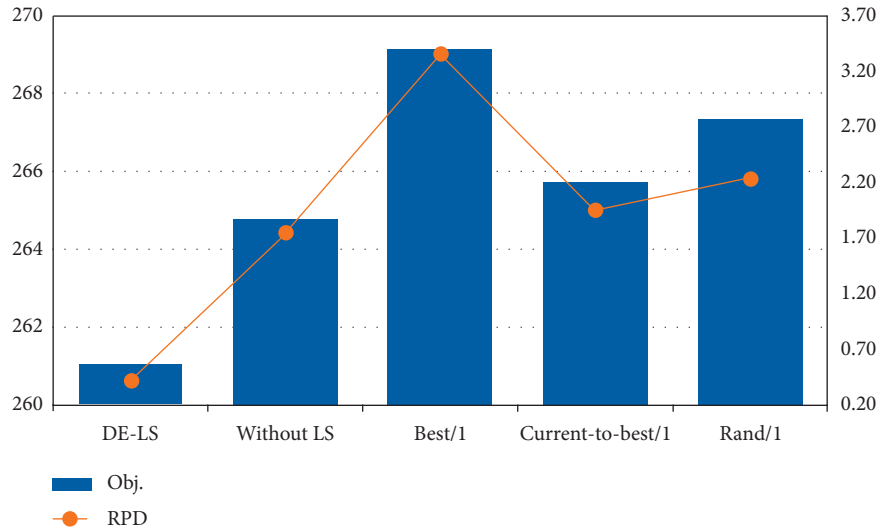


FIGURE 7: Two-dimensional diagram when the number of machines equals 4.

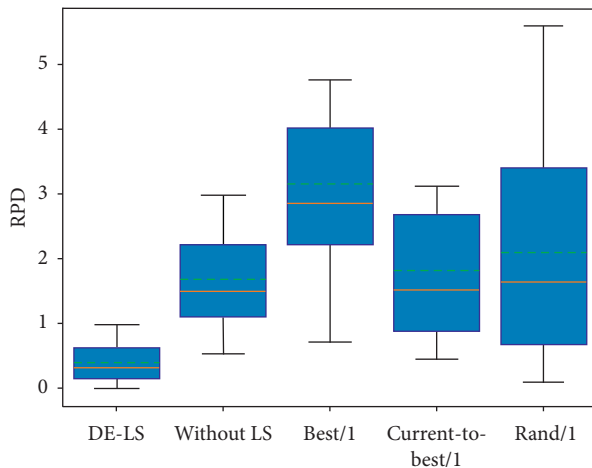


FIGURE 8: Box-plot of RPD for five compared algorithms.

the proposed DE-LS has outperformed other compared algorithms whatever in objective values or RPD values.

In addition, consider the stochastic features of the metaheuristic algorithms; the stability of the proposed algorithm has been further evaluated using the box-plot analysis. The box-plot can describe the discrete distribution of the data in a relatively stable way, which means the plot will not be influenced by the abnormal value. In this part, the box-plot of the RPD value for five algorithms for 16 instances in 10 independent runs is shown in Figure 8. Applying the concept of quartile, there are six elements having been drawn for each box plot, i.e., the up limb, the up quartile, the average value, the median, the low quartile, and the low limb. In the figure, we can find that all elements in the RPD box-plot graphics of the proposed DE-LS are smaller than other compared algorithms, which shows that there is significant difference in RPD values between the proposed DE-LS and other algorithms. Moreover, the gap between the up limb and low limb of the proposed DE-LS is the smallest. On the

contrary, the gap between the up limb and low limb of the rand/1 algorithm is the largest.

5. Conclusions

This paper addresses a hybrid flow-shop scheduling problem with considering different requirements of customers, which reflects the procedures of the productions. A mixed-integer programming model has been formulated for the concerned problem. The objective of the studied problem is to minimize the makespan and cost of delay. Different requirements of customers indicate the diverse productions need to be processed in different procedures and have the specific number of the procedures, which increases the complexity of the studied problem. Moreover, a hybrid metaheuristic algorithm which combined differential evolution and local search named DE-LS has been proposed in the paper to solve the studied problem. The differential evolution and local search are efficient metaheuristic algorithms, which can search for the better solution in a short time. Thus, the paper combines these two algorithms in order to find the optimal solution of the concerned problem in an available time. Moreover, the proposed DE-LS has been compared with four variant algorithms, and the experimental results show that the proposed DE-LS outperforms the compared algorithms. In future study, more practical features in flow-shop scheduling problem will be investigated, such as the delivery requirements of customers, the deteriorating effect of the machines, etc.

Data Availability

The raw/processed data required to reproduce these findings cannot be shared at this time as the data forms part of an ongoing study.

Conflicts of Interest

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

References

- [1] Q.-K. Pan, M. Fatih Tasgetiren, P. N. Suganthan, and T. J. Chua, "A discrete artificial bee colony algorithm for the lot-streaming flow shop scheduling problem," *Information Sciences*, vol. 181, no. 12, pp. 2455–2468, 2011.
- [2] J. Deng, L. Wang, S.-Y. Wang, and X.-L. Zheng, "A competitive memetic algorithm for the distributed two-stage assembly flow-shop scheduling problem," *International Journal of Production Research*, vol. 54, no. 12, pp. 3561–3577, 2016.
- [3] J. Lin, Z.-J. Wang, and X. Li, "A backtracking search hyper-heuristic for the distributed assembly flow-shop scheduling problem," *Swarm and Evolutionary Computation*, vol. 36, pp. 124–135, 2017.
- [4] J. Deng and L. Wang, "A competitive memetic algorithm for multi-objective distributed permutation flow shop scheduling problem," *Swarm and Evolutionary Computation*, vol. 32, pp. 121–131, 2017.
- [5] S. M. Johnson, "Optimal two- and three-stage production schedules with setup times included," *Naval Research Logistics Quarterly*, vol. 1, no. 1, pp. 61–68, 1954.
- [6] V. Riahi and M. Kazemi, "A new hybrid ant colony algorithm for scheduling of no-wait flowshop," *Operational Research*, vol. 18, no. 1, pp. 55–74, 2018.
- [7] M. K. Marichelvam, Ö. Tosun, and M. Geetha, "Hybrid monkey search algorithm for flow shop scheduling problem under makespan and total flow time," *Applied Soft Computing*, vol. 55, pp. 82–92, 2017.
- [8] R. Ruiz and J. A. Vázquez-Rodríguez, "The hybrid flow shop scheduling problem," *European Journal of Operational Research*, vol. 205, no. 1, pp. 1–18, 2010.
- [9] I. Ribas, R. Leisten, and J. M. Framiñan, "Review and classification of hybrid flow shop scheduling problems from a production system and a solutions procedure perspective," *Computers & Operations Research*, vol. 37, no. 8, pp. 1439–1454, 2010.
- [10] W. Qin, J. Zhang, and D. Song, "An improved ant colony algorithm for dynamic hybrid flow shop scheduling with uncertain processing time," *Journal of Intelligent Manufacturing*, vol. 29, no. 4, pp. 891–904, 2018.
- [11] D. Lei and Y. Zheng, "Hybrid flow shop scheduling with assembly operations and key objectives: a novel neighborhood search," *Applied Soft Computing*, vol. 61, pp. 122–128, 2017.
- [12] B. Zhang, Q.-K. Pan, L. Gao, X.-L. Zhang, H.-Y. Sang, and J.-Q. Li, "An effective modified migrating birds optimization for hybrid flowshop scheduling problem with lot streaming," *Applied Soft Computing*, vol. 52, pp. 14–27, 2017.
- [13] C. Yu, Q. Semeraro, and A. Matta, "A genetic algorithm for the hybrid flow shop scheduling with unrelated machines and machine eligibility," *Computers & Operations Research*, vol. 100, pp. 211–229, 2018.
- [14] Q.-K. Pan, L. Gao, X.-Y. Li, and K.-Z. Gao, "Effective metaheuristics for scheduling a hybrid flowshop with sequence-dependent setup times," *Applied Mathematics and Computation*, vol. 303, pp. 89–112, 2017.
- [15] A. Ebrahimi, H. W. Jeon, S. Lee, and C. Wang, "Minimizing total energy cost and tardiness penalty for a scheduling-layout problem in a flexible job shop system: a comparison of four metaheuristic algorithms," *Computers & Industrial Engineering*, vol. 141, Article ID 106295, 2020.
- [16] S. Kavitha and P. Venkumar, "A vibrant crossbreed social spider optimization with genetic algorithm tactic for flexible job shop scheduling problem," *Measurement and Control*, vol. 53, no. 1-2, pp. 93–103, 2020.
- [17] T. R. Chinnusamy, P. Kammar, F. Praveen et al., "Scheduling of flexible manufacturing system by hybridizing petri net with improved scatter search algorithm," in *Recent Trends in Mechanical Engineering*, G. S. V. L. Narasimham, A. V. Babu, S. S. Reddy, and R. Dhanasekaran, Eds., pp. 305–332, Springer Singapore, Singapore, 2020.
- [18] G. Kliewer and S. Tschoke, "A general parallel simulated annealing library and its application in airline industry," in *Proceedings of the 14th International Parallel and Distributed Processing Symposium. IPDPS 2000*, pp. 55–61, Cancun, Mexico, May 2000.
- [19] O. Ezzinbi, M. Sarhani, A. El Afia, and Y. Benadada, "A metaheuristic approach for solving the airline maintenance routing with aircraft on ground problem," in *Proceedings of the 2014 International Conference on Logistics Operations Management*, pp. 48–52, Rabat, Morocco, June 2014.
- [20] E. Teymourian, A. Sadeghi, and F. Taghipourian, "A dynamic virtual hub location problem in airline networks—formulation and metaheuristic solution approaches," in *Proceedings of the First International Technology Management Conference*, pp. 1061–1068, San Jose, CA, USA, June 2011.
- [21] R. Mansi, S. Hanafi, C. Wilbaut, and F. Clautiaux, "Disruptions in the airline industry: math-heuristics for re-assigning aircraft and passengers simultaneously," *European Journal of Industrial Engineering*, vol. 6, no. 6, p. 690, 2012.
- [22] N. R. Das, S. C. Rai, S. C. Rai, and A. Nayak, "Intelligent scheduling of demand side energy usage in smart grid using a metaheuristic approach," *International Journal of Intelligent Systems and Applications*, vol. 10, no. 6, pp. 30–39, 2018.
- [23] M. A. Djema, M. Boudour, K. Agbossou, A. Cardenas, and M. L. Doumbia, "Adaptive direct power control based on ANN-GWO for grid interactive renewable energy systems with an improved synchronization technique," *International Transactions on Electrical Energy Systems*, vol. 29, Article ID e2766, 2019.
- [24] F. Ekinci, T. Demirdelen, I. O. Aksu, K. Aygul, B. Esenboga, and M. Bilgili, "A novel hybrid metaheuristic optimization method to estimate medium-term output power for horizontal axis wind turbine," *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, vol. 233, no. 5, pp. 646–658, 2019.
- [25] D. Dabhi and K. Pandya, "Enhanced velocity differential evolutionary particle swarm optimization for optimal scheduling of a distributed energy resources with uncertain scenarios," *IEEE Access*, vol. 8, pp. 27001–27017, 2020.
- [26] R. Storn and K. Price, "Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.
- [27] J. Brest, S. Greiner, B. Boskovic, M. Mernik, and V. Zumer, "Self-adapting control parameters in differential evolution: a comparative study on numerical benchmark problems," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 6, pp. 646–657, 2006.
- [28] S. Das and P. N. Suganthan, "Differential evolution: a survey of the state-of-the-art," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 1, pp. 4–31, 2011.
- [29] G. Zhang and K. Xing, "Differential evolution metaheuristics for distributed limited-buffer flowshop scheduling with makespan criterion," *Computers & Operations Research*, vol. 108, pp. 33–43, 2019.
- [30] B.-H. Zhou, L.-M. Hu, and Z.-Y. Zhong, "A hybrid differential evolution algorithm with estimation of distribution algorithm

- for reentrant hybrid flow shop scheduling problem,” *Neural Computing and Applications*, vol. 30, no. 1, pp. 193–209, 2018.
- [31] R. Ruiz and T. Stützle, “A simple and effective iterated greedy algorithm for the permutation flowshop scheduling problem,” *European Journal of Operational Research*, vol. 177, no. 3, pp. 2033–2049, 2007.
- [32] J. Xu, C.-C. Wu, Y. Yin, and W.-C. Lin, “An iterated local search for the multi-objective permutation flowshop scheduling problem with sequence-dependent setup times,” *Applied Soft Computing*, vol. 52, pp. 39–47, 2017.
- [33] M. F. Tasgetiren, D. Kizilay, Q.-K. Pan, and P. N. Suganthan, “Iterated greedy algorithms for the blocking flowshop scheduling problem with makespan criterion,” *Computers & Operations Research*, vol. 77, pp. 111–126, 2017.
- [34] B. Yagmahan and M. M. Yenisey, “A multi-objective ant colony system algorithm for flow shop scheduling problem,” *Expert Systems with Applications*, vol. 37, no. 2, pp. 1361–1368, 2010.
- [35] J. Bruno and P. Downey, “Complexity of task sequencing with deadlines, set-up times and changeover costs,” *SIAM Journal on Computing*, vol. 7, no. 4, pp. 393–404, 1978.
- [36] J. N. D. Gupta, “Two-stage, hybrid flowshop scheduling problem,” *Journal of the Operational Research Society*, vol. 39, no. 4, pp. 359–364, 1988.

Research Article

Cost-Benefit Models on Integrating Information Technology Services in Automotive Production Management

Bin Hu, Jianlin Lv , and Kun Yang

School of Management Studies, Shanghai University of Engineering Science, Shanghai 201620, China

Correspondence should be addressed to Jianlin Lv; m030218112@sues.edu.cn

Received 10 May 2020; Revised 26 May 2020; Accepted 5 June 2020; Published 15 July 2020

Academic Editor: Lu Zhen

Copyright © 2020 Bin Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The integration of the new-generation information technology and the automobile manufacturing industry has significantly improved the production efficiency of the automobile manufacturing industry, but it will also increase the technology application cost of the automobile manufacturing industry. The boundary value of the income change of the automobile manufacturing industry can be obtained by examining the influence of new-generation information technology on the price of parts, the price of automobiles, and the quantity of production in the upstream and downstream of the automobile manufacturing industry chain. The study found that the benefit of the automobile manufacturing industry that meets the conditions of technology application costs has increased. The value added to the downstream enterprises in the industrial chain is greater than the value added to the upstream companies. The lower the cost of technology application, the greater the impact on the number of automobile production. In the end, an example is used to verify the reliability of the research results.

1. Introduction

Nowadays, a new round of scientific and technological revolution which is represented by big data analysis, mobile Internet of Things, cloud computing, etc., is sweeping the world. A new manufacturing system of interconnection, capability collaboration, resource sharing, and open cooperation is being built, which greatly expands the space for innovation and development of the manufacturing industry. As an important support of China's industrial economy, the automobile manufacturing industry is characterized by a large number of parts, long industrial chains, and large-scale and complex production organizations. It is closely related to other industries and plays an important role in leading economic growth. The development of the new-generation information technology will drive China's automobile manufacturing industry to a new stage of transformation and upgrading. Under the role of the new-generation information technology, the entire industrial chain, the entire production process, and the entire manufacturing life cycle of the automobile manufacturing industry have been optimized and

upgraded, which will accelerate the pace of China's transformation from a big automobile manufacturing country to a powerful automobile manufacturing country. The upstream of the automobile manufacturing industry chain is automobile parts manufacturing enterprises, and the downstream is automobile manufacturing enterprises. The integrated development of the new-generation information technology and automobile manufacturing industry will optimize the production mode of the automobile manufacturing industry chain, improve the industrial chain interaction mode, and bring new economic growth points [1]. Under the guidance of the "Made in China 2025" action plan, the whole industry chain of automobile manufacturing will be deeply integrated and applied with the new-generation information technology. The automobile parts manufacturing enterprises will achieve the refined production, the automobile manufacturing enterprises will achieve the intelligent manufacturing, and the whole industry chain will achieve the real-time interconnection and intelligent logistics and jointly promote the transformation and upgrading of the automobile manufacturing industry [2].

2. Literature Review

The deep integration of new-generation information technology and manufacturing makes the development of manufacturing and the innovation of science and technology present unprecedented features and trends. Based on the application of new-generation information technology, the manufacturing industry will achieve digital, networked, and intelligent manufacturing, which will have a profound impact on the high-quality development of Chinese manufacturing [3, 4]. Thirupathi and Vinodh [5] took auto parts manufacturing enterprises in southern India as an example, and a structural equation model was established to study the importance of sustainable manufacturing for auto parts manufacturing companies to maintain a competitive advantage and analyze the relationship between the sustainable development factors of auto parts manufacturing companies. Huang and Zhang [6] used the gray correlation analysis and analytic hierarchy process to establish the hierarchical structure model of the multiangle indices system of the supply chain management of the automobile manufacturing industry, proposed that efficient supply chain management and the application of information technology have a significant impact on the economic efficiency of automobile manufacturing companies. Li et al. [7] proposed that the manufacturing industry is moving towards smart manufacturing through new-generation information technology. The deep integration and development of new-generation information technology and manufacturing would lead to the innovation of manufacturing methods, organizational forms, and manufacturing models in the manufacturing industry, thereby promoting the transformation and upgrading of the manufacturing industry. Tao et al. [8] believed that the integration of new-generation information technology and manufacturing can promote manufacturing to achieve manufacturing collaboration, product design personalization, service management customization, and system resource integration, which would help manufacturing achieve smart manufacturing and create higher economic value. Cainelli et al. [9] conducted empirical research on Spanish manufacturing companies and found that manufacturing companies continue to carry out innovative R&D activities with the help of new-generation information technology, which would help manufacturing achieve intelligent manufacturing. From the perspective of the application of information technology, Liu and Gu [10] deeply analyzed the development process of advanced manufacturing and the application of information technology. He found that information technology has an important impact on the manufacturing process and manufacturing technology, and the key to the transformation and upgrading of the manufacturing industry lays in the integrated application of information technology.

In summary, the research on the integration of new-generation information technology and the automobile manufacturing industry is gradually accumulating, and the research of more experts and scholars is mainly distributed in the significance of the integration of new-generation information technology and the manufacturing industry. There are few specific studies on how the automobile

manufacturing industry applies new-generation information technology and the specific impact path of new-generation information technology on the automobile manufacturing industry, and this kind of study precisely determines whether the automobile manufacturing industry integrates and applies new-generation information technology. This article starts from the automobile manufacturing industry chain and analyzes the costs and benefits of the automobile parts manufacturing industry located upstream of the industry chain and the vehicle manufacturing industry located downstream of the industry chain. This paper will focus on comparing the impact on the automobile manufacturing industry before and after the application of new-generation information technology and guide the integration of new-generation information technology and automobile manufacturing.

3. Model Assumptions and Design

The automobile manufacturing industry chain includes upstream automobile component manufacturing and downstream vehicle manufacturing. The automobile component manufacturing industry uses material resources to produce components suitable for automobiles and deliver them to the downstream vehicle manufacturing industry [11]. The vehicle manufacturing industry uses the supplied parts to produce passenger cars or freight cars for the market [12]. Automobile manufacturing companies need to purchase parts from auto parts manufacturing companies at a price b to manufacture q number of cars. Auto parts manufacturing companies determine auto parts shipments and prices based on their manufacturing costs and logistics costs. After receiving the parts and components, the automobile manufacturing enterprise carries out manufacturing. Automobile manufacturers determine the market sales price of automobiles P based on market demand, processing and manufacturing costs, logistics costs, and advertising effectiveness.

3.1. Model Assumptions

3.1.1. Logistics Costs. In the automobile manufacturing industry chain, supposing that the logistics cost of transporting material resources for the production of parts by the auto parts manufacturing company is borne by the auto parts manufacturing company, which is $c_1(t)$. The logistics cost of auto parts manufacturing enterprises to deliver parts to auto manufacturing enterprises is borne by auto manufacturing enterprises, which is $c_2(t)$. The logistics cost function is positively related to transportation time. The longer the transportation time, the higher the logistics cost that needs to be paid [13]. Therefore, assuming that the logistics cost function is a function with transportation time as an independent variable, some of the characteristics of the function are $\partial c_1(t)/\partial t \geq 0$ and $\partial c_2(t)/\partial t \geq 0$.

3.1.2. Manufacturing Costs. The manufacturing cost of the automobile manufacturing industry includes the rental fee of the production workshop, the salary of labor workers, the

depreciation of production equipment and machinery, and other expenses in the manufacturing process [14]. Assuming that the manufacturing cost of the process of producing and processing auto parts for the auto parts manufacturing enterprise is $\alpha_1(t)$, the manufacturing cost of auto manufacturing companies using resources such as parts and components to produce a car is $\alpha_2(t)$. There is a positive correlation between manufacturing costs and time spent. The longer the time required for manufacturing, the higher the salary the enterprise needs to pay for workers, and other costs incurred in manufacturing increase with time, and the higher the manufacturing cost. Therefore, assuming that the manufacturing cost is a function of the required time as an independent variable, some features of the function are $\partial\alpha_1(t)/\partial t \geq 0$ and $\partial\alpha_2(t)/\partial t \geq 0$.

3.1.3. Material Costs and Other Costs. The automobile manufacturing industry requires a certain amount of materials when manufacturing products. Assuming that material costs and other costs are a , they are jointly borne by the automobile parts manufacturing company and the automobile manufacturing company. The decision-makers in the operation and management of the automobile manufacturing industry are completely rational, and the risk appetite is neutral.

3.1.4. Advertising Marketing Effect. To increase the sales volume of automobiles and expand the scale of market demand, automobile manufacturers will actively put advertisements on the market to promote the popularity of automobile brands, so that let potential consumers in the market have a good impression on the brand car, stimulate consumers' desire to buy, and help the brand car open the market [15, 16]. After applying new-generation information technology, Internet propaganda will make the advertising of auto manufacturing companies achieve social marketing effects. Automobile manufacturers use big data analysis technology to build potential customer relationship graphs to fully understand the actual market demand. The cloud computing platform will use targeted data mining to precisely target and deliver targeted advertisements to gain users' goodwill and effectively enhance brand value [17, 18]. Supposing the advertising marketing effect function of automobile manufacturing enterprises is $\varnothing(t)$ ($0 \leq \varnothing(t) \leq Z$), when $\varnothing(t) = 0$, advertisement marketing does not affect; when $\varnothing(t) = Z$, advertising marketing is fully effective. The longer the advertising period, the more likely the potential consumers in the market to see this advertisement, and the greater the number of potential consumers who see this advertisement. Therefore, the advertising marketing effect function is a function with time as its independent variable, and some of the features of the function are $\partial\varnothing(t)/\partial t \geq 0$.

3.1.5. New-Generation Information Technology Application Costs. The integration of new-generation information technology and the automobile manufacturing industry will help to promote the automobile manufacturing industry to complete technological innovation and upgrade, personalized product

design, diversified customer service, and integrated manufacturing. Increased production efficiency will release labor, intelligent logistics will reduce transportation time, and the interconnection of production equipment and machines will realize intelligent manufacturing [19, 20]. The application of new-generation information technology in the automobile manufacturing industry chain has to pay a certain application cost. Assuming that the application cost of new-generation information technology is $C_{IT} = C_t + C_y$, where C_t is fixed and C_y is a variable cost. The application cost of new-generation information technology is jointly borne by the upstream automobile parts manufacturing enterprises and the downstream vehicle manufacturing enterprises. After applying new-generation information technology, the overall manufacturing efficiency of the automobile manufacturing industry has improved, the logistics time has been shortened, and the advertising and marketing effects have become better and better.

3.1.6. Demand Function. The demand of potential consumers in the market for cars is affected by the sales price of cars and the effectiveness of advertising and marketing. With the same price, the better the car advertising effect, the more likely the potential consumers will buy. Based on the research of Abu-Eisheh and Mannering [21] and Yu et al. [22], we can assume that the market's demand function for cars is $n = AP^{-k}\varnothing(t)$, where A is the market size coefficient, k is the price elasticity coefficient, and $k > 1$.

3.1.7. Cost Function and Benefit Function. We set the cost function of the automobile manufacturing industry to be represented by C_y^x and the profit function to be represented by L_y^x . Among them, $x \in \{E, F\}$, where E stands for automobile parts manufacturing enterprises and F stands for automobile manufacturing enterprises; $y \in \{0, 1\}$, where 0 represents that the automobile manufacturing industry has not applied new-generation information technology and 1 represents that the automobile manufacturing industry has applied new-generation information technology.

Assuming that the time spent without applying new-generation information technology is t_1 , and the time spent by the automobile manufacturing industry after the application of new-generation information technology is t_2 , the logistics cost and manufacturing cost of auto parts manufacturing companies and auto manufacturing companies, respectively, meet $0 \leq c_1(t_2) \leq c_1(t_1)$, $0 \leq \alpha_1(t_2) \leq \alpha_1(t_1)$, $0 \leq c_2(t_2) \leq c_2(t_1)$, and $0 \leq \alpha_2(t_2) \leq \alpha_2(t_1)$. The advertising marketing performance function satisfies $0 \leq \varnothing(t_1) \leq \varnothing(t_2) \leq 1$.

3.2. Model Design

3.2.1. Cost-Benefit Model of Automobile Manufacturing. Automobile manufacturers purchase parts from auto parts manufacturers at a price b_0 to produce q_0 quantity of cars. The unified assumption of logistics time and manufacturing time is t_1 . The cost of an automobile manufacturing

company is $C_0^F = [b_0 + \alpha_2(t_1) + c_2(t_1)]q_0 + a$, and the benefit is $L_0^F = [p_0 - b_0 - \alpha_2(t_1) - c_2(t_1)]q_0 - a$. Because the market demand function is $n = AP^{-k}\varnothing(t)$, to pursue maximum profits, automobile manufacturers will set appropriate sales prices based on market demand, and then it can be concluded $P_0 = [A\varnothing(t)/q_0]^{1/k}$. Substituting this formula into the benefit function yields

$$L_0^F = \left\{ \left[\frac{A\phi(t_1)}{q_0} \right]^{1/k} - b_0 - \alpha_2(t_1) - c_2(t_1) \right\} q_0 - a. \quad (1)$$

To maximize the profits of the automobile manufacturing enterprise, there must be a suitable q_0 so that the benefit function reaches the maximum value, and obviously, the second derivative of the benefit function is less than zero. If the first derivative of the benefit function is zero, then we have

$$q_0^F = A\phi(t_1) \left\{ \frac{k-1}{k[b_0 + \alpha_2(t_1) + c_2(t_1)]} \right\}^k. \quad (2)$$

Substituting (2) into the price function gives

$$P_0^F = \frac{k[b_0 + \alpha_2(t_1) + c_2(t_1)]}{k-1}. \quad (3)$$

According to this, we can get the cost and benefit of auto manufacturers when the price of parts is b_0 :

$$C_0^F = A\phi(t_1)[b_0 + \alpha_2(t_1) + c_2(t_1)] \left\{ \frac{k-1}{k[b_0 + \alpha_2(t_1) + c_2(t_1)]} \right\}^k + a, \quad (4)$$

$$L_0^F = A\phi(t_1) \left[\frac{b_0 + \alpha_2(t_1) + c_2(t_1)}{k-1} \right] \left\{ \frac{k-1}{k[b_0 + \alpha_2(t_1) + c_2(t_1)]} \right\}^k - a. \quad (5)$$

The costs and benefits of auto parts manufacturing companies are as follows:

$$C_0^E = [c_1(t_1) + \alpha_1(t_1)]q_0 + a, \quad (6)$$

$$L_0^E = [b_0 - c_1(t_1) - \alpha_1(t_1)]q_0 - a. \quad (7)$$

There is a great value in the benefit of auto parts manufacturing enterprises. Obviously, the second derivative is less than zero, so that the first derivative is zero, and we can get

$$\frac{\partial L_0^E}{\partial b_0} = q_0 + b_0 \frac{\partial q_0}{\partial b_0} - [c_1(t_1) + \alpha_1(t_1)] \frac{\partial q_0}{\partial b_0}. \quad (8)$$

From formula (2), we can get

$$\frac{\partial q_0}{\partial b_0} = q_0 \left[\frac{-k}{b_0 + \alpha_2(t_1) + c_2(t_1)} \right]. \quad (9)$$

Therefore, the optimal parts prices for auto parts manufacturing companies are as follows:

$$b_0 = \frac{k}{k-1} \left[c_1(t_1) + \alpha_1(t_1) + \frac{c_2(t_1) + \alpha_2(t_1)}{k} \right]. \quad (10)$$

It can be seen that the price of auto parts will affect the number of cars produced. The lower the price of parts and components, the more cars are produced, and the higher the price of parts and components, the fewer cars are produced. The price of parts is related to the logistics cost and manufacturing cost of auto parts manufacturing companies and auto manufacturing companies. The higher the logistics cost and manufacturing cost, the higher the price of parts.

Substituting equation (10) into equations (4) and (5) can further obtain the specific expressions of the cost and benefit of automobile manufacturing enterprises:

$$C_0^F = \left(\frac{k-1}{k} \right)^{2k-1} A\phi(t_1) [c_1(t_1) + \alpha_1(t_1) + \alpha_2(t_1) + c_2(t_1)]^{1-k} + a, \quad (11)$$

$$L_0^F = \left(\frac{k-1}{k} \right)^{2k-1} \frac{A\phi(t_1)}{k-1} [c_1(t_1) + \alpha_1(t_1) + \alpha_2(t_1) + c_2(t_1)]^{1-k} - a. \quad (12)$$

Substituting equations (2) and (10) into equations (6) and (7), we can obtain the optimal expressions of the cost function and benefit function of the auto parts manufacturing enterprise:

$$C_0^E = \left(\frac{k-1}{k} \right)^{2k} A\phi(t_1) [c_1(t_1) + \alpha_1(t_1)] [c_1(t_1) + \alpha_1(t_1) + \alpha_2(t_1) + c_2(t_1)]^{-k} + a, \quad (13)$$

$$L_0^E = \left(\frac{k-1}{k} \right)^{2k} \frac{A\phi(t_1)}{k-1} [c_1(t_1) + \alpha_1(t_1) + \alpha_2(t_1) + c_2(t_1)]^{1-k} - a. \quad (14)$$

As can be seen from the above model, the costs and benefits of auto parts manufacturing companies and auto manufacturing companies are closely related to logistics costs, manufacturing costs, and advertising and marketing effects. The number of cars produced by automobile manufacturing companies is affected by the price of parts, logistics costs, and manufacturing costs and is directly

proportional to the effectiveness of advertising and marketing. In the development of the automobile manufacturing industry, to reduce costs and increase profits, the logistics costs and manufacturing costs should be controlled as much as possible to improve the production efficiency of the automobile manufacturing industry.

3.2.2. Cost-Benefit Model of the Integration of New-Generation Information Technology and Automobile Manufacturing. After the application of new-generation information technology in the automotive industry, the automobile manufacturing industry comprehensively integrates big data, cloud computing, Internet of Things, and other new-generation information technologies in the entire process of R&D, design, management, manufacturing, sales, and service. The construction of intelligent manufacturing capacity in the automobile manufacturing industry chain has been strengthened, the interconnected network system has been improved, and intelligent management and control have been achieved [23]. After applying new-generation information technology, market demand information for automotive products will be effectively communicated to auto manufacturing companies. Automobile manufacturers continue to transmit information to upstream of the industry chain. The auto parts manufacturing company accepts and processes the information and sends the feedback information and an appropriate number of auto parts to the auto manufacturing company [24]. Automobile manufacturers produce automotive products that meet market expectations and put them on the market [25]. The information processing efficiency of this process is improved, logistics time is reduced, manufacturing costs are reduced accordingly, and advertising and marketing effects

are also improved. However, this process requires a certain technical cost to achieve. The application cost of new-generation information technology is jointly borne by the upstream and downstream companies in the automobile manufacturing industry chain.

After bearing the application cost of new-generation information technology, the cost function and benefit function of the auto parts manufacturing enterprise are as follows:

$$\begin{aligned} C_1^E &= [c_1(t_2) + \alpha_1(t_2) + C_t + C_\gamma]q_1 + a, \\ L_1^E &= [b_1 - c_1(t_2) - \alpha_1(t_2) - C_t - C_\gamma]q_1 - a. \end{aligned} \quad (15)$$

According to this, the prices of auto parts manufacturers can be obtained as follows:

$$b_1 = \frac{k}{k-1} \left[c_1(t_2) + \alpha_1(t_2) + C_t + C_\gamma + \frac{c_2(t_2) + \alpha_2(t_2)}{k} \right]. \quad (16)$$

It can be seen from the above model, in the case of other factors unchanged, the higher the application cost of new-generation information technology, the higher the price of auto parts. Substituting equation (16) into equation (2) can obtain the optimal number of vehicles produced by automobile manufacturing enterprises:

$$q_1^F = \left(\frac{k-1}{k} \right)^{2k} A\phi(t_2) [c_1(t_2) + \alpha_1(t_2) + C_t + C_\gamma + c_2(t_2) + \alpha_2(t_2)]^{-k}. \quad (17)$$

Substituting equation (16) into equation (3), we can obtain the optimal sales price set by the automobile manufacturing enterprise after the application of new-generation information technology:

$$P_1^F = \left(\frac{k}{k-1} \right)^2 [c_1(t_2) + \alpha_1(t_2) + C_t + C_\gamma + c_2(t_2) + \alpha_2(t_2)]. \quad (18)$$

According to (16)–(18), the cost function and benefit function of the automobile manufacturing enterprise can be obtained as follows:

$$C_1^F = \left(\frac{k-1}{k} \right)^{2k-1} A\phi(t_2) [c_1(t_2) + \alpha_1(t_2) + C_t + C_\gamma + c_2(t_2) + \alpha_2(t_2)]^{1-k} + a, \quad (19)$$

$$L_1^F = \left(\frac{k-1}{k} \right)^{2k-1} \frac{A\phi(t_2)}{k-1} [c_1(t_2) + \alpha_1(t_2) + C_t + C_\gamma + c_2(t_2) + \alpha_2(t_2)]^{1-k} - a. \quad (20)$$

The cost function and benefit function of auto parts manufacturing enterprise are as follows:

$$C_1^E = \left(\frac{k-1}{k} \right)^{2k} A\phi(t_2) [c_1(t_2) + \alpha_1(t_2) + C_t + C_\gamma] [c_1(t_2) + \alpha_1(t_2) + C_t + C_\gamma + c_2(t_2) + \alpha_2(t_2)]^{-k} + a, \quad (21)$$

$$L_1^E = \left(\frac{k-1}{k} \right)^{2k} \frac{A\phi(t_2)}{k-1} [c_1(t_2) + \alpha_1(t_2) + C_t + C_\gamma + c_2(t_2) + \alpha_2(t_2)]^{1-k} - a. \quad (22)$$

4. Results

The cost of the automobile manufacturing industry chain is $C_0 = C_0^E + C_0^F$, and the benefit is $L_0 = L_0^E + L_0^F$. The total cost of the automotive manufacturing industry chain after the integration of new-generation information technology is $C_1 = C_1^E + C_1^F$, and the total benefit is $L_1 = L_1^E + L_1^F$.

Theorem 1. *When $C_{IT} \leq C_{IT}^T$, the automobile manufacturing industry's total benefit increases after the application of new-generation information technology.*

Proof. The decision-making condition for the automobile manufacturing industry to realize the increase in benefit after the application of new-generation information technology is $L_1 \geq L_0$; combining (20), (22), (12), and (14), we can get

$$C_{IT} \leq \left[\frac{\phi(t_2)}{\phi(t_1)} \right]^{k-1} [c_1(t_1) + \alpha_1(t_1) + \alpha_2(t_1) + c_2(t_1)] - [c_1(t_2) + \alpha_1(t_2) + c_2(t_2) + \alpha_2(t_2)]. \quad (23)$$

Set the right side of the above formula to C_{IT}^T , and this value is the boundary of whether the automobile manufacturing industry can improve the profit after the application of new-generation information technology. \square

Theorem 2. *When $C_{IT} \leq C_{IT}^S$, upstream auto parts manufacturing companies in the automotive manufacturing industry chain will lower the price of parts. When $C_{IT} > C_{IT}^S$, auto parts manufacturers will increase the price of parts.*

Proof. The decision-making condition for auto parts manufacturing companies to reduce the price of parts after applying new-generation information technology is $b_1 \leq b_0$, and combining (16) and (10) and $k > 1$, we can get

$$C_{IT} \leq c_1(t_1) + \alpha_1(t_1) - c_1(t_2) - \alpha_1(t_2) + \frac{c_2(t_1) + \alpha_2(t_1) - c_2(t_2) - \alpha_2(t_2)}{k}. \quad (24)$$

Set the right side of the above formula to C_{IT}^S , and this value is the decision-making boundary for auto parts manufacturers to reduce the price of parts after applying new-generation information technology. \square

Theorem 3. *When the application cost of new-generation information technology meets $C_{IT} \leq C_{IT}^R$, automobile manufacturers will lower the market sales price of cars and vice versa will increase the selling price.*

Proof. The integration and development of new-generation information technology and automobile manufacturing companies have made the decision-making condition for automobile manufacturing companies to reduce car sales prices is $P_1 \leq P_0$, and combining (18) and (3), we can get

$$C_{IT} \leq c_1(t_1) + \alpha_1(t_1) - c_1(t_2) - \alpha_1(t_2) + c_2(t_1) + \alpha_2(t_1) - c_2(t_2) - \alpha_2(t_2). \quad (25)$$

Set the right side of the above formula to C_{IT}^R , and then this value is the boundary value of whether automobile manufacturing enterprises will reduce the sales price of automobiles after applying new-generation information technology. \square

Theorem 4. *When the application cost of new-generation information technology meets $C_{IT} \leq C_{IT}^Q$, the number of cars produced by automobile manufacturing companies will increase, and otherwise, it will decrease.*

Proof. Analyzing the optimal number of cars produced by automobile manufacturing enterprises before and after the application of new-generation information technology (2), (17), and $q_1 \geq q_0$, then we can get

$$C_{IT} \leq \left[\frac{\phi(t_2)}{\phi(t_1)} \right]^{1/k} [c_1(t_1) + \alpha_1(t_1) + \alpha_2(t_1) + c_2(t_1)] - [c_1(t_2) + \alpha_1(t_2) + c_2(t_2) + \alpha_2(t_2)]. \quad (26)$$

Set the right side of the above formula to C_{IT}^Q , and this value is the basis for the decision of whether to increase the output of automobile manufacturing enterprises after applying new-generation information technology. \square

Theorem 5. $C_{IT}^S \leq C_{IT}^R \leq C_{IT}^Q \leq C_{IT}^T$

Proof. Because $k > 1$, we can get $C_{IT}^S \leq C_{IT}^R$, $C_{IT}^Q \leq C_{IT}^T$. Because $k > 1$ and $0 \leq \phi(t_1) \leq \phi(t_2) \leq 1$, we can get $C_{IT}^R \leq C_{IT}^Q$, thus $C_{IT}^S \leq C_{IT}^R \leq C_{IT}^Q \leq C_{IT}^T$.

After the application of new-generation information technology in the automotive industry, the greater the improvement in advertising marketing effectiveness, the greater the increase in the benefits of technology upgrades, and the higher the application costs that car manufacturers are willing to bear. The new-generation information technology improves the overall production efficiency of the automotive industry. The lower the logistics cost and manufacturing cost, the greater the technology application cost that the automotive manufacturing industry chain can bear. When the technology application cost is less than C_{IT}^S , both the price of parts and the price of automobiles have been reduced, and both upstream and downstream enterprises in the automobile manufacturing industry chain have benefited. When $C_{IT}^S \leq C_{IT} \leq C_{IT}^R$, auto parts prices rise. In order to develop market demand, automobile manufacturers still sell cars at prices lower than before applying new-generation information technology. When $C_{IT}^R \leq C_{IT} \leq C_{IT}^Q$, both auto parts prices and auto selling prices have increased, and the number of auto production has increased. The automobile manufacturing industry tries to transfer the cost of technology application to the market, let consumers share the cost for it, and make the automobile manufacturing

TABLE 1: Impact of the value of C_{IT} on the automobile manufacturing industry.

		Ranges (C_{IT})				
		$[0, C_{IT}^S]$	$[C_{IT}^S, C_{IT}^R]$	$[C_{IT}^R, C_{IT}^Q]$	$[C_{IT}^Q, C_{IT}^T]$	$[C_{IT}^T, +\infty)$
Parameter changes	Parts price	↓	↑	↑	↑	↑
	Auto price	↓	↓	↑	↑	↑
	Number of auto production	↑	↑	↑	↓	↓
	Benefit of the automobile manufacturing industry	↑	↑	↑	↑	↓

TABLE 2: Changes in auto industry’s decision-making parameters.

New-generation information technology	Parts price	Car price	Number of cars produced	Auto parts manufacturing company benefit	Car manufacturing company earnings	The total benefit of the automobile manufacturing industry chain
Before application	23650	81225	51350	9.27×10^8	1.41×10^9	2.34×10^9
After application	22950	65025	278872	4.00×10^9	6.13×10^9	1.01×10^{10}

industry get the most benefits. When $C_{IT}^Q \leq C_{IT} \leq C_{IT}^T$, as auto parts prices and auto prices increase, the market’s demand for autos declines, but the benefit can still be improved. When $C_{IT} > C_{IT}^T$, the income of the automobile manufacturing industry chain has declined, and it is not recommended to apply new-generation information technology. The relationship between the application cost of new-generation information technology and the benefit of the automobile manufacturing industry is shown in Table 1. □

5. Example Analysis

Drawing on the relevant research of experts and scholars [26, 27], according to the data of Chinese automobile manufacturing enterprises, the example parameters are used to verify the model constructed in this paper.

Assuming the scale of demand in the automobile market is $A = 30000000$, demand elasticity is $k = 3$, the logistics cost of the auto parts manufacturing company is $c_1(t_1) = 600$, the manufacturing cost is $\alpha_1(t_1) = 5000$; the logistics cost of the automobile manufacturing company is $c_2(t_1) = 500$, the manufacturing cost is $\alpha_2(t_1) = 30000$, and the advertising marketing effect is $\varnothing(t_1) = A^{1.6}$. In order to promote the technological innovation and upgrade of the automobile manufacturing industry to achieve intelligent manufacturing, it is necessary to integrate and apply new-generation information technology to accelerate its transformation and upgrading process. After the application of new-generation information technology in the automotive industry, the logistics cost of auto parts manufacturing company is $c_1(t_2) = 500$, and the manufacturing cost is $\alpha_1(t_2) = 3500$; the logistics cost of the automobile manufacturing company is $c_2(t_2) = 400$, the manufacturing cost is $\alpha_2(t_2) = 20000$, the advertising marketing effect is $\varnothing(t_2) = A^{1.66}$, and the application cost of new-generation information technology is $C_{IT} = 4500$. Because material costs and other costs are less affected by new-generation information technology, they will not be included in the example analysis. Substituting the above parameters into the model of this paper, the relationship of the obtained parameters is shown in Table 2.

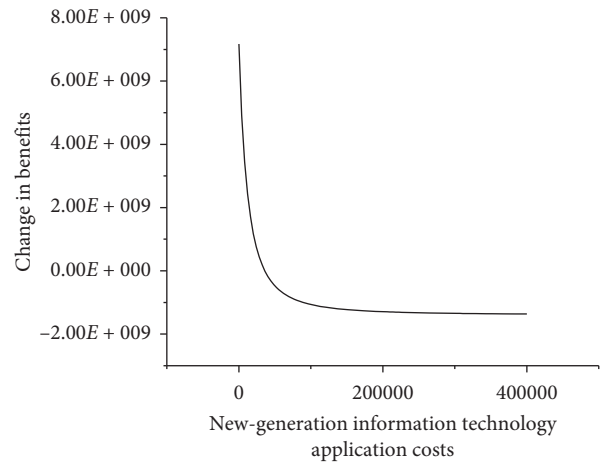


FIGURE 1: Changes in the benefit of the automotive industry.

From Table 1, we can find that when the cost of applying new-generation information technology in the automotive industry is $C_{IT} = 4500$, auto parts prices fall by 2.96%, car prices fall by 19.94%, the number of automobile production increases by more than 4 times, and the benefit of auto parts manufacturing companies and auto manufacturing companies have increased by more than 3 times. According to the model, we can also calculate $C_{IT}^S = 4966.67$, $C_{IT}^R = 11700$, $C_{IT}^Q = 26414.36$, and $C_{IT}^T = 259974.14$, and these satisfy the boundary relationships in Theorem 5.

In order to further verify the model constructed in this paper, we can observe the trend of changes in the benefit of the automobile manufacturing industry as the application cost of new-generation information technology changes while other parameters remain unchanged. The specific changes are shown in Figure 1. When the application cost of new-generation information technology is less than C_{IT}^T , the integration of new-generation information technology and the automobile manufacturing industry will increase economic benefits and positive changes in benefit. When the technology application cost is greater than the boundary value C_{IT}^T , the high cost of technology application undertaken by the automobile manufacturing industry will result

in a decline in benefit. At this time, the benefit change is negative, and the automobile manufacturing industry will abandon the application of new-generation information technology.

6. Conclusion

By building a cost-benefit model before and after the integration of new-generation information technology and the automobile industry, this article focuses on the analysis of the impact of new-generation information technology on upstream and downstream enterprises in the automotive manufacturing industry chain and compares the changes in logistics costs, manufacturing costs, and benefits of the automobile manufacturing industry before and after the integration of new-generation information technology. This paper analyzes the changes in the benefit of the automobile manufacturing industry based on the upper limit of the application cost of new-generation information technology (C_{IT}^T) and obtains the basis for the fusion development of new-generation information technology and automobile manufacturing:

- (1) When the application cost of new-generation information technology is less than the boundary value (C_{IT}^T), the benefit of the entire industry chain of automobile manufacturing will be improved.
- (2) After applying new-generation information technology, the value added to the downstream automobile manufacturing enterprises in the automobile manufacturing industry chain is greater than the value added to the upstream automobile component manufacturing enterprises. When the technology application cost is less than C_{IT}^S , both parts prices and car prices will decrease. When technology application costs are met $C_{IT} \in [C_{IT}^S, C_{IT}^R]$, the price of auto parts will increase, but the selling price of automobiles will still decrease. Auto parts manufacturing enterprises transfer part of the technology application costs to auto manufacturing enterprises.
- (3) The lower the application cost of the new-generation information technology, the impact on the number of automobile production will be greater. When the technology application cost is less than C_{IT}^Q , automobile production is increasing. In the process of integration and development of new-generation information technology and the automobile manufacturing industry, it is necessary to reasonably control the cost of technology application and put more automobile products on the market to obtain greater benefits.

The integration and development of new-generation information technology and the automobile manufacturing industry will transform the automobile manufacturing industry chain into an intelligent environment, enabling interconnection and real-time interaction between upstream and downstream enterprises in the industry chain. Under the effect of digital end-to-end integration technology, the

automobile manufacturing industry realizes digital, networked, and intelligent production. People, machinery equipment, and resources collaborate efficiently in smart factories to promote the automobile manufacturing industry to complete the upgrade of smart manufacturing technology.

In future research, we will consider adding car pricing contracts and technology upgrade paths. The upstream and downstream enterprises of the automobile manufacturing industry chain and the consumers in the market form a three-level supply chain to analyze the changes in the main parameters of the fusion development of new-generation information technology and the automobile manufacturing industry and infer a more realistic decision basis.

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.

Authors' Contributions

All authors have read and agreed to the published version of the manuscript.

Acknowledgments

The authors would like to thank the Research Team of Deep Integration of New-generation Information Technology and Manufacturing from the School of Management Studies in Shanghai University of Engineering Science for their discussion and technical support. This research was supported by the Ministry of Education Humanities and Social Sciences Planning Foundation of China (Grant no. 19YJA790028) and Key Projects in Soft Science Research Field of the "Science and Technology Innovation Action Plan" of Shanghai in 2019 (Grant no. 19692100800).


References

- [1] D. Marc and Y. Masaru, "The emergence of hybrid electric cars: innovation path creation through co-evolution of supply and demand," *Technological Forecasting and Social Change*, vol. 77, no. 8, pp. 1371–1390, 2010.
- [2] S. Sahoo, "Assessing lean implementation and benefits within Indian automotive component manufacturing SMEs," *Benchmarking: An International Journal*, vol. 27, no. 3, pp. 1042–1084, 2020.
- [3] D. Yu, "New trends of global manufacturing development in the era of new industrial revolution and its impact on China," *Tianjin Social Sciences*, vol. 2, pp. 90–102, 2019, in Chinese.
- [4] M. Ghobakhloo, "Determinants of information and digital technology implementation for smart manufacturing," *International Journal of Production Research*, vol. 58, no. 8, pp. 2384–2405, 2020.
- [5] R. M. Thirupathi and S. Vinodh, "Application of interpretive structural modelling and structural equation modelling for analysis of sustainable manufacturing factors in Indian

- automotive component sector,” *International Journal of Production Research*, vol. 54, no. 22, pp. 6661–6682, 2016.
- [6] S. Huang and H. Zhang, “Green supply chain management of automotive manufacturing industry considering multi-perspective indices,” *IEEE Transactions on Electrical and Electronic Engineering*, vol. 14, no. 12, pp. 1787–1795, 2019.
- [7] H. Li, Y. Tian, and W. Li, “Internet thinking and traditional enterprise reengineering,” *China Industrial Economics*, vol. 10, pp. 135–146, 2014, in Chinese.
- [8] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui, “Digital twin-driven product design, manufacturing and service with big data,” *The International Journal of Advanced Manufacturing Technology*, vol. 94, no. 9–12, pp. 3563–3576, 2018.
- [9] G. Cainelli, V. De Marchi, and R. Grandinetti, “Does the development of environmental innovation require different resources? Evidence from Spanish manufacturing firms,” *Journal of Cleaner Production*, vol. 94, pp. 211–220, 2015.
- [10] M. Liu and Q. Gu, “The innovation development of advanced manufacturing industry from the perspective of supply-side reform—comparison of major economies in the world and its enlightenment to China,” *Comparative Economic & Social Systems*, vol. 1, pp. 19–29, 2016, in Chinese.
- [11] T. Laosirihongthong, K.-C. Tan, and D. Adebajo, “Supply chain management in ASEAN automotive manufacturing industry,” *International Journal of Logistics Research and Applications*, vol. 14, no. 5, pp. 317–333, 2011.
- [12] J. Chu, T. Zhang, P. Cui et al., “Analysis of the development patterns of Chinese automobile remanufacturing industry,” *Forum on Science and Technology in China*, vol. 165, no. 1, pp. 33–78, 2010, in Chinese.
- [13] X. Ji, W. Zhou, and T. Ma, “Innovation of logistics cost management and competitiveness of manufacturing enterprises in the Yangtze River delta—based on comparison of manufacturing efficiency between China and America,” *Modern Economic Research*, vol. 440, no. 8, pp. 91–97, 2018, in Chinese.
- [14] J. Zhang and D. Yu, “The cost of China’s manufacturing industry: evolution, characteristics and future trend,” *Henan Social Sciences*, vol. 26, no. 3, pp. 57–62, 2018, in Chinese.
- [15] L. Qian, C. Liu, and H. Ding, “Corporate philanthropy, marketing activities and firm performance,” *Soft Science*, vol. 29, no. 8, pp. 97–100, 2015, in Chinese.
- [16] N. Singh and D. Raina, “Impact of advertisement expenses on net sales for the selected manufacturing companies,” *Global Journal of Enterprise Information System*, vol. 7, no. 3, pp. 38–48, 2015.
- [17] R. L. Gruner, A. Vomberg, C. Homburg, and B. A. Lukas, “Supporting new product launches with social media communication and online advertising: sales volume and profit implications,” *Journal of Product Innovation Management*, vol. 36, no. 2, pp. 172–195, 2019.
- [18] D. Z. Mohammad and K. Shahram, “Predicting user click behaviour in search engine advertisements,” *New Review of Hypermedia & Multimedia*, vol. 21, no. 3, pp. 301–319, 2015.
- [19] Y. Hu, X. Zhou, and C. Li, “Internet-based intelligent service-oriented system architecture for collaborative product development,” *International Journal of Computer Integrated Manufacturing*, vol. 23, no. 2, pp. 113–125, 2010.
- [20] Y. Gao and Y. Song, “Research on the interactive relationship between information communication technology and manufacturing industry,” *Cluster Computing*, vol. 22, no. 3, pp. 5719–5729, 2019.
- [21] S. A. Abu-Eisheh and F. L. Mannering, “Forecasting automobile demand for economies in transition: a dynamic simultaneous-equation system approach,” *Transportation Planning and Technology*, vol. 25, no. 4, pp. 311–331, 2002.
- [22] S. Yu, S. Lu, and R. Shen, “Simulation of demand and effects of merger and reorganization in China’s passenger car market,” *Systems Engineering*, vol. 36, no. 12, pp. 32–39, 2018, in Chinese.
- [23] J. Chen, K. Zhang, Y. Zhou et al., “Exploring the development of research, technology and business of machine tool domain in new-generation information technology environment based on machine learning,” *Sustainability*, vol. 11, no. 12, p. 3316, 2019.
- [24] M. Khouja and R. L. Kumar, “Information technology investments and volume-flexibility in production systems,” *International Journal of Production Research*, vol. 40, no. 1, pp. 205–221, 2002.
- [25] B. Zhang and H. Li, “Promoting the convergence of internet and manufacturing industry—based on the path and mechanism of innovation by “internet plus”,” *Research on Economics and Management*, vol. 38, no. 2, pp. 87–96, 2017, in Chinese.
- [26] S. H. Hymans, G. Ackley, and F. T. Juster, “Consumer durable spending: explanation and prediction,” *Brookings Papers on Economic Activity*, vol. 1970, no. 2, pp. 173–199, 1970.
- [27] D. Chen and W. Liu, “Analysis of demand and its elasticity and forecasting about Chinese automobile market,” *Journal of Chongqing University (Natural Science Edition)*, vol. 12, pp. 138–142, 2005, in Chinese.

Research Article

Intelligent Differential Evolution Scheme for Network Resources in IoT

Huu Dang Quoc ¹, **Loc Nguyen The**,² **Cuong Nguyen Doan**,³ **Toan Phan Thanh**,² and **Neal N. Xiong**⁴

¹*Thuong Mai University, Ha Noi, Vietnam*

²*Ha Noi National University of Education, Ha Noi, Vietnam*

³*Military Institute of Science and Technology, Ha Noi, Vietnam*

⁴*Northeastern State University, Tahlequah, OK, USA*

Correspondence should be addressed to Huu Dang Quoc; huudq@tmu.edu.vn

Received 5 June 2020; Accepted 19 June 2020; Published 14 July 2020

Academic Editor: Tingsong Wang

Copyright © 2020 Huu Dang Quoc et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Scheduling is a fundamental factor in managing the network resource of the Internet of things. For IoT systems such as production lines, to operate effectively, it is necessary to find an intelligent management scheme, i.e., schedule, for network resources. In this study, we focus on multiskill resource-constrained project scheduling problem (MS-RCPSP), a combinatorial optimization problem that has received extensive attention from the research community due to its advantages in network resource management. In recent years, many approaches have been utilized for solving this problem such as genetic algorithm and ant colony optimization. Although these approaches introduce various optimization techniques, premature convergence issue also occurs. Moreover, previous studies have only been verified on simulation data but not on real data of factories. This paper formulated the MS-RCPSP and then proposed a novel algorithm called DEM to solve it. The proposed algorithm was developed based on the differential evolution metaheuristic. Besides, we build the reallocate function to improve the solution quality so that the proposed algorithm converges rapidly to global extremum while also avoiding getting trapped in a local extremum. To evaluate the performance of the proposed algorithm, we conduct the experiment on iMOPSE, the simulated dataset used by previous authors in their research studies. In addition, DEM was verified on real dataset supported by a famous textile industry factory. Experimental results on both simulated data and real data show that the proposed algorithm not only finds a better schedule compared with related works but also can reduce the processing time of the production line currently used at the textile industry factory.

1. Introduction

Internet of things systems include not only sensors and automated machines like robots but also intelligent control and management algorithms [1]. Intelligent algorithms for managing network resource play a key role in controlling the operation of automated machines such as robots, enabling them to devote their full capacity to serve users.

Multiskill resource-constrained project scheduling problem (MS-RCPSP), a combinatorial optimization problem, has received extensive attention from the research community in

recent years due to its advantages in network resources management and scheduling for Internet of things system. The main purpose of the MS-RCPSP is to find the way for assigning tasks to resources so that the execution time is minimal.

In recent years, MS-RCPSP is widely used, and this problem is found not only in critical civilian fields but also in the military setting. MS-RCPSP appears on a broad scale of military operations such as planning an onboard mission or routing an unmanned aerial vehicle (UAV) [2, 3].

The rest of the paper is organized as follows. In Section 2, we present some related works such as previous approaches

to solving the MS-RCPSP. Specifically, the summary of the differential evolution method is included in this section. Differential evolution is a well-known evolutionary algorithm that has been applied effectively to solve NP-hard problems. In this paper, the proposed algorithm is inspired by the differential evolution method. In Section 3, we present the system model and the statement of MS-RCPSP. The formulation of MS-RCPSP is introduced in this section. Section 4 describes the proposed algorithm, which is built on differential evolution strategy [4]. In this section, we focus on presenting the reallocate function, which is the key component that creates the strength of the proposed algorithm. Besides, this section introduces our design for schedule representation and a new measurement model for measuring the difference between two schedules. The proposed algorithm (DEM) is also introduced in this section. Simulation results examining the proposed algorithms along with computational and comparative results are given and analyzed in Section 5. Simulations conducted by using the iMOPSE dataset [5] are presented in this section. In addition, this section describes experiments conducted on the factory's dataset collected from TNG, a well-known national textile company. Finally, Section 6 concludes the paper and outlines some directions for future work.

2. Related Works

2.1. Approximation Algorithms for MS-RCPSP. Scheduling problem arises in many practical situations [6–11]. In the general case, the scheduling is proved to be NP-hard. The authors of [5, 12, 13] show that MS-RCPSP, the problem we mentioned in this paper, is an NP-hard scheduling problem. In addition, MS-RCPSP is multimodal or discontinuous and very challenging to solve with traditional optimization methods. In the past few years, various approaches have been proposed to solve MS-RCPSP, thereby finding out intelligent algorithms to allocate and manage network resources in the Internet of things systems [14–17]. Some popular metaheuristic algorithms are the genetic algorithm (GA), ant colony optimization (ACO), and particle swarm optimization [18–22].

Su et al. [23] used a mixed-integer model and discrete constraints to solve the problems. Maghsoudlou et al. [24] and Bibiks et al. [25] applied the cuckoo Search algorithm to plan for the multirisk project with three distinct evaluation objectives. Lin et al. [26] proposed a new solution based on GA and some other heuristic algorithms.

Myszkowski et al. [5, 13] proposed the hybrid algorithm which is the combination of the differential evolution strategy and the greedy approach for arranging the human resources in product manufacturing projects and calculating time-of-use electricity tariffs and shift differential payments. The authors aimed to achieve the shortest makespan and low cost and gave a new benchmark dataset called iMOPSE [5], an artificially created dataset based on real-world instances for the MS-RCPSP.

The MS-RCPSP is studied and applied in the planning of many practical areas of the Internet of things. Some authors also study new variants of the MS-RCPSP and apply them in

different fields. Mejia [3] developed a new variant of the MS-RCPSP to coordinate research activities in the nuclear laboratory. Hosseinian [27] proposed two new algorithms, P-GWO and MOFA, to solve the MS-RCPSP with the deterioration effect and financial constraints (a case study of a gas treating company). In another study, Hosseinian et al. introduced a new mixed-integer formulation for the time-dependent MS-RCPSP considering the learning effect [28].

Nemati-Lafmejani et al. [29] developed an integrated biobjective optimization model to deal with multimode RCPSP. The objective of the proposed model is to minimize both the costs and the makespan of the project. Tahrir and Yang [30] proposed a hybrid metaheuristic algorithm that combined ant colony optimization and variable neighbourhood search approaches for job scheduling in grid computing.

Several prior research has also been conducted to solve other subproblems of RCPSP, the general problem of MS-RCPSP, to improve the efficiency of network resource management in specific fields of the Internet of things. Ballestin and Leus [31] and Javanmard [32] and their colleagues studied specific cases of RCPSPs such as multiskill stochastic and preemptive problem to calculate project implementation time, i.e., makespan, and build mathematical models for the project resource investment.

2.2. Differential Evolution Method. Differential evolution (DE) is a population-based stochastic optimization algorithm introduced by Storn and Price [4]. The advantages of DE lies in its simplicity, efficiency, and speed thanks to the local search method. DE deals with the old generation of the original population by using the mutation operator to create better solutions in the new generation. Until now, DE is clearly recognized to be the approach that has the potential to solve NP-hard [4, 33] problems.

Tanabe et al. and Ghosh and Das [34, 35] proposed a DE-based method for scheduling problems in grid computing environment whose purpose is to minimize the execution time and add more parameter adjustment to get high effects.

Like other evolutionary approaches, DE performs the evolution of a population by using three kinds of the operator: crossover (recombination), mutation, and selection. The major difference between DE and genetic algorithm, a classical evolutionary algorithm, is due to the mutation operator. DE's mutation operator uses orientation information to change solutions of the current population as follows.

Given the population that is composed of solutions, each of them consists of D components, and thus a particular solution is represented by a vector $x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,D})$ where $x_{i,j} \in R$, $i = 1, 2, \dots, N$; $j = 1, 2, \dots, D$.

To create the mutated solution v_i of a given solution x_i , DE picks out three random different solutions from the current population: $x_{r_1} \neq x_{r_2} \neq x_{r_3} \neq x_i$, and then the mutation solution v_i is determined as follows:

$$v_i = x_{r_1} + F(x_{r_2} - x_{r_3}), \quad (1)$$

where the value of the constant F , which plays the role of the mutant factor, is in the range $[0, 1]$. Because the mutant factor F is used to adjust the size of the directional vector, it is also called the directional hop length.

After executing the mutation step, the crossover operation is carried out by combining the parent solution x_i and the mutation solution v_i . The crossover operator is performed by selecting a random number CR ($CR \in [0, 1]$) as the probability of crossover. The result of the crossover step is represented by the vector $(u_{1,1}, u_{1,2}, \dots, u_{N,D})$ where

$$u_{i,j} = \begin{cases} v_{i,j} & \text{if } \text{rand}_{i,j} \leq CR \text{ or } i = I_{\text{rand}} \\ x_{i,j} & \text{if } \text{rand}_{i,j} \geq CR \text{ or } i \neq I_{\text{rand}} \end{cases} \quad (2)$$

- (i) $i = 1, 2, \dots, N; j = 1, 2, \dots, D$.
- (ii) $\text{rand}_{i,j}$: mutant factor.
- (iii) I_{rand} : a random number varies within the range $[1, D]$. Thanks to I_{rand} , the mutation solution v_i is always different from the parent solution x_i .

Finally, the next generation is created from the x_i and the u_i as follows:

$$x_i(t+1) = \begin{cases} u_i(t), & \text{if } f(u_i(t)) < f(x_i(t)), \\ x_i(t), & \text{in other cases.} \end{cases} \quad (3)$$

Since 1997, various versions of DE have been developed by researchers. This paper proposes a variant of the DE algorithm which uses the following terminologies and conventions:

- (i) A solution represents a schedule, that is, a plan for performing given tasks
- (ii) The number of search space's dimensions is equal to the number of tasks

3. System Model and Definitions

3.1. RCPSP. MS-RCPSP is just a special case of RCPSP. Before the outline of the multitask RCPSP, the description of the classical RCPSP would be introduced as follows.

- (i) Assume a given project represented by a directed noncyclical graph $G(V, E)$, where each node depicts a task and the arc represents finish-to-start precedence relationship between two tasks. The arc $(u, v) \in E$ shows that the task u must be completed before task v begins (Figure 1).
- (ii) Without any difference, two empty tasks were added to the project. The first one is placed at the beginning of the project as the predecessor of every other task, whereas the second empty task is placed at the end of the process as the successor of other tasks.
- (iii) The duration (also called execution time) of a certain task is calculated from the beginning time to the ending time of that task.

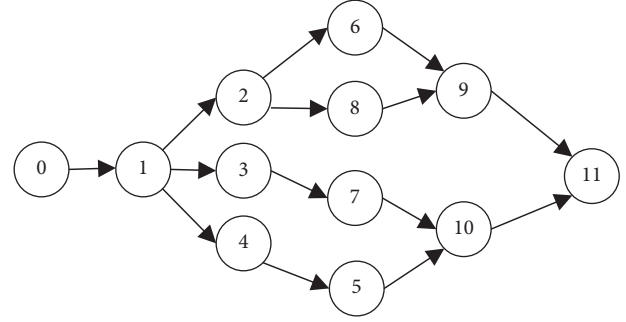


FIGURE 1: The relationship between tasks.

- (iv) The execution progress of every task cannot be stopped until it is done.
- (v) Each task needs an amount of resources for its execution.

The goal of the RCPSP is to find out a schedule of the task such that the precedence and resource constraints are satisfied and the makespan is minimized. RCPSP has been proved to be NP-hard [12]. Throughout the paper, we use the term “makespan” to refer to the execution time of the project, also called project duration. In the process of looking for the optimal schedule, both priority constraint between operations and resource constraint have to be met.

3.2. MS-RCPSP Formulation. In this paper, MS-RCPSP [5, 13, 36] was chosen because it is more practical than its general problem, RCPSP. MS-RCPSP uses the additional attribute, skill, in the problem formulation. Each task requires particular skills at the given levels to be executed, whereas each resource has some skills with given levels. Thus, the resource A is capable of performing the task B if A has skills required by B at the identical or higher level.

Before deepening the study of MS-RCPSP, we take a look at some notations as follows:

- (i) d_j : the duration of task j ; P_j : subset of the predecessors of task j
- (ii) q_j : subset of skills required by task j
- (iii) Q_k : the subset of skills owned by the resource k
- (iv) Q : the set of skills, $Q_k \subseteq Q$
- (v) RS : the set of resources
- (vi) T : the set of tasks
- (vii) T_k : the set of tasks that resource k can perform, $T_k \subseteq T$
- (viii) RS_k : the set of resources that can perform task k , $RS_k \subseteq RS$
- (ix) S_j, F_j : start time and finish time of task j
- (x) $U_{j,k}^t$: $U_{j,k}^t = 1$ in case of resource k is assigned to task j at given time t ; $U_{j,k}^t = 0$ in other cases
- (xi) l_q : the level of the given skill q ; h_q : type of the skill q
- (xii) q_j : the skill required by task j
- (xiii) τ : duration of a schedule

- (xiv) PS: a feasible schedule
- (xv) PS_{all} : the set of all feasible schedules
- (xvi) $f(PS)$: makespan of schedule PS

Each schedule (PS) is responsible for coordinating n tasks and m resources.

Formally, MS-RCPSP could be stated as follows:

$$f(PS) \longrightarrow \min, \quad (4)$$

where

$$f(PS) = \max_{F_i \in T} \{F_i\} - \min_{F_k \in T} \{F_k\}, \quad (5)$$

$$Q_k \neq \emptyset \forall k \in RS, \quad (6)$$

$$d_j \geq 0 \forall j \in T, \quad (7)$$

$$F_j \geq 0 \forall j \in T, \quad (8)$$

$$F_i \leq F_j \quad \forall j \in T, j \neq 1, i \in P_j, \quad (9)$$

$$\forall i \in T_k \exists q \in Q^k : h_q = h_{q_i} \text{ and } l_q \geq l_{q_i}, \quad (10)$$

$$\forall k \in RS, \quad \forall t \in \tau: \sum_{i=1}^n U_{i,k}^t \leq 1, \quad (11)$$

$$\forall j \in T \exists! t \in \tau, \quad !k \in RS: U_{j,k}^t = 1, \text{ where } U_{j,k}^t \in \{0; 1\}. \quad (12)$$

4. Proposed Algorithm

4.1. Schedule Representation. In this study, every schedule or solution is represented by a vector that consists of elements; the number of elements (size of the vector) and the number of tasks are the same. The value of a certain element depicts the resource that can execute the corresponding task.

Example 1. Given a set of task $T = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ and a set of resource $RS = \{1, 2, 3\}$. The MS-RCPSP's goal is to find out a schedule in which the makespan can be minimized while meeting priority constraints between tasks (Figure 1).

Consider a set of task $T = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ and set of resource $RS = \{1, 2, 3\}$. Priority constraints between tasks are depicted in Figure 1, whereby task 1 is the predecessor of task 2, and thus task 2 cannot be performed until task 1 is finished: $S_2 \geq F_1$. There is no relation between task 3 and task 4; therefore, they are executed concurrently or sequentially.

Besides, in this example, we assume that

- (i) $s_k = s_i \forall k, i \in RS$
- (ii) $Q_k = Q \forall k \in RS$
- (iii) $RS_k = RS \forall k \in RS$

The duration of the tasks is given in Table 1.

TABLE 1: Task duration.

Task	1	2	3	4	5	6	7	8	9	10
Duration	2	4	3	5	2	2	5	3	4	2

A feasible schedule for the project is demonstrated in Figure 2, where it can be seen that the makespan is 15.

The above schedule is demonstrated in Table 2, where resource 1 performed tasks number 1, 2, 6, 8, and 9; resource 2 is assigned to execute tasks number 3 and 7; resource 3 deals with the remaining tasks.

4.2. Measurement Model. The differential evolution algorithm was primarily designed for real-valued data. However, like most of the other scheduling problems, MS-RCPSP deals with discrete values, and thus there is a need for the discrete variant of DE. In addition, this discrete DE algorithm needs a model for representing the MS-RCPSP's factors such as the difference between two schedules. In this paper, we build a new measurement model for measuring the difference between the two schedules as follows:

- (i) Unit vector $P = (p_1, p_2, \dots, p_n)$; $p_i = 100/(k_i - 1)$, where k_i is the number of resources that can execute task i
- (ii) The difference between schedule $X = (x_1, x_2, \dots, x_n)$ and schedule $Y = (y_1, y_2, \dots, y_n)$ is denoted by the differential vector $D = X - Y = (d_1, d_2, \dots, d_n)$

where

$$d_i = p_i \times (\text{order}(x_i) - \text{order}(y_i)), \quad (13)$$

$\text{order}(x_i)$: the position of the resource (x_i) in the RS_i

- (iii) The sum of the schedule Y and differential vector D is the schedule $X = (x_1, x_2, \dots, x_n)$

where

$$x_i = \text{position}(\text{round}(y_i + d_i)), \quad (14)$$

$\text{position}(i)$: the resource corresponding to the position i

Example 2. Assume that there are 6 resources that could handle task 1: $RS_1 = \{R_1, R_3, R_4, R_5, R_9, R_{10}\}$. Thus, $k_1 = 6$; $p_1 = 100/(6-1) = 20$. Assume that there are 5 resources that could handle task 2: $RS_2 = \{R_3, R_5, R_7, R_8, R_{10}\}$. Thus, $k_2 = 5$; $p_2 = 100/(5-1) = 25$. The resource order in measurement model is presented in Table 3.

Consider 2 schedules: $X = (3, 5)$; $Y = (5, 10)$. The task assignment is shown in Table 4.

$D = X - Y = (d_1, d_2)$, where $d_1 = p_1$. Abs ($\text{order}(x_1) - \text{order}(y_1)$) = 20. Abs ($1-3$) = 40.

$d_2 = p_2$. Abs ($\text{order}(x_2) - \text{order}(y_2)$) = 25. Abs ($1-4$) = 75. Thus, $D = (40, 75)$. Consider $D' = (35.71, 5.23)$. As shown in Table 5, we have $Z = X + D' = (5, 8)$.

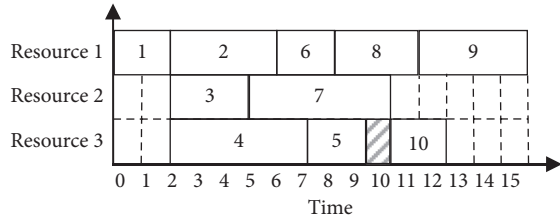


FIGURE 2: A feasible schedule.

TABLE 2: Resource-task assignment.

Task	1	2	3	4	5	6	7	8	9	10
Resource	1	1	2	3	3	1	2	1	1	3

TABLE 3: Resource order in measurement model.

Order	0	1	2	3	4	5
Resource	R_1	R_3	R_4	R_5	R_9	R_{10}
Resource	R_3	R_5	R_7	R_8	R_{10}	

TABLE 4: Schedule’s task-resource assignment.

Schedule	Task	
	1	2
X	R_3	R_5
Y	R_5	R_{10}

TABLE 5: Measurement value.

Task	1	2
X	R_3	R_5
D'	35.71	5.23
Pi. order (x_i)	20	60
$X + D'$	55.71	65.23
$Z = X + D'$	R_5	R_8

4.3. *Proposed DEM Algorithm.* The proposed algorithm (differential evolution multiskill, or DEM) is described in Algorithm 1:

- (i) Stop criterion: the loop *while-end while* would be stopped if the difference between *makespan* of some continuous generations is smaller than a threshold
- (ii) Size: number of solutions in the population
- (iii) F : parameter of the mutation operator
- (iv) CR: crossover probability
- (v) n : number of tasks

As shown in line 25, after classical DE steps such as mutation and selection are executed, function *Reallocate()* is called to reduce the makespan of the candidate schedule *bestnest*. Procedures of this function will be dealt with in more detail in the next section (Algorithm 2):

currentBest: the best schedule among the population of the current generation.

R_b : According to the arrangement of schedule *newbest* (line 3), R_b is the last resource that finishes its job. In Figure 3, before the *reallocate* function was called, R_b is resource 1.

maxResource(): the function which returns the last resource to finish (line 3).

size(): the function which returns the number of elements of the given set or array (line 5).

newmakespan: makespan of the new schedule (line 11), which is obtained from the old schedule by reassigning task i . The task i is moved from resource R_b (line 9) to resource $R_i[j]$ (line 10).

4.4. *Function Reallocate.* Line 12 and 13 demonstrate that the new schedule (*newbest*) is better than the old schedule (*currentBest*) in terms of makespan, which means that the function *Reallocate* certainly returns a better schedule. Moreover, the function *Reallocate* guides DEM algorithm in the right direction by finding a new schedule (*newbest*) based on the best solution (*currentBest*); therefore, it inherits the advanced genomes of the population as well as the old schedule.

Example 3. Given set of task $T = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, set of resource $RS = \{1, 2, 3\}$. Resource 1 can execute tasks 1, 2, 3, 4, 6, 8, 9, and 10; resource 2 can perform tasks 1, 3, 7, and 9; resource 3 can handle tasks 1, 4, 5, 8, 9, and 10. The relationship between tasks is depicted in Figure 1, and the duration of the tasks is given in Table 1.

Consider a schedule (PS) as shown in Table 6. From observation, Figure 4 shows that makespan $f(PS) = 17$.

New schedule, the result of the function *Reallocate*, is described in Table 7.

Figure 4 illustrates that function *Reallocate* converts the schedule PS into the new schedule by assigning task 9 to resource 2 instead of resource 1 as before. Thanks to this allocation, the new schedule achieves better makespan (15) compared to the old schedule (17).

5. Experiment

Experiments were conducted on the simulated dataset and the real dataset to evaluate the performance of proposed algorithms in comparison with previous algorithms such as GreedyDO and GA. We first utilize the simulated iMOPSE dataset, a collection of project instances that are artificially created based on a database obtained from the international company, to evaluate the performance of the proposed algorithm.

5.1. Simulation on Simulated Dataset

5.1.1. *iMOPSE Dataset.* The simulated dataset iMOPSE [5] has been used to investigate existing algorithms GreedyDO and GA [13, 36]; therefore, we conduct simulations on iMOPSE dataset to fairly compare between DEM and those algorithms.

As shown in Table 8, iMOPSE’s instances are classified into 15 categories according to the following attributes:

```

(1) Begin
(2) Load and validate iMOPSE dataset
(3)  $t \leftarrow 0$ 
(4) Size  $\leftarrow$  size of the population
(5)  $P(t) \leftarrow$  Initial population
(6)  $f(t) \leftarrow$  Calculate the fitness and makespan
(7) while (Stop criterion)
(8)   for (int  $i = 0; i < \text{Size}; i++$ )
(9)      $x_{r_1} \neq x_{r_2} \neq x_{r_3} \neq x_i \leftarrow \text{rand}(1, \text{Size})$ 
(10)     $F \leftarrow \text{rand}(0, 1)$ 
(11)     $v_i(t) \leftarrow x_{r_1} + F \times (x_{r_2} - x_{r_3})$  // parameter of the mutation operator
(12)    for ( $j = 0; j < n; j++$ )
(13)       $u_{i,j} = \begin{cases} v_{i,j} & \text{if } \text{rand}_{i,j} \leq \text{CR or } i = I_{\text{rand}} \\ x_{i,j} & \text{if } \text{rand}_{i,j} \geq \text{CR or } i \neq I_{\text{rand}} \end{cases}$ 
(14)    end for
(15)    if ( $f(u_i(t)) \leq f(x_i(t))$ )
(16)       $x_i(t+1) = u_i(t)$ 
(17)    else
(18)       $x_i(t+1) = x_i(t)$ 
(19)    end if
(20)  end for
(21)  Calculate the fitness and bestnest
(22)  If (makespan > min (fitness))
(23)    makespan = min (fitness)
(24)  End if
(25)  bestnest  $\leftarrow$  Reallocate (bestnest)
(26)   $t \leftarrow t + 1$ 
(27) end while
(28) return makespan
(29) End

```

ALGORITHM 1: DEM algorithm.

Input: currentBest //the best schedule among the current population
Output: //the improved schedule

```

(1) Begin
(2) makespan =  $f(\text{best})$ 
(3) newbest = currentBest;
(4)  $R_b \leftarrow \text{maxResource}(\text{newbest})$  //the last resource to finish its job
(5)  $T_b \leftarrow$  set of tasks is performed by resource  $R_b$ 
(6) For  $i = 1$  to size( $T_b$ ) // Consider each task in  $T_b$ , the set of tasks performed by resource  $R_b$ 
(7)    $T_i = T_b[i]$ ;
(8)    $R_i \leftarrow$  set of resource that are skilled enough to execute the task  $i$  except  $R_b$ 
(9)   For  $j = 1$  to size( $R_i$ ) // Consider each resource in turn
(10)     $R_i[j] = R_i[j] + \{i\}$  // Reassign task  $i$  to resource  $R_i[j]$ 
(11)     $R_b = R_b - \{i\}$  // Remove task  $i$  from the task list of  $R_b$ 
(12)    newmakespan =  $f(\text{newbest})$  // The makespan of the new schedule
(13)    If newmakespan < makespan
(14)      makespan = newmakespan
(15)      Return newbest; // Successful Reallocate, return the new and better schedule
(16)    End if
(17)    newbest = currentBest; // Unsuccessful Reallocate, reuse the original schedule for the next iteration
(18)  End for
(19) End for
(20) Return best
(21) End Function

```

ALGORITHM 2: Function reallocate.

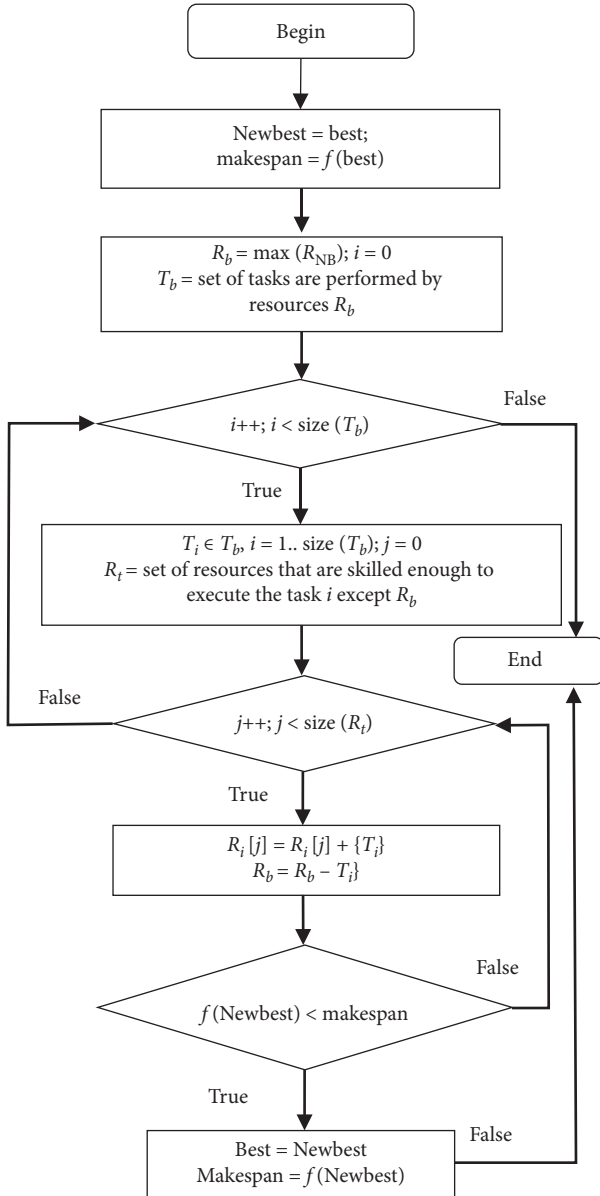


FIGURE 3: Function reallocate.

- (i) The number of tasks and resources
- (ii) The number of precedence relations between tasks
- (iii) The number of skills

All simulations were run on a PC with Intel Core i5-CPU 2.2 GHz, 6 GB RAM, and operating system Windows 10. All of iMOPSE instances used by our simulations are listed in Table 8.

5.1.2. Parameters and System Settings. The simulation is conducted by using the following:

- (i) Simulation environment: Matlab ver. 2014
- (ii) Simulation tool: the performance of previous algorithms such as GA is fairly evaluated by using GARunner [13]; the simulation tool was provided by the author of those algorithms

TABLE 6: Resource-task assignment of PS.

Task	1	2	3	4	5	6	7	8	9	10
Resource	1	1	2	3	3	1	2	1	3	3

- (iii) Input data: 15 iMOPSE instances that are described in Table 8
- (iv) Number of solution in the population N_p : 100
- (v) Number of generations N_g : 50,000
- (vi) Since all of the considered algorithms are approximate, we run on each instance 35 times to get average value, the standard deviation value, and the best value

5.1.3. Simulation Results. Simulation results are shown in Tables 9 and 10, where the makespan of the best solutions found by GreedyDO, GA, and the proposed algorithm (DEM) is illustrated (in hours). Table 9 shows that the makespans acquired by GA are always better than those acquired by GreedyDO.

Table 10 illustrates the average value, the best value, and the standard deviation of the makespan found by GA and the proposed DEM algorithm.

From the observation, the result in Tables 9 and 10 shows that

- (i) The solutions found by DEM are better than Greedy from 16% to 78% and GA from 0% to 21%.
- (ii) Regarding the average value (Table 9): DEM is better than Greedy from 15% to 77.5% and GA from 0.7% to 20.8%.
- (iii) Inspired by the differential evolution strategy, the proposed algorithm uses two parameters to guide the generation of the solution generated in the next generation: (i) the experience vector of both the population and (ii) the experience vector of the solution across different generations. DEM not only inherits those advantages of differential evolution approach but also utilizes the function reallocate, which improves its performance. Therefore, DEM has faster convergence and greater stability, which is shown by the standard deviation in Table 9 (column Std). The total standard deviation of DEM is 19.5 while the value of GA is 55.7; this result proves that the stability of DEM algorithm is better than GA.
- (iv) In addition, the use of a scale to represent solutions helps DEM calculate the deviations between solutions more accurately and create better solutions in the next generations.

5.2. Simulation on Real Dataset

5.2.1. TNG Dataset. Experiments conducted on the simulated dataset are sometimes inaccurate and unsatisfactory. To overcome this drawback, we collected the dataset of TNG Investment and Trading Joint Stock Company [37], a well-

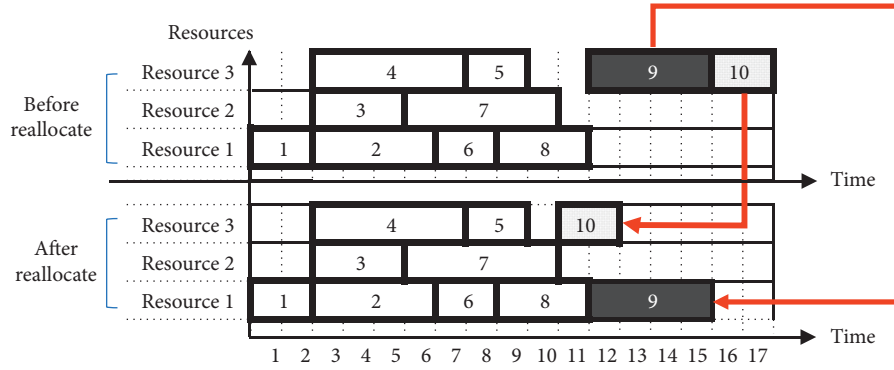
FIGURE 4: Gantt diagram of schedule before and after function *Reallocate*.

TABLE 7: Resource-task assignment of new schedule.

Task	1	2	3	4	5	6	7	8	9	10
Resource	1	1	2	3	3	1	2	1	1	3

TABLE 8: MS-RCPSP d36 iMOPSE dataset.

Dataset instance	Tasks	Resources	Precedence relations	Skills
100_5_48_9	100	5	48	9
100_5_64_15	100	5	64	15
100_5_64_9	100	5	64	9
100_10_64_9	100	10	64	9
100_10_65_15	100	10	65	15
100_20_65_15	100	20	65	15
100_20_65_9	100	20	65	9
200_10_84_9	200	10	84	9
200_10_85_15	200	10	85	15
200_20_55_9	200	20	55	9
200_20_97_15	200	20	97	15
200_20_97_9	200	20	97	9
200_40_45_9	200	40	45	9
200_40_90_9	200	40	90	9
200_40_91_15	200	40	91	15

known national textile company. TNG dataset instances have the following characteristics and parameters:

- (i) The company has product sewing contracts under the business partners' contracts; each order places a sample of a product in large quantities
- (ii) Each product has a certain number of stages, each of which requires a certain time to complete
- (iii) Each order will be assigned to a group of workers
- (iv) Each group consists of many workers
- (v) Each worker's skills are rated based on his rank

The above parameters of the dataset have been converted to a suitable format as shown in Table 11.

Table 11 demonstrates that the TNG dataset parameters will be represented as follows:

TABLE 9: The makespan of the best solutions found by GreedyDO and GA.

Dataset instance	GreedyDO	GA
100_5_48_9	779	528
100_5_64_15	640	527
100_5_64_9	597	508
100_10_64_9	533	296
100_10_65_15	426	286
100_20_65_15	310	240
100_20_65_9	408	181
200_10_84_9	999	567
200_10_85_15	706	549
200_20_55_9	999	312
200_20_97_15	680	424
200_20_97_9	816	321
200_40_45_9	821	209
200_40_90_9	963	211
200_40_91_15	519	200

- (i) The order from the business partners' is represented by a project
- (ii) The stage of the product is represented by a task
- (iii) A worker in the factory is represented by a resource in the problem formulation
- (iv) The worker grade is represented by the skill level of the resource
- (v) The order in which the tasks are executed is represented by the priority of the tasks
- (vi) Order execution time is represented by the total time of project implementation, i.e., the makespan

After the conversion, the dataset is shown in Table 12

5.2.2. Parameters and System Settings. Experiment setting:

- (i) Input data: 8 instances that are listed in Table 12
- (ii) Number of the solution in the population N_p : 100
- (iii) Number of generations N_g : 50,000

TABLE 10: The average value, the best value, and the standard deviation of makespan obtained from GA and DEM.

Dataset instance	GA			DEM		
	Avg	Best	Std	Avg	Best	Std
100_5_48_9	535	528	9.7	498	492	0.50
100_5_64_15	530	527	2.5	510	504	1.50
100_5_64_9	521	508	9.9	488	485	1.00
100_10_64_9	305	296	6.6	264	257	1.50
100_10_65_15	290	286	5.0	263	258	0.50
100_20_65_15	240	240	0.0	211	210	0.50
100_20_65_9	187	181	4.5	173	158	1.00
200_10_84_9	583	567	11.4	537	527	3.00
200_10_85_15	555	549	4.9	492	490	0.50
200_20_55_9	318	312	4.2	310	290	0.50
200_20_97_15	438	424	9.7	336	330	0.00
200_20_97_9	326	321	6.2	314	306	3.00
200_40_45_9	213	209	2.9	217	207	1.50
200_40_90_9	215	211	3.1	217	206	1.00
200_40_91_15	205	200	3.4	205	191	3.50

TABLE 11: Parameters conversion.

Real factor or parameter	The role
Order	Project
Stage of the product	Task
Time to execute the stage	Duration of the given task
Worker	Resource
The worker grade	Skill level of the resource
The order in which the tasks are executed	The priority of the tasks
Order execution time	The total time of project implementation, i.e., the makespan

TABLE 12: TNG dataset.

Dataset instance	Number of tasks	Number of resources	Precedence relations	Skill levels	Project time
TNG1	71	37	1026	6	409
TNG2	71	39	1026	6	325
TNG3	71	41	1026	6	296
TNG4	71	45	1026	6	392
TNG5	137	37	1894	6	1174
TNG6	137	39	1894	6	1052
TNG7	137	41	1894	6	871
TNG8	137	45	1894	6	996

- (iv) Like simulation on the simulated dataset, we run on each TNG dataset instance 35 times to get the best value

5.2.3. Simulation Results. To prove the effectiveness of the proposed algorithm (DEM), we carry out experiments by using the real-world dataset collected from the TNG company. DEM will be compared with other algorithms from previous research studies, GreedyDO and GA. The performance of these algorithms is recorded and observed by using GARunner, the tool created by Myszkowski et al. [13].

Table 13 contains the makespan (measured in hours) of the best solution achieved by GreedyDO, GA, and DEM. In addition, the performing intervals of the actual projects at the TNG factory are shown in column TNG.

Table 13 proves that schedules found by DEM are better than those acquired by GA and GreedyDO. Table 13 and Figure 5 also highlight that the schedules found by evolutionary algorithms are always shorter than the execution time of the actual projects at the TNG factory.

Compared to the execution time at the TNG factory, the GreedyDO and GA algorithms reduce the makespan by 4%–42% and 7%–51%, respectively, while DEM has the best

TABLE 13: Comparison between the makespan of real TNG project and makespan of schedule found by GreedyDO, GA, and DEM.

Dataset instance	TNG	GreedyDO	GA	DEM
TNG1	409	236	201	165
TNG2	325	243	198	166
TNG3	296	258	212	166
TNG4	392	248	176	165
TNG5	1174	972	751	712
TNG6	1052	963	791	712
TNG7	871	834	810	728
TNG8	996	906	720	675

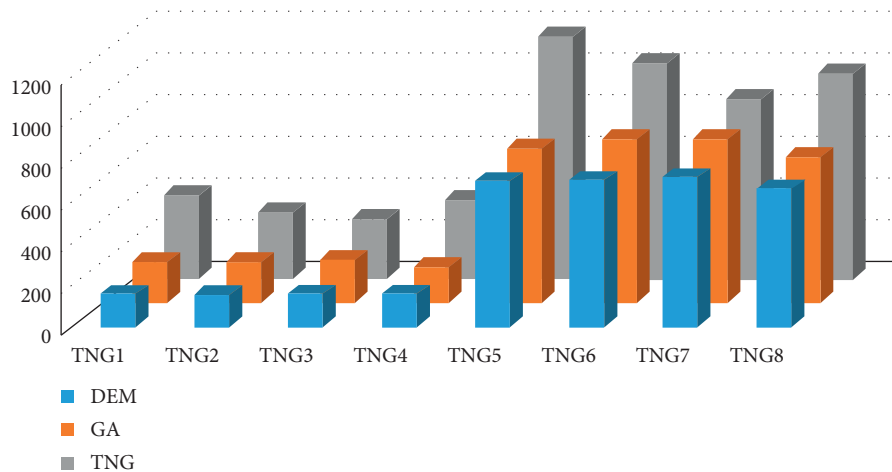


FIGURE 5: The comparison of makespan between DEM, GA, and TNG's solution.

results of 16%–61%. This result is understandable because so far all projects have been manually scheduled based on the scheduler's experience.

This statistical analysis of the experimental result proves that DEM achieved better solutions than the previous algorithms. Thanks to the directed nature, DEM not only ensures fast convergence but also averts getting trapped in local extrema. In experiments conducted on the TNG dataset, DEM reduces the makespan from 16% to 61% compared to the current factory schedule.

6. Conclusion and Future Work

This paper proposes to investigate the MS-RCPSPP, one of the combinatorial optimization problems that have a wide range of practical applications in science, technology, and life. The MS-RCPSPP is formulated and presented as a mathematical model, and then a novel algorithm called DEM which is based on differential evolution strategy is proposed.

The MS-RCPSPP, which is the subject of the investigation in this paper, has significant meaning for network resource management in the Internet of things. The MS-RCPSPP we study takes into account the resource's skill level, an important factor of the practical projects. Inspired by differential evolution metaheuristic, we have devised a new evolutionary algorithm that has the ability to outperform all the previous approaches. We build a function called

reallocate, which is the key component that creates the strength of the proposed algorithm. This function avoids the idle time intervals in a given schedule, improving the solution quality.

To analyze the efficiency of the proposed algorithm, various simulations are conducted and compared with the previous algorithms. Besides the simulated dataset, the real dataset supported by the TNG company has also been utilized for the assessment. The simulation result shows that our developed algorithm is more effective than existing algorithms. DEM achieves not only better quality solution but also faster convergence to global extremum in comparison with the prior algorithms.

In the next steps, we are planning to improve DEM algorithm by using a blended random walk and the Gauss statistical function. We will implement local search operations to further improve the quality of the solutions. At the same time, we will integrate the Gaussian probability distribution to improve the search direction modification in the solution space. Thanks to that improvement, we expect that the proposed algorithm will not fall into the local extreme trap.

In the future, we will utilize the DE metaheuristic with the integration of the Gauss distribution and Gauss statistical function [38, 39] to improve the effectiveness of the searching for the optimum solution and avoid the local extreme trap.

Data Availability

The paper uses the standard iMOPSE dataset to test the efficiency of the algorithm. This dataset is publicly available at <http://imopse.i.pwr.wroc.pl/> and is free of charge. In addition, we also tested the algorithm with TNG's garment manufacturing dataset. We have obtained permission from TNG to use these data.

Conflicts of Interest

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

Acknowledgments

We are grateful to the TNG Investment and Trading Joint Stock Company for giving us the information regarding their projects as well as providing us with their factory dataset.

References

- [1] Z. Ning, G. Xu, N. Xiong et al., "TAW: cost-effective threshold authentication with weights for Internet of Things," *IEEE Access*, vol. 7, pp. 30112–30125, 2019.
- [2] M. Alirezaei, S. T. A. Niaki, and S. A. A. Niaki, "A bi-objective hybrid optimization algorithm to reduce noise and data dimension in diabetes diagnosis using support vector machines," *Expert Systems with Applications*, vol. 127, pp. 47–57, 2019.
- [3] O. P. Mejia, "A new RCPSp variant to schedule research activities in a nuclear laboratory," in *Proceedings of the 47th International Conference on Computers and Industrial Engineering (CIE47)*, Lisbon, Portugal, October 2017.
- [4] K. Price, R. Storn, and J. Lampinen, *Differential Evolution-A Practical Approach to Global Optimization*, Springer, Berlin, Germany, 2005.
- [5] Y. Myszkowski, B. Paweł, and E. S. Marek, "A new benchmark dataset for multi-skill resource-constrained project scheduling problem," in *Proceedings of the 2015 Federated Conference on Computer Science and Information Systems (FedCSIS)*, IEEE, Sofia, Bulgaria, September 2015.
- [6] Y. Zeng, N. Xiong, and T. H. Kim, "Channel assignment and scheduling in multichannel wireless sensor networks," in *Proceedings of the 2008 33rd IEEE Conference On Local Computer Networks (LCN)*, pp. 512–513, Montreal, Canada, October 2008.
- [7] H. Cheng, N. Xiong, and L. T. Yang, "Distributed access scheduling algorithms in wireless mesh networks," in *Proceedings of the 22nd International Conference On Advanced Information Networking And Applications (Aina 2008)*, pp. 509–516, Okinawa, Japan, November 2008.
- [8] G. Wei, A. V. Vasilakos, and N. Xiong, "Scheduling parallel cloud computing services: an evolutionary game," in *Proceedings of the 2009 First International Conference on Information Science and Engineering*, pp. 376–379, Nanjing, China, December 2009.
- [9] W. Nie, N. Xiong, J. H. Park, and S. Yeo, "A fair-oriented two-level scheduling scheme for downlink traffic in wimax network," in *Proceedings of the 2010 2nd International Conference on Information Technology Convergence and Services*, pp. 1–6, Cebu, Philippines, January 2010.
- [10] W. Zheng, N. Xiong, N. Ghani, M. Peng, A. V. Vasilakos, and L. Zhou, "Adaptive scheduling for wireless video transmission in high-speed networks," in *Proceedings of the 2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 180–185, Shanghai, China, January 2011.
- [11] P. Dong, J. Xie, W. Tang, N. Xiong, H. Zhong, and A. V. Vasilakos, "Performance evaluation of multipath TCP scheduling algorithms," *IEEE Access*, vol. 7, pp. 29818–29825, 2019.
- [12] R. Klein, *Scheduling of Resource Constrained Project*, Springer Science Business Media, NewYork, NY, USA, 2000.
- [13] P. B. Myszkowski, M. Laszczyk, I. Nikulin, and E. Skowronski, "'iMOPSE: a library for bicriteria optimization in multi-skill resource-constrained project scheduling problem'" *Soft Computing*, vol. 23, no. 10, pp. 3397–3410, 2019.
- [14] H. Cheng, N. Xiong, X. Huang, and L. T. Yang, "An efficient scheduling model for broadcasting in wireless sensor networks," in *Proceedings of the 2013 IEEE International Symposium on Parallel & Distributed Processing*, pp. 1417–1428, IEEE, Cambridge, MA, USA, May 2013.
- [15] B. Lin, W. Guo, N. Xiong, G. Chen, A. V. Vasilakos, and H. Zhang, "A pretreatment workflow scheduling approach for big data applications in multicloud environments," *IEEE Transactions on Network and Service Management*, vol. 13, no. 3, pp. 581–594, 2016.
- [16] B. Lin, W. Guo, G. Chen, N. Xiong, and R. Li, "Cost-driven scheduling for deadline-constrained workflow on multi-clouds," in *Proceedings of the 2015 IEEE International Parallel and Distributed Processing Symposium Workshop*, pp. 1191–1198, IEEE, Hyderabad, India, May 2015.
- [17] L. Tan, Z. Zhu, F. Ge, and N. Xiong, "Utility maximization resource allocation in wireless networks: methods and algorithms," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 7, pp. 1018–1034, 2015.
- [18] P. Guo, W. Cheng, and Y. Wang, "A general variable neighborhood search for single-machine total tardiness scheduling problem with step-deteriorating jobs," *Journal of Industrial and Management Optimization*, vol. 10, no. 4, pp. 1071–1090, 2014.
- [19] S. Kavitha and P. Venkumar, "A vibrant crossbreed social spider optimization with genetic algorithm tactic for flexible job shop scheduling problem," *Measurement and Control*, vol. 53, no. 1-2, 2020.
- [20] A. P. Agrawal, A. Choudhary, and A. Kaur, "An effective regression test case selection using hybrid whale optimization algorithm," *International Journal of Distributed Systems and Technologies*, vol. 11, no. 1, pp. 53–67, 2020.
- [21] W. Guo, J. H. Park, L. T. Yang, A. V. Vasilakos, N. Xiong, and G. Chen, "Design and analysis of a MST-based topology control scheme with PSO for wireless sensor networks," in *Proceedings of the 2011 IEEE Asia-Pacific Services Computing Conference*, pp. 360–367, IEEE, Jeju, Korea, November 2011.
- [22] X. Zhuang, H. Cheng, N. Xiong, and L. T. Yang, "Channel assignment in multi-radio wireless networks based on pso algorithm," in *Proceedings of the 2010 5th International Conference on Future Information Technology*, pp. 1–6, Busan, Korea, November 2010.
- [23] C. T. Su, M. C. Santoro, and A. Mendes, "Constructive heuristics for project scheduling resource availability cost problem with tardiness," *Journal of Construction Engineering and Management*, vol. 144, Article ID 4018074, p. 8, 2018.
- [24] H. Maghsoudlou, B. Afshar-Nadjafi, S. T. Akhavan Niaki, and S. Niaki, "Multi-skilled project scheduling with level-

- dependent rework risk; three multi-objective mechanisms based on cuckoo search,” *Applied Soft Computing*, vol. 54, pp. 46–61, 2017.
- [25] K. Bibiks, Y.-F. Hu, J.-P. Li, P. Pillai, and A. Smith, “Improved discrete cuckoo search for the resource-constrained project scheduling problem,” *Applied Soft Computing*, vol. 69, pp. 493–503, 2018.
- [26] J. Lin, L. Zhu, and K. Gao, “A genetic programming hyper-heuristic approach for the multi-skill resource constrained project scheduling problem,” *Expert Systems with Applications*, vol. 140, p. 112915, 2020.
- [27] A. H. Hosseinian and V. Baradaran, “P-GWO and MOFA: two new algorithms for the MSRCPSP with the deterioration effect and financial constraints (case study of a gas treating company),” *Applied Intelligence*, vol. 50, pp. 2151–2176, 2020.
- [28] A. H. Hosseinian, V. Baradaran, and M. Bashiri, “Modeling of the time-dependent multi-skilled RCPSP considering learning effect,” *Journal of Modelling in Management*, vol. 10, 2019.
- [29] R. Nemati-Lafmejani, H. Davari-Ardakani, and H. Najafzad, “Multi-mode resource constrained project scheduling and contractor selection: mathematical formulation and metaheuristic algorithms,” *Applied Soft Computing*, vol. 81, Article ID 105533, 2019.
- [30] Y. Tahrir and S. Yang, “Hybrid meta-heuristic algorithms for independent job scheduling in grid computing,” *Applied Soft Computing*, vol. 72, pp. 498–517, 2018.
- [31] F. Ballestin and R. Leus, “Resource-Constrained project scheduling for timely project completion with stochastic activity durations,” *Production and Operations Management*, vol. 18, no. 4, pp. 459–474, 2009.
- [32] S. Javanmard, B. Afshar-Nadjafi, S. T. Akhavan Niaki, and S.T.K. Niaki, “Preemptive multi-skilled resource investment project scheduling problem: mathematical modelling and solution approaches,” *Computers & Chemical Engineering*, vol. 96, pp. 55–68, 2017.
- [33] V. Feoktistov, *Differential Evolution*, Springer, New York, NY, USA, 2006.
- [34] S. Tanabe, F. Ryoji, and A. Fukunaga, “Success-history based parameter adaptation for differential evolution,” in *Proceedings of the 2013 IEEE congress on evolutionary computation*, IEEE, Cancun, Mexico, November 2013.
- [35] T. K. Ghosh and S. Das, “A novel hybrid algorithm based on firefly algorithm and differential evolution for job scheduling in computational grid,” *International Journal of Distributed Systems and Technologies*, vol. 9, no. 2, pp. 1–15, 2018.
- [36] H. Najafzad, H. Davari-Ardakani, and R. Nemati-Lafmejani, “Multi-skill project scheduling problem under time-of-use electricity tariffs and shift differential payments,” *Energy*, vol. 168, pp. 619–636, 2019.
- [37] <http://www.tng.vn> TNG Investment and Trading Joint Stock Company, 434/1 bac kan street, Thai nguyen city, Viet Nam.
- [38] L. Kang, R.-S. Chen, N. Xiong, Y. C. Chen, Y.-X. Hu, and C.-M. Chen, “Selecting hyper-parameters of Gaussian process regression based on non-inertial particle Swarm optimization in internet of things,” *IEEE Access*, vol. 7, pp. 59504–59513, 2019.
- [39] M. Wu, N. Xiong, and L. Tan, “Adaptive range-based target localization using diffusion gauss–newton method in industrial environments,” *IEEE Transactions on Industrial Informatics*, vol. 15, no. 11, pp. 5919–5930, 2019.

Research Article

Addressing the Bike Repositioning Problem in Bike Sharing System: A Two-Stage Stochastic Programming Model

Qiong Tang,^{1,2,3} Zhuo Fu ,^{1,3} Dezhi Zhang,^{1,3} Hao Guo ,⁴ and Minyi Li⁵

¹School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China

²College of Economics and Management, Hengyang Normal University, Hengyang 421002, China

³Smart Transport Key Laboratory of Hunan Province, Changsha 410075, China

⁴School of Management, Jinan University, Guangzhou 510632, China

⁵School of Science, RMIT University, Melbourne, VIC 3000, Australia

Correspondence should be addressed to Zhuo Fu; zhfu@csu.edu.cn

Received 10 April 2020; Revised 26 April 2020; Accepted 5 May 2020; Published 20 May 2020

Academic Editor: Lu Zhen

Copyright © 2020 Qiong Tang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, a bike repositioning problem with stochastic demand is studied. The problem is formulated as a two-stage stochastic programming model to optimize the routing and loading/unloading decisions of the repositioning truck at each station and depot under stochastic demands. The goal of the model is to minimize the expected total sum of the transportation costs, the expected penalty costs at all stations, and the holding cost of the depot. A simulated annealing algorithm is developed to solve the model. Numerical experiments are conducted on a set of instances from 20 to 90 stations to demonstrate the effectiveness of the solution algorithm and the accuracy of the proposed two-stage stochastic model.

1. Introduction

Bike sharing system (BSS) as a means of sustainable and carbon-free transportation can effectively solve the “the first and last mile” problem of urban public transportation. BSS has gained increasing popularity in recent years due to the fact that it not only reduces urban pollution emissions and traffic congestion but also is considered as an effective way to improve the health of users [1]. A BSS consists of a depot and several stations, which scatter in different streets of the city. The success of a BSS depends on the availability of bikes. Due to the rental cycling of the user, there are often unbalanced situations at some stations of the system, i.e., there are either surplus or not enough bikes at these stations. In other words, either when there are surplus bikes in a station, they will be wasted as no users will need them, or in some stations where there are not enough bikes, a user’s demand cannot be met. To ensure the availability of bikes at all stations in the system, the BSS needs to employ redistribution trucks to transfer bikes from stations to stations, to balance the number of bikes according to the demand in each station. This problem

is commonly defined as the bike repositioning problem (BRP) in the BSS. The main objective of the BRP is to balance the supply and demand of bikes across stations as a mean to improve user satisfaction, while in the meantime, to reduce the transportation cost of redistribution trucks as much as possible. According to the classification of Berbeglia et al. [2], the bike repositioning problem is essentially a many-to-many pickup and delivery problem, and it is first proven to be an NP-hard problem by Benchimol et al. [3].

Typically, existing work assumes the redistribution demand at each station in the BSS is deterministic input parameters [4–21]. That is to say, these models heavily rely on that the loading and unloading quantities of redistribution trucks at each station be given in advance. However, this strong assumption generally could not hold in real-world conditions where the demand in urban public transportation commonly contains uncertainty and randomness [22]. As pointed out by Fricker and Gast [23], the effect of the uncertainty in user demand should not be neglected when studying the performance of a BSS. The authors have also demonstrated that incorporating knowledge of future

demands can lead to more accurate and robust decisions in repositioning.

The other challenge that might not have received enough attention in BRP is the consideration of the holding cost of the depot. In the BRP, bikes transferred from a bike surplus station to another bike deficient station is required to achieve the balance of supply and demand at all bike stations in the BSS to satisfy the user's demand. This perfect balance of supply and demand after repositioning operation must first ensure the internal balance in the BSS, i.e., the number of bikes at bike surplus stations should be equal to the number of bikes at bike deficient stations. In the actual operation process, some bikes might be lost or damaged after they are put into use. This happens quite often and will likely lead to the inherent imbalance between supply and demand in the BSS. For the bikes which cannot be balanced among the stations in the BSS, they are either picked up from or delivered back to the depot at the cost of increasing holding cost of the depot. Some BRP models have simply assumed the capacity of the depot is sufficient. As such, some bikes can be picked up from the depot to meet the demand of bike deficient stations, and/or some bikes at bike surplus stations can be brought back to the depot, in order to achieve a perfect balance between supply and demand at all stations in the BSS [4–7]. Unfortunately, these existing works have not yet considered the holding cost of the depot in the BRP model.

Stochastic programming is a technology to design mathematical programming models with stochastic demand, which helps decision makers to make more accurate decisions when the effect of the randomness of demand cannot be ignored. Although the optimization results obtained from stochastic programming cannot satisfy all future scenarios, they help to achieve better-than-average performance in dealing with stochastic demand.

To address the abovementioned concerns, in this paper, we take into account the stochastic nature of the BRP by capturing and modeling the uncertainty of redistribution demand within the system using stochastic programming. In particular, we propose a two-stage stochastic programming model for the bike repositioning problem with stochastic demand (BRPSD). In the first stage, we determine the routing decision, where “here-and-now” decision must be made before the realization of redistribution demand is known. Then, moving to the second stage, we further determine the loading/unloading decisions at each station and depot, where “wait-and-see” decisions are made taking future uncertainty into account. We incorporate the holding cost into the model such that the overall objective of the BRPSD is to determine the best routes of the repositioning truck and the optimal loading/unloading quantities at each station and depot, such that the expected total sum of the transportation costs, the penalty costs at all stations, and holding cost of the depot are minimized.

In this paper, stochastic redistribution demands are modeled by a set of discrete scenarios, and a predefined probability of occurrence is given for each scenario. The main contributions of this paper are summarized as follows:

- (1) A general scenario-based two-stage stochastic programming model for the bike repositioning problem with stochastic demands (BRPSD) is introduced. The proposed model takes the inventory holding cost of the depot into account in contexts where the loading and unloading activities are allowed in the depot.
- (2) An efficient metaheuristic, namely, simulated annealing algorithm is proposed to solve the defined BRPSD.
- (3) The necessity of considering redistribution demand as a stochastic demand in the BRPSD is further discussed with numerical experiments.

The remainder of this paper is organized as follows. We review the relevant literature in Section 2 and describe the mathematical formulation of the problem in Section 3. In section 4, we present the improved simulated annealing algorithm to solve the model. Then, we present experimental results in Section 5 and conclude this paper with discussions on potential future research in Section 6.

2. Literature Review

To the best of our knowledge, Benchimol et al. first formalise and address the BRP in [3]. Since then, various models and solution algorithms for the BRP have been discussed over the last decades.

We briefly review studies that are related to this paper, including heuristic and metaheuristic solution methods, for solving the deterministic BRP and models for the BRP with stochastic demands.

2.1. Heuristic Solution Algorithm for the Deterministic BRP.

According to Pal and Zhang [8], the BRP can be categorized into complete and partial repositioning problems according to the rigor of a repositioning that needs to be performed. In the former problem, the constraint for achieving a perfect balance among stations is considered as a hard constraint, and in the latter problem, it is a soft constraint.

In the complete repositioning problem, the optimal inventory of each station must be met after the repositioning operation. An iterated local search algorithm was first applied by Cruz et al. [9] to solve the single-vehicle BRP and then used by Bulhões et al. [10] to solve the multivehicle version of the BRP. Later, Pal and Zhang [8] modeled a static BRP in a free-floating bike sharing system and presented a hybrid nested large neighborhood search with a variable neighborhood descent algorithm to tackle the BRP. Dell'Amico et al. [11] developed a destroy and repair metaheuristic algorithm, which makes use of effective constructive heuristic and several local search procedures together for solving the BRP. Overall, the objective of the complete repositioning problem is mainly to pursue the minimum transportation cost by determining routing decisions. It has been a challenge to apply such methods in real-world systems effectively as there are other equally important criteria that have been neglected in these models.

In the partial repositioning problem, the ideal inventory of each station is not necessarily achieved after the repositioning operation. Nonetheless, the goal of these problems is not only to pursue the minimum transportation cost by determining routing decisions but also to minimize user dissatisfaction by determining loading and unloading decisions at each station. In the literature, a 3-step math heuristic for the multivehicle BRP is proposed by Forma et al. [6]. In their approach, stations are firstly clustered by a saving heuristic. Then, the routes through the clusters are constructed. Finally, the routing sequence through the stations in each cluster and the sequence through the clusters are determined. The last two steps are formulated as a mixed integer linear model solving by CPLEX. You [12] proposes a two-phase heuristic for the multivehicle BRP under a minimum service requirement over planning. In his proposed model, each stage represents a decision process. In the first phase, loading and unloading decisions for all stations for each time slot are developed by a linear programming model. Then, in the second phase, routing decisions are developed by an iterative approach through two parameter sensitive mathematical models.

In addition to the phased heuristic algorithm, many heuristics and metaheuristic solution methods have been presented to solve the deterministic BRP, to name a few, tabu search [5,13], genetic algorithm [14], chemical reaction optimization algorithm [15,16], artificial bee colony algorithm [17], constraint programming [18,19], variable neighborhood search [20], iterated local search [21], large neighborhood search algorithm [7], etc. Nonetheless, the practical applications of these existing approaches remain difficult and often bring in unexpected results. This is because uncertainty about user demands commonly exists in real-world BRPs; however, it is not yet captured by such deterministic models.

2.2. The BRP with Stochastic Demands. To the best of our knowledge, limited literature considers stochastic demands in the BRP. Two-stage stochastic programming with recourse models have been successfully used in solving bike repositioning problems as they allow the modeler to represent routing plans and loading and unloading activities together via first- and second-stage decision variables. Dell'Amico et al. [24] appear to be the pioneer paper that proposed a scenario-based two-stage stochastic model for bike repositioning problem, considering redistribution demands of stations as random variables. The authors provide five exact procedures as well as a heuristic algorithm that combines correlation-based constructive procedures with VND local search approach to solve their BRPSD. The heuristic algorithm only considers transportation costs while evaluating the local search move and only accept a feasible solution where the feasibility of the local search move is checked by an efficient strategy.

However, the BRPSD differs from our problem in the following three aspects. First, the objective of this model is to minimize the expected total travel distance of the truck and the penalty costs at all stations, while our problem minimizes

the expected total travel cost of the truck, the penalty costs at all stations, and the handling cost at the depot. Second, in our problem, we measure the local search move by using not only transportation costs but also the expected total costs. Third, we develop a method to quickly determine the feasible loading and unloading quantities at all stations and depot under any given route. Therefore, it is not necessary to check the feasibility of the solution in the local search.

There are also a few recent works that study a variation of the stochastic BRP where the stochastic demand is defined by origin-destination (OD) pair rather than the redistribution of stations. For instance, Maggioni et al. [25] propose two-stage stochastic programming models to determine the optimal number of bikes to assign to each station and the optimal number of transshipped bikes from one station to another station, respectively. They use AMPL and CPLEX to solve the model. Yan et al. [26] apply the time-space network model to determine the optimal locations of bike stations, bike fleet allocation, and bike routing. Their solution algorithm is based on a threshold-accepting-based heuristic.

From the abovementioned review of literature, we can see that the current scholars have proposed a rich heuristic algorithm to solve the deterministic BRP. Nonetheless, to the best of our knowledge, there is no work yet that effectively employs a metaheuristic algorithm to solve the BRP with stochastic demand. The study of this paper fills this theoretical gap.

3. Mathematical Formulation

The bike sharing system studied in this paper comprises a depot and a set of stations. Each station has a redistribution demand for each scenario, indicating the difference between the current inventory level and the optimal inventory level for each scenario. For a certain scenario, if the redistribution demand at a station is positive, the station is defined as a pickup station; if the redistribution demand at a station is negative, then the station is defined as a delivery station and should be supplied with bikes from pickup stations; if the redistribution demand at a station is equal to 0, then the station is called a balanced station. Balanced stations must also be visited.

We consider that only a single truck with a given capacity is available in the BSS. The truck collects bikes from pickup stations and transports them to delivery stations. Some bikes can be picked up from or delivered back to the depot, and each station is allowed to be visited only once. In addition, similar to Liu et al. [16], bikes are assumed to be parked anywhere in the BSS; therefore, the station capacity constraint is removed.

The decisions regarding the loading and unloading quantities at all stations and the depot in each scenario are made independently, according to the corresponding demand scenario. However, the routing decisions of the truck have to take into consideration all possible demand scenarios. This means that routing decisions are the same for all scenarios. Because the routing decisions to all scenarios are the same but the loading and unloading decisions vary in each scenario; there may exist a shortage or excess of bikes

for each station in each scenario. In this context, the following two-stage bike repositioning problem with stochastic demands (BRPSD) naturally arises. In the first stage, the BRPSD aims to construct a route for the truck that starts from and ends at the depot where each station is visited only once. In the second stage, the BRPSD aims to determine the optimal loading and unloading quantities at all stations and the depot to ensure the feasibility of the solution after the given route in the first stage. The objective of the BRPSD is to minimize the sum of the travel costs for the truck (first-stage objective function) and the expected total penalty costs at all stations and holding costs at the depot (second-stage objective function).

3.1. Two-Stage Stochastic Programming Model. Our objective is to determine the sequence of all stations to visit to minimize the transportation cost, as well as the loading/unloading decisions in each scenario to minimize the total expected penalty costs and holding costs. We formulate the problem as a two-stage stochastic programming problem. Given the discrete probability distribution of the scenario occurrences, the formulation for the SPBSD addressed in this study is as follows:

$$\min \sum_{i \in N_0} \sum_{j \in N_0, i \neq j} C_i t_{ij} x_{ij} + \sum_{\xi \in \Xi} p^\xi \left(\sum_{i \in N} (C_d (s_i^{\xi+} + s_i^{\xi-}) + C_h (y_{0i}^\xi + y_{i0}^\xi)) \right), \quad (1)$$

subject to

$$\sum_{j \in N} x_{0j} = \sum_{j \in N} x_{j0} = 1, \quad (2)$$

$$\sum_{j \in N_0, j \neq i} x_{ij} = 1, \quad i \in N, \quad (3)$$

$$\sum_{j \in N_0, j \neq i} x_{ij} = \sum_{j \in N_0, j \neq i} x_{ji}, \quad i \in N, \quad (4)$$

$$q_j \geq q_i + 1 - M(1 - x_{ij}), \quad \forall i, j \in N_0, i \neq j, \quad (5)$$

$$y_{ij}^\xi \leq Q x_{ij}, \quad \forall \xi \in \Xi, i \in N_0, j \in N_0, i \neq j, \quad (6)$$

$$\sum_{j \in N_0, j \neq i} y_{ij}^\xi - \sum_{j \in N_0, j \neq i} y_{ji}^\xi = d_i^\xi + s_i^{\xi-} - s_i^{\xi+}, \quad \forall \xi \in \Xi, i \in N, \quad (7)$$

$$q_i \geq 0, \quad \forall i \in N_0, \quad (8)$$

$$s_i^{\xi-}, s_i^{\xi+} \geq 0, \quad \forall i \in N, \xi \in \Xi, \quad (9)$$

$$y_{ij}^\xi \geq 0, \quad \forall i, j \in N_0, i \neq j, \xi \in \Xi, \quad (10)$$

$$x_{ij} \in \{0, 1\}, \quad \forall i, j \in N_0, i \neq j. \quad (11)$$

The objective function (1) is defined as the weighted sum of the transportation costs and the expected penalty costs at stations and holding costs at the depot over all scenarios. Constraint (2) states that the truck departs from the depot only once and returns to the depot. Constraint (3) ensures that the truck can visit a station only once. Constraint (4) ensures that if the truck visits a station, it must leave that station. Constraint (5) is the subtour elimination constraint [27]. Constraint (6) ensures that the load on the vehicle cannot be greater than the truck capacity. Constraint (7) requires that the number of bikes loaded onto or unloaded from the truck at a given station equals the difference between the truck load before and after the station visits. Constraint (8) ensures the auxiliary variable is nonnegative. Constraints (9)-(10) ensure that the excess/slack quantities of bikes at station and the number of bikes on the truck are nonnegative integers. Constraint (11) defines x_{ij} to be a binary variable.

4. Solution Algorithm

When the number of scenarios equals 1, the studied BRPSD reduces to the classical deterministic bike repositioning problems, an NP-hard combinatorial optimization problem. We employ simulated annealing (SA) to solve the studied BRPSD. SA is a metaheuristic algorithm based on local search, which can avoid falling into local optimization by accepting poor solutions with less probability in the process of iteration. It has been successfully applied to a variety of deterministic combinatorial problems [28,29] and stochastic combinatorial problems [30,31].

Our proposed algorithm, called SA_{BRPSD}, consists of two phases: constructing the initial solution and improving the initial solution by neighborhood search mechanism. The structure of this section is organized as follows. In Section 4.1, we propose the heuristic method to generate an initial solution. During initialization, two main decisions must be made: the routing decision (in Section 4.1.1) and the loading and unloading decision (i.e., inventory decision) (in Section 4.1.2). Then, in Section 4.2, we describe in detail of the local search operators used in this SA_{BRPSD}, namely, Swap, Relocate, and 2-opt. After that the main algorithm flow based on the SA_{BRPSD} framework is presented in Section 4.3.

4.1. Construct Initial Solution. Since the BRPSD is a two-stage decision-making process, the decision-making in this paper is divided into two stages: routing decisions are determined in the first stage, and the loading and unloading decisions are determined in the second stage. The process of constructing the initial solution is as follows, which is roughly divided into two main phases.

4.1.1. Construct the Initial Route for all Scenarios. Since the route in all scenarios is the same, the greedy heuristic algorithm is used to quickly construct the initial route by ignoring the stochastic redistribution demand. In this work, the nearest principle is used to construct the initial route by inserting stations one by one.

4.1.2. Determine Loading and Unloading Decision for Each Scenario. For a given scenario ξ and the corresponding redistribution demand d_i^ξ at station i in the scenario, the **Second-Stage-Opt** procedure uses a simple principle to quickly generate the initial feasible loading and unloading decision and then uses a reoptimize heuristic operator to improve the initial loading and unloading decision. The details of the **Second-Stage-Opt** procedure are given as follows.

Step 1: determine the initial loading and unloading quantities of all stations and the depot. The loading and unloading quantities at each station are defined as the corresponding redistribution demand. Because the demand of the depot is 0, the loading and unloading quantities are also defined as 0. As a result, the initial solution of the BRPSD is generated, including two-stage decisions, namely, the routing decision and the loading and unloading decision. It is obvious that the initial solution is often not feasible.

Step 2: adjust the loading and unloading quantities of some stations to achieve a feasible initial solution. If the number of bikes on the truck when it travels directly from station i to station j in scenario ξ $y_{ij}^\xi > Q$ or $y_{ij}^\xi < 0$, then the initial solution is not feasible. In this case, the loading and unloading quantity of station i should be adjusted. Set $\gamma = y_{ij}^\xi - Q$, if $y_{ij}^\xi > Q$ or $\gamma = y_{ij}^\xi$, if $y_{ij}^\xi < 0$, then the loading and unloading quantity of station i is adjusted to subtract γ from the original quantity.

Step 3 (reoptimize): the basic idea of the reoptimize operator is to optimize the solution by a heuristic method that adjusts the loading and unloading quantities between stations and the depot [5,7]. In this paper, some bikes can be collected from the depot to meet the delivery demands of some stations, and also some bikes collected from some stations can be delivered to the depot at the end of the repositioning. In this procedure, there are two cases. In the first case, the truck picks up extra bikes when it starts from the depot and drops them off at some stations. In the second case, the truck picks up more bikes from some stations and drops them off when it ends at the depot.

4.2. Neighborhood Structure. In this paper, the local search operator aims at routing decisions. Three kinds of classical operators Swap move, Relocate move, and 2-opt move are used. These operators are also often used in other BRP problems. Each new neighborhood solution is generated by these three moves with equal probability. Note that, after a local search operator is executed, the new solution needs to be improved by *Second-Stage-Opt* procedure to obtain the second-stage decision.

Swap- N^1 : randomly select two stations from the route and then exchange them

Relocate- N^2 : randomly select one station, delete it from the current route, and then insert it into another position

2-Opt- N^3 : randomly select two stations, exchange the positions of the two stations, and change the orientation of the route between the two stations

4.3. The Proposed SA_{BRPSD}. Algorithm 1 provides the pseudocodes of the SA_{BRPSD}. Parameter T_0 denotes the initial temperature; T_E is the final temperature; α is the coefficient used to control the cooling; I_{iter} indicates the number of generated new solutions under a certain temperature; and Max_{iter} is the max number of the accepted new solutions under a certain temperature. In this section, we use S_0 and S^{best} to denote the initial solution and the best solution obtained so far, respectively. The best objective value $F(S^{best})$ is set to be the objective function value of the best solution S^{best} .

As shown in Algorithm 1, in the inner loop of the SA_{BRPSD} algorithm, a random new neighborhood solution S^{new} is initially generated by *Local search* procedure to improve the first-stage decision and then followed by *Second-Stage-Opt* procedure to generate the second-stage decision. The stopping criteria for the inner loop are either it has generated I_{iter} new solutions or accepted Max_{iter} new solution under the current temperature T . When the current temperature T decreases to T_E , the SA_{BRPSD} algorithm terminates.

5. Numerical Examples

In this section, numerical examples are conducted to examine the validity and efficiency of the proposed model and solving algorithm. In addition, the necessity of considering stochastic demands is examined by analyzing the expected value of perfect information (EVPI) and the value of the stochastic solution (VSS).

In this study, we consider the instances as in [24] to conduct the numerical experiments (the instances are available at <http://www.or.unimore.it/site/home/online-resources>). The size of the instances ranges from 20 to 90 stations. There are a total of 20 instances and 30 scenarios for each instance. In any instance, the first station is chosen to be the depot. To better analyze the impact of relevant parameters on the solution of the stochastic model, for a given set N_0 , station i and station j belong the set N_0 , and we redefine the parameter transportation time t_{ij} as $t_{ij} = t_{ij}/\min\{t_{ij}\}$.

The SA_{BRPSD} was coded in Matlab, and all computational experiments were carried out on a computer with an Intel Core i5-4590 CPU @ 3.30 GHz and 4 GB RAM. Each instance was run 10 times, the best and average of the solutions, and the average computing time of 10 runs were used to evaluate the performance of the algorithm solution.

The SA_{BRPSD} proposed in this study relies on five parameters, namely, the initial temperature T_0 ; the final temperature T_E ; the coefficient used to control the cooling α ; the number of generated solutions under a certain temperature I_{iter} ; and the number of accepted solutions under a certain temperature Max_{iter} . After some preliminary experiments, the results from the algorithm were obtained by

```

Obtain the initial solution  $S_0$ .
Set the parameter for the simulated annealing,  $T_0, T_E, \alpha, I_{iter}, \text{Max}_{iter}$ 
Set  $S^{\text{best}} \leftarrow S_0; S^{\text{current}} \leftarrow S_0; T \leftarrow T_0$ .
While  $T \geq T_E$  do
   $m \leftarrow 0; n \leftarrow 0;$ 
  While  $m \leq I_{iter}$  &  $n \leq \text{Max}_{iter}$  do
     $S \leftarrow \text{Local search}(S^{\text{current}}, N^k);$ 
     $S^{\text{new}} \leftarrow \text{Second-Stage-Opt}(S);$ 
     $\Delta \leftarrow F(S^{\text{new}}) - F(S^{\text{current}});$ 
    If  $\Delta \leq 0$  then
       $S^{\text{current}} \leftarrow S^{\text{new}}, n \leftarrow n + 1;$ 
      If  $F(S^{\text{current}}) \leq F(S^{\text{best}})$  then
         $S^{\text{best}} \leftarrow S^{\text{current}};$ 
      end
    else
      If  $e^{-\Delta/T} \geq \text{random}$  then
         $S^{\text{current}} \leftarrow S^{\text{new}}, n \leftarrow n + 1;$ 
      end
    end
     $m \leftarrow m + 1;$ 
  end
   $T \leftarrow \alpha \square T;$ 
end
Return:  $S^{\text{best}}$ 

```

ALGORITHM 1: SA_{BRPSD}.

setting $T_0 = 20$, $T_E = 0.1$, $\alpha = 0.97$, $I_{iter} = 3 * |N_0|$, and $\text{Max}_{iter} = |N_0|$.

5.1. The Performance Comparison between Lingo and SA Algorithm. In terms of the quality of the solution and the calculation efficiency, Tables 1–4 show the comparison results obtained by Lingo and SA_{BRPSD} under different truck capacity Q , the holding cost C_h , and the number of stations. The CPU (in seconds) indicates the average computing time of the SA_{BRPSD} or the computing time for Lingo, respectively. Gap_{Best} (%) and Gap_{Avg} (%) indicate the performance of the SABRP relative to that of Lingo based on the best and average objective values, respectively. When $|N_0| \geq 50$, Lingo is not able to obtain any feasible solutions for the problems, and thus, the results are not shown in the tables. It also generally fails to obtain any global optimal solutions within 3600 seconds even for small-size problems (with four exceptions for the results obtained under instance *Washington (20)2*). From Tables 1–4, we can see that SA_{BRPSD} is significantly faster than Lingo in all instances, i.e., it has better time efficiency to solve the BRPSD compared to Lingo. For the various sizes of problems from small to large, the average computing time of the SA_{BRPSD} varies from 66.14 s to 70.22 s.

As shown in Tables 1–4, all of the values of Gap_{Best} and Gap_{Avg} are negative. This implies that the SA_{BRPSD} performs better than Lingo. The absolute values of the average Gap_{Best} and Gap_{Avg} values obtained by the SA_{BRPSD} for the tests with $Q = 20$ are larger than those with $Q = 10$. As the capacity of the repositioning truck increases, the objective function value decreases. This may be explained by the fact that if the

truck capacity increases, the truck can load/unload more bikes at one station, resulting in an expansion of the solution space. Overall, we can draw the conclusion that the proposed SA_{BRPSD} can obtain better feasible optimal solutions much faster than Lingo.

5.2. The Necessity of Considering Demand as a Stochastic Parameter. To evaluate the performance of the stochastic model we proposed in this paper, we use the two concepts of the evaluation approach: the expected value of perfect information (EVPI) and the value of the stochastic solution (VSS) [32–34]. EVPI measures the maximum amount a decision maker would be ready to pay in return for perfect information about the future, and VSS measures the amount of cost saving when the decision maker uses expected values of stochastic parameters instead of stochastic parameters in the model (the amount of cost due to ignoring the uncertainty). EVPI is defined as the difference between the wait-and-see solution (WS) and the here-and-now solution (HN). In the EVPI, the WS solution is the expected value of the foresight. We obtain the value by averaging the results of each stochastic scenario. The HN solution is the objective value of this two-stage stochastic model. VSS is defined as the difference between the expected results of using the expected value problem solution (EEV) and the here-and-now solution. In the VSS, the value of EEV is obtained by inputting the optimal solution of the optimization problem associated with the mean value of stochastic demands into the two-stage stochastic model. The gaps Gap_{VSS} and Gap_{EVPI} for VSS and EVPI are calculated by VSS/HN and EVPI/WS , respectively.

TABLE 1: Comparison of the results between Lingo and SA_{BRPSD} (Q = 20 and C_h = 1).

Instance	Lingo		Best value	SA _{BRPSD}		Gap _{Best} (%)	Gap _{Avg} (%)
	Optimal value	CPU		Average value	CPU		
Washington (20)1	639.15	3600	639.16	639.17	10.15	0.00	0.00
Washington (20)2	576.79	159	576.79	576.79	24.58	0.00	0.00
Chicago (20)1	166.24	3600	153.40	153.40	20.85	-7.72	-7.72
Chicago (20)2	115.19	3600	113.46	113.46	14.04	-1.50	-1.50
Washington (30)1	450.19	3600	387.78	388.03	30.77	-13.86	-13.81
Washington (30)2	520.39	3600	403.12	405.09	25.22	-22.53	-22.16
Chicago (30)1	518.02	3600	265.71	266.96	54.21	-48.71	-48.47
Chicago (30)2	204.14	3600	202.21	204.14	29.48	-0.94	0.00
Washington (40)1	797.87	3600	634.07	638.52	44.71	-20.53	-19.97
Washington (40)2	1036.51	3600	826.68	831.62	29.94	-20.24	-19.77
Chicago (40)1	1287.45	3600	424.34	424.34	52.13	-67.04	-67.04
Chicago (40)2	478.01	3600	213.10	215.92	48.37	-55.42	-54.83
Washington (50)1			546.18	554.30	105.62		
Washington (50)2			603.40	611.72	77.76		
Chicago (50)			326.32	328.59	70.84		
Washington (66)			682.51	685.62	101.87		
Chicago (66)			6243.75	6494.41	122.49		
Washington (80)1			719.65	732.09	121.20		
Washington (80)2			678.58	684.66	159.47		
Washington (90)			925.21	938.56	196.93		
Avg		3335			66.14	-24.15	-23.85

TABLE 2: Comparison of the results between Lingo and SA_{BRPSD} (Q = 10 and C_h = 1).

Instance	Lingo		Best value	SA _{BRPSD}		Gap _{Best} (%)	Gap _{Avg} (%)
	Optimal value	CPU		Average value	CPU		
Washington (20)1	668.94	3600	662.46	662.50	12.63	-0.97	-0.96
Washington (20)2	576.79	663	576.79	576.79	24.63	0.00	0.00
Chicago (20)1	164.14	3600	164.14	164.14	19.19	0.00	0.00
Chicago (20)2	136.83	3600	135.19	135.22	16.13	-1.20	-1.17
Washington (30)1	447.37	3600	393.75	393.75	34.19	-11.99	-11.99
Washington (30)2	501.14	3600	414.52	417.24	28.02	-17.28	-16.74
Chicago (30)1	300.48	3600	299.42	300.05	34.42	-0.35	-0.14
Chicago (30)2	299.76	3600	242.00	242.69	33.23	-19.27	-19.04
Washington (40)1	736.88	3600	688.63	692.12	60.31	-6.55	-6.07
Washington (40)2	911.88	3600	855.64	863.20	38.21	-6.17	-5.34
Chicago (40)1	576.94	3600	449.41	449.41	47.47	-22.10	-22.10
Chicago (40)2	538.80	3600	243.16	245.98	50.99	-54.87	-54.35
Washington (50)1			581.86	588.84	85.98		
Washington (50)2			647.38	650.63	75.82		
Chicago (50)			353.99	354.51	69.17		
Washington (66)			676.54	683.93	90.98		
Chicago (66)			6192.78	6468.38	120.21		
Washington (80)1			733.58	744.35	114.69		
Washington (80)2			736.77	745.27	183.02		
Washington (90)			984.90	996.38	228.76		
Avg		3355			68.40	-11.73	-11.49

The test instance *Chicago 20(1)* used for explaining the total costs of WS, HN, EEV, EVPI, and VSS is a small-size instance consisting of 20 stations. We set the truck capacity $Q = 10$, $C_h = 1$, $C_d = 1$, and $C_t = 1$.

In this test instance, the total costs can be solved by the SA_{BRPSD} as follows: WS = 145.31, HN = 153.40, and EEV = 161.31. Then, calculate $VSS = EEV - HN = 7.90$, $EVPI = HN - WS = 8.09$, $Gap_{VSS} = 5.15\%$, and $Gap_{EVPI} = 5.57\%$. Both values of EVPI and VSS lead to the conclusion that

managers must pay 5.57% more of the total cost of the stochastic solutions to obtain the perfect redistribution demand information about the future. Ignoring the uncertainty will let managers in a situation where they have a risk to pay at least 5.15% higher if they only focus on a scenario only (deterministic solutions).

The results of the sensitivity analysis on Gap_{VSS} and Gap_{EVPI} for the truck capacity Q , the holding cost C_h , the penalty cost C_d , and the travel cost C_t are explained in the

TABLE 3: Comparison of the results between Lingo and SA_{BRPSD} ($Q=20$ and $C_h=0$).

Instance	Lingo		SA _{BRPSD}			Gap _{Best} (%)	Gap _{Avg} (%)
	Optimal value	CPU	Best value	Average value	CPU		
Washington(20)1	645.10	3600	627.85	627.90	13.76	-2.67	-2.67
Washington(20)2	556.26	40	556.26	556.26	33.82	0.00	0.00
Chicago(20)1	134.73	3600	134.73	134.73	20.53	0.00	0.00
Chicago(20)2	100.80	3600	99.60	99.76	17.95	-1.19	-1.04
Washington(30)1	414.18	3600	369.72	369.72	32.90	-10.73	-10.73
Washington(30)2	432.25	3600	392.72	393.61	27.89	-9.15	-8.94
Chicago(30)1	489.71	3600	249.47	249.80	35.62	-49.06	-48.99
Chicago(30)2	321.25	3600	189.40	191.21	31.73	-41.04	-40.48
Washington(40)1	721.11	3600	617.41	622.76	50.49	-14.38	-13.64
Washington(40)2	1021.87	3600	814.77	819.79	42.10	-20.27	-19.78
Chicago(40)1	412.36	3600	408.94	408.94	57.98	-0.83	-0.83
Chicago(40)2	397.27	3600	201.52	203.18	50.88	-49.27	-48.86
Washington(50)1			526.11	534.88	109.26		
Washington(50)2			592.72	595.78	79.45		
Chicago(50)			321.13	324.60	69.87		
Washington(66)			656.99	663.09	105.25		
Chicago(66)			6243.75	6494.41	122.49		
Washington(80)1			702.93	716.69	126.66		
Washington(80)2			661.88	673.31	171.68		
Washington(90)			914.40	924.30	204.03		
Avg		3303.33			70.22	-16.55	-16.33

TABLE 4: Comparison of the results between Lingo and SA_{BRPSD} ($Q=10$ and $C_h=0$).

Instance	Lingo		SA _{BRPSD}			Gap _{Best} (%)	Gap _{Avg} (%)
	Optimal value	CPU	Best value	Average value	CPU		
Washington(20)1	665.08	3600	654.91	655.01	10.76	-1.53	-1.52
Washington(20)2	566.26	30	566.26	566.26	25.00	0.00	0.00
Chicago(20)1	155.17	3600	155.17	155.17	19.20	0.00	0.00
Chicago(20)2	127.97	3600	127.97	127.97	17.38	0.00	0.00
Washington(30)1	449.66	3600	382.56	382.56	33.10	-14.92	-14.92
Washington(30)2	442.69	3600	410.45	411.18	28.40	-7.28	-7.12
Chicago(30)1	292.28	3600	289.57	290.67	33.25	-0.93	-0.55
Chicago(30)2	236.60	3600	234.10	234.10	32.92	-1.06	-1.06
Washington(40)1	818.25	3600	681.61	685.40	61.54	-16.70	-16.24
Washington(40)2	859.04	3600	853.74	857.84	47.10	-0.62	-0.14
Chicago(40)1	500.44	3600	442.74	442.74	47.85	-11.53	-11.53
Chicago(40)2	349.26	3600	239.51	240.61	53.41	-31.42	-31.11
Washington(50)1			571.78	579.44	106.77		
Washington(50)2			638.40	640.92	75.32		
Chicago(50)			352.67	354.72	72.96		
Washington(66)			664.31	671.87	89.90		
Chicago(66)			6175.14	6472.54	117.22		
Washington(80)1			728.64	739.63	114.94		
Washington(80)2			722.88	743.14	183.62		
Washington(90)			984.90	996.38	228.76		
Avg		3303			69.97	-7.17	-7.02

TABLE 5: Results of different vehicle capacities (Q).

Q	WS	HN	EEV	VSS	Gap _{VSS} (%)	EVPI	Gap _{EVPI} (%)
10	153.86	164.14	168.57	4.43	2.70	10.28	6.68
15	147.66	156.87	163.24	6.37	4.06	9.22	6.24
20	145.31	153.40	161.31	7.90	5.15	8.09	5.57
25	144.69	151.60	160.97	9.37	6.18	6.91	4.78
30	144.35	150.87	160.41	9.54	6.32	6.53	4.52
Avg				7.52	4.88	8.21	5.56

TABLE 6: Results of different holding costs (C_h).

C_h	WS	HN	EEV	VSS	Gap _{VSS} (%)	EVPI	Gap _{EVPI} (%)
0.5	137.68	144.73	150.29	5.56	3.84	7.05	5.12
0.75	141.89	149.73	155.80	6.07	4.05	7.84	5.53
1	145.31	153.40	161.31	7.90	5.15	8.09	5.57
1.25	147.43	157.02	166.82	9.79	6.24	9.60	6.51
1.5	148.80	160.64	172.32	11.69	7.28	11.84	7.96
Avg				8.20	5.31	8.91	6.16

TABLE 7: Results of different penalty costs (C_d).

C_d	WS	HN	EEV	VSS	Gap _{VSS} (%)	EVPI	Gap _{EVPI} (%)
0.5	120.92	125.00	128.24	3.24	2.59	4.08	3.38
0.75	133.14	139.57	144.77	5.21	3.73	6.43	4.83
1	145.31	153.40	161.31	7.90	5.15	8.09	5.57
1.25	157.43	167.19	177.84	10.65	6.37	9.76	6.20
1.5	169.50	180.97	194.37	13.40	7.41	11.47	6.77
Avg				8.08	5.05	7.97	5.35

TABLE 8: Results of different travel costs (C_t).

C_t	WS	HN	EEV	VSS	Gap _{VSS} (%)	EVPI	Gap _{EVPI} (%)
0.5	96.81	104.30	113.72	9.42	9.04	7.49	7.74
0.75	121.09	128.84	137.51	8.68	6.74	7.75	6.40
1	145.31	153.40	161.31	7.90	5.15	8.09	5.57
1.25	169.49	177.97	185.10	7.13	4.01	8.49	5.01
1.5	193.61	202.07	208.90	6.83	3.38	8.47	4.37
Avg				7.99	5.66	8.06	5.82

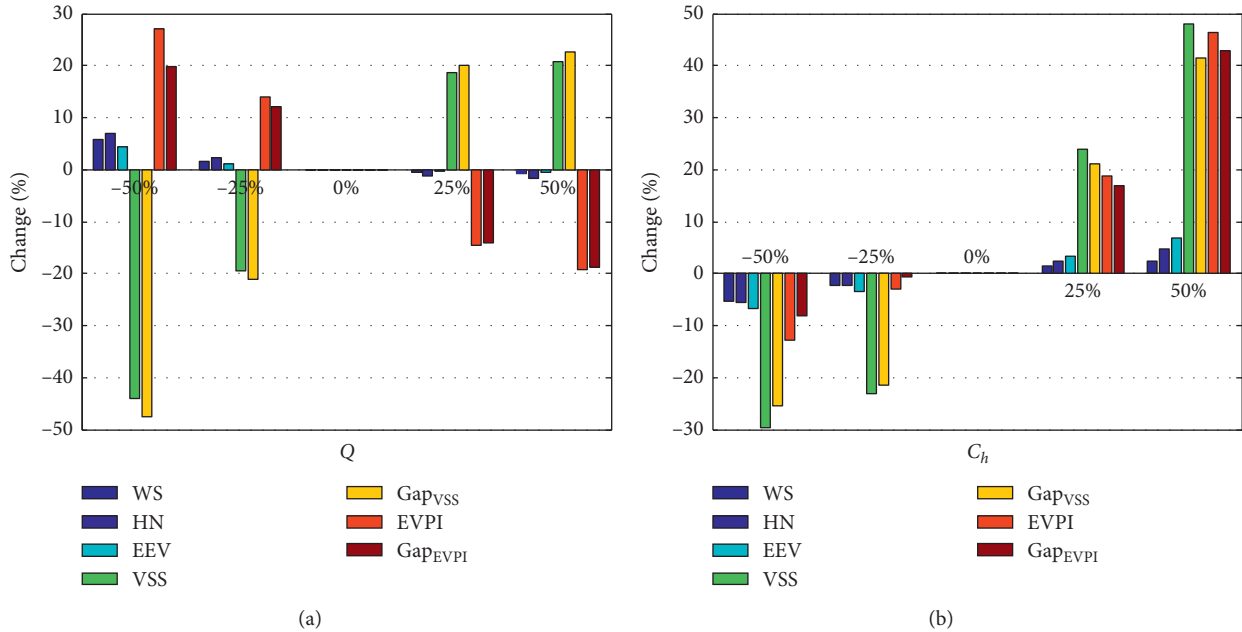


FIGURE 1: Continued.

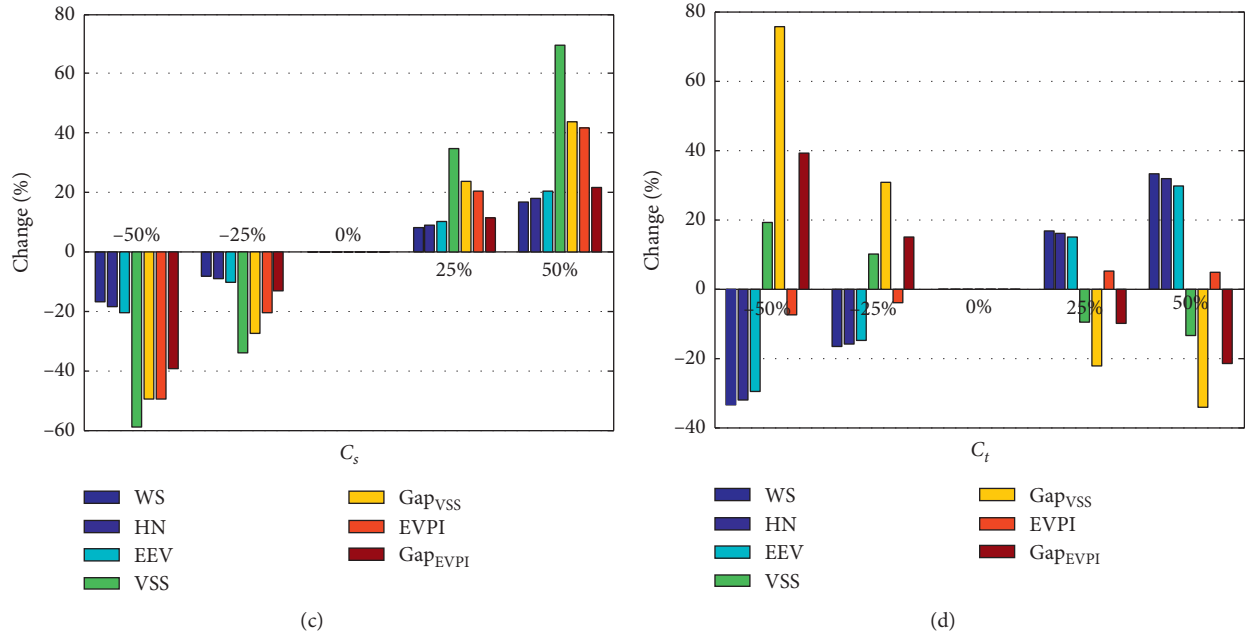


FIGURE 1: Sensitivity analysis on different values of parameters: (a) Q ; (b) C_h ; (c) C_d ; (d) C_t .

following. The value of the four parameters are increased by -50% , -25% , $+25\%$, and $+50\%$, respectively. Tables 5–8 summarize the total costs of WS, HN, EEV, EVPI and VSS, GAP_{EVPI} , and GAP_{VSS} obtained from all tests. Figure 1 shows the sensitivity analysis on the results for the four parameters. Under current settings, the abovementioned four parameters have a significant impact on the solutions. From Tables 5–8, consistent with theoretical results, it can be founded that $WS < HN < EEV$ in all tests. We can see that the average GAP_{VSS} are all greater than the average GAP_{EVPI} . The GAP_{EVPI} of the stochastic model in all tests is non-negative, indicating the problem of the BRPSD is stochastic by nature. Besides, the GAP_{VSS} in all tests is nonnegative for the BRPSD, indicating that the problem in the study is suitable for the stochastic model.

From Figure 1, we can see that the increase of C_h , C_d and C_t rends the increase of HN. We also see that the increase of Q rends the decrease of HN. The parameter that has the greatest impact on HN is the unit transportation cost C_t . This is likely because the contribution of transportation cost to the total objective cost is greater than the penalty cost and the holding cost.

From Figure 1, we can see that, as Q increases, GAP_{EVPI} tends to decrease, but GAP_{VSS} increases. It is also noticed that GAP_{EVPI} and GAP_{VSS} tend to increase when C_h and C_d increase. However, the value of GAP_{EVPI} and GAP_{VSS} contradicts the increase of C_t . The influence of the change of the abovementioned four parameters on GAP_{VSS} is significantly greater than that of GAP_{EVPI} .

The results illustrate that the GAP_{EVPI} increases when C_h and C_d increase and Q and C_t decrease, which means that the decision makers are willing to pay more to get accurate redistribution demand information about the future. GAP_{VSS} increases by increasing the values of C_h , C_d , and Q or

decreasing the values of C_t , and consequently, the decision makers can obtain more cost saving resulting from the two-stage stochastic model.

6. Conclusions

In this paper, a general scenario-based two-stage programming model is proposed for bike repositioning problems with stochastic redistribution demands. In the proposed model, the first-stage decisions correspond to the routing decisions and the second-stage decisions correspond to the loading and unloading decisions. The model aims to find the best routes and the ideal quantities of loading and unloading at each station and depot in order to minimize the weighted sum of the expected transportation costs, penalty costs, and holding costs. Then, a simulated annealing algorithm is proposed to solve the model. Number examples are performed to evaluate the proposed model and algorithm, followed by a detailed sensitivity analysis that studies how the change of several important parameters affects the performance of the proposed model and solution algorithm. The proposed model and algorithm have important theoretical and practical significance for the BSS operators, which can reduce the operation costs of the BSS and improve user satisfaction. On the contrary, in addition to bike sharing, other sharing facilities in the sharing economy, such as car sharing, also need to be relocated, so this kind of problem has a strong application background.

In the future, we will apply robust optimization model to solve the bike repositioning problem with stochastic demand in the bike sharing system [35]. We also consider environmental issues by adding the carbon emissions to the objective function of the two-stage stochastic programming model [36–38].

Notations

Sets

- N : The set of stations
 N_0 : The set of nodes, including the stations (indexed by i , $i \in N$) and the depot (indexed by 0)
 Ξ : Set of scenarios, indexed by $\xi \in \Xi$

Parameters

- C_t : The travel cost per unit travel time
 t_{ij} : The travel time from station i to station j
 C_d : The penalty cost per bike at the station
 C_h : The holding cost per bike at the depot
 Q : The truck capacity
 M : A very large number
 d_i^ξ : The relocation demand at station i in scenario ξ
 p^ξ : The occurrence probability of the scenario ξ

Decision Variables

- x_{ij} : Binary variable that equals to 1 if the truck travels directly from station i to station j and 0 otherwise
 y_{ij}^ξ : The number of bikes on the truck when it travels directly from station i to station j in scenario ξ
 $s_i^{\xi+}$: The excess quantities of bikes at station i after relocation operation in scenario ξ
 $s_i^{\xi-}$: The slack quantities of bikes at station i after relocation operation in scenario ξ
 q_i : Auxiliary variables associated with station i used for subtour elimination constraints.

Data Availability

All data included in 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

The authors acknowledge the National Natural Science Foundation of China (Grant no. 71271220) and the Project of Social Science Achievement Review Committee of Hunan Province (Grant no. XSP20YBZ165).

References

- [1] Y. Yao, Y. Zhang, L. Tian, N. Zhou, and Z. Li, "Analysis of network structure of urban bike-sharing systems: a case study based on real-time data of a public bicycle system," *Sustainability*, vol. 11, no. 19, Article ID 5425, 2019.
- [2] G. Berbeglia, J.-F. Cordeau, I. Gribkovskaia, and G. Laporte, "Static pickup and delivery problems: a classification scheme and survey," *Top*, vol. 15, no. 1, pp. 1–31, 2007.
- [3] M. Benchimol, P. Benchimol, B. Chappert et al., "Balancing the stations of a self service "bike hire" system," *RAIRO—Operations Research*, vol. 45, no. 1, pp. 37–61, 2011.
- [4] T. Raviv, M. Tzur, and I. A. Forma, "Static repositioning in a bike-sharing system: models and solution approaches," *EURO Journal on Transportation and Logistics*, vol. 2, no. 3, pp. 187–229, 2013.
- [5] S. C. Ho and W. Y. Szeto, "Solving a static repositioning problem in bike-sharing systems using iterated tabu search," *Transportation Research Part E: Logistics and Transportation Review*, vol. 69, pp. 180–198, 2014.
- [6] I. A. Forma, T. Raviv, and M. Tzur, "A 3-step math heuristic for the static repositioning problem in bike-sharing systems," *Transportation Research Part B: Methodological*, vol. 71, pp. 230–247, 2015.
- [7] S. C. Ho and W. Y. Szeto, "A hybrid large neighborhood search for the static multi-vehicle bike-repositioning problem," *Transportation Research Part B: Methodological*, vol. 95, pp. 340–363, 2017.
- [8] 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.
- [9] F. Cruz, A. Subramanian, B. P. Bruck, and M. Iori, "A heuristic algorithm for a single vehicle static bike sharing rebalancing problem," *Computers & Operations Research*, vol. 79, pp. 19–33, 2017.
- [10] T. Bulhões, A. Subramanian, G. Erdoğan, and G. Laporte, "The static bike relocation problem with multiple vehicles and visits," *European Journal of Operational Research*, vol. 264, no. 2, pp. 508–523, 2018.
- [11] M. Dell'Amico, M. Iori, and S. Novellani, "A destroy and repair algorithm for the bike sharing rebalancing problem," *Computer & Operations Research*, vol. 71, pp. 149–162, 2016.
- [12] P.-S. You, "A two-phase heuristic approach to the bike repositioning problem," *Applied Mathematical Modelling*, vol. 73, pp. 651–667, 2019.
- [13] D. Chemla, F. Meunier, and R. Wolfler Calvo, "Bike sharing systems: solving the static rebalancing problem," *Discrete Optimization*, vol. 10, no. 2, pp. 120–146, 2013.
- [14] Y. Li, W. Y. Szeto, J. Long, and C. S. Shui, "A multiple type bike repositioning problem," *Transportation Research Part B: Methodological*, vol. 90, pp. 263–278, 2016.
- [15] W. Y. Szeto, Y. Liu, and S. C. Ho, "Chemical reaction optimization for solving a static bike repositioning problem," *Transportation Research Part D: Transport and Environment*, vol. 47, pp. 104–135, 2016.
- [16] Y. Liu, W. Y. Szeto, and S. C. Ho, "A static free-floating bike repositioning problem with multiple heterogeneous vehicles, multiple depots, and multiple visits," *Transportation Research Part C: Emerging Technologies*, vol. 92, pp. 208–242, 2018.
- [17] W. Y. Szeto and C. S. Shui, "Exact loading and unloading strategies for the static multi-vehicle bike repositioning problem," *Transportation Research Part B: Methodological*, vol. 109, pp. 176–211, 2018.
- [18] L. D. Gaspero, A. Rendl, and T. Urli, "Balancing bike sharing systems with constraint programming," *Constraints*, vol. 21, no. 2, pp. 318–348, 2016.
- [19] L. Di Gaspero, A. Rendl, and T. Urli, "Constraint-based approaches for balancing bike sharing systems," in *Principles and Practice of Constraint Programming*, vol. 8124, pp. 758–773, Springer, Berlin, Germany, 2013.
- [20] M. Rainer-Harbach, P. Papazek, G. R. Raidl, B. Hu, and C. Kloimüller, "PILOT, GRASP, and VNS approaches for the

- static balancing of bicycle sharing systems,” *Journal of Global Optimization*, vol. 63, no. 3, pp. 597–629, 2015.
- [21] Q. Tang, Z. Fu, and M. Qiu, “A bilevel programming model and algorithm for the static bike repositioning problem,” *Journal of Advanced Transportation*, vol. 2019, Article ID 8641492, 19 pages, 2019.
- [22] J. Warrington and D. Ruchti, “Two-stage stochastic approximation for dynamic rebalancing of shared mobility systems,” *Transportation Research Part C: Emerging Technologies*, vol. 104, pp. 110–134, 2019.
- [23] C. Fricker and N. Gast, “Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity,” *EURO Journal on Transportation and Logistics*, vol. 5, no. 3, pp. 261–291, 2016.
- [24] M. Dell’Amico, M. Iori, S. Novellani, and A. Subramanian, “The bike sharing rebalancing problem with stochastic demands,” *Transportation Research Part B: Methodological*, vol. 118, pp. 362–380, 2018.
- [25] F. Maggioni, M. Cagnolari, L. Bertazzi, and S. W. Wallace, “Stochastic optimization models for a bike-sharing problem with transshipment,” *European Journal of Operational Research*, vol. 276, no. 1, pp. 272–283, 2019.
- [26] S. Yan, J.-R. Lin, Y.-C. Chen, and F.-R. Xie, “Rental bike location and allocation under stochastic demands,” *Computers & Industrial Engineering*, vol. 107, pp. 1–11, 2017.
- [27] C. E. Miller, A. W. Tucker, and R. A. Zemlin, “Integer programming formulation of traveling salesman problems,” *Journal of the ACM (JACM)*, vol. 7, no. 4, pp. 326–329, 1960.
- [28] Y. Xiao, Q. Zhao, I. Kaku, and Y. Xu, “Development of a fuel consumption optimization model for the capacitated vehicle routing problem,” *Computers & Operations Research*, vol. 39, no. 7, pp. 1419–1431, 2012.
- [29] H. Guo, C. Li, Y. Zhang, C. Zhang, and Y. Wang, “A nonlinear integer programming model for integrated location, inventory, and routing decisions in a closed-loop supply chain,” *Complexity*, vol. 2018, Article ID 2726070, 17 pages, 2018.
- [30] Y. Zhang, M. Qi, W.-H. Lin, and L. Miao, “A metaheuristic approach to the reliable location routing problem under disruptions,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 83, pp. 90–110, 2015.
- [31] 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.
- [32] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, Springer Science & Business Media, Berlin, Germany, 2011.
- [33] S.-L. Hu, C.-F. Han, and L.-P. Meng, “Stochastic optimization for joint decision making of inventory and procurement in humanitarian relief,” *Computers & Industrial Engineering*, vol. 111, pp. 39–49, 2017.
- [34] E. Nikzad, M. Bashiri, and F. Oliveira, “Two-stage stochastic programming approach for the medical drug inventory routing problem under uncertainty,” *Computers & Industrial Engineering*, vol. 128, pp. 358–370, 2019.
- [35] Y. Liu, H. Lei, Z. Wu, and D. Zhang, “A robust model predictive control approach for post-disaster relief distribution,” *Computers & Industrial Engineering*, vol. 135, pp. 1253–1270, 2019.
- [36] J. Jiang, D. Zhang, S. Li, and Y. Liu, “Multimodal green logistics network design of urban agglomeration with stochastic demand,” *Journal of Advanced Transportation*, vol. 2019, Article ID 4165942, 2019.
- [37] S. Li, Z. Wang, X. Wang, D. Zhang, and Y. Liu, “Integrated optimization model of a biomass feedstock delivery problem with carbon emissions constraints and split loads,” *Computers & Industrial Engineering*, vol. 137, Article ID 106013, 2019.
- [38] Y. Wang and W. Y. Szeto, “Static green repositioning in bike sharing systems with broken bikes,” *Transportation Research Part D: Transport and Environment*, vol. 65, pp. 438–457, 2018.