

Discrete Dynamics in Nature and Society

# Discrete Optimization for Dynamic Systems of Operations Management in Data-Driven Society

Special Issue Editor in Chief: Lu Zhen

Guest Editors: Shuaian Wang, Xiaobo Qu, and Xinchang Wang





---

**Discrete Optimization for Dynamic  
Systems of Operations Management  
in Data-Driven Society**

Discrete Dynamics in Nature and Society

---

**Discrete Optimization for Dynamic  
Systems of Operations Management  
in Data-Driven Society**

Special Issue Editor in Chief: Lu Zhen

Guest Editors: Shuaian Wang, Xiaobo Qu, and Xinchang Wang



---

Copyright © 2019 Hindawi. All rights reserved.

This is a special issue published in "Discrete Dynamics in Nature and Society." 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.

## Editorial Board

Douglas R. Anderson, USA  
David Arroyo, Spain  
Viktor Avrutin, Germany  
Stefan Balint, Romania  
Kamel Barkaoui, France  
Gian I. Bischi, Italy  
Gabriele Bonanno, Italy  
Florentino Borondo, Spain  
Driss Boutat, France  
Filippo Cacace, Italy  
Pasquale Candito, Italy  
Giulio E. Cantarella, Italy  
Antonia Chinnì, Italy  
Cengiz Çinar, Turkey  
Carmen Coll, Spain  
Giancarlo Consolo, Italy  
Alicia Cordero, Spain  
Giuseppina D'Agui, Italy  
Manuel De la Sen, Spain  
Beatrice Di Bella, Italy  
Luisa Di Paola, Italy  
Josef Diblík, Czech Republic  
Xiaohua Ding, China  
Tien Van Do, Hungary  
Marcio Eisencraft, Brazil  
Elmetwally Elabbasy, Egypt  
Hassan A. El-Morshedy, Egypt  
D. Fournier-Prunaret, France  
Genni Fragnelli, Italy  
Ciprian G. Gal, USA  
Marek Galewski, Poland  
Caristi Giuseppe, Italy  
David M. Goldsman, USA  
Gisèle R Goldstein, USA  
Vladimir Gontar, Israel  
Chris Goodrich, USA

Pilar R. Gordo, Spain  
Kannan Govindan, Denmark  
Luca Guerrini, Italy  
Juan L. G. Guirao, Spain  
Antonio Iannizzotto, Italy  
Giuseppe Izzo, Italy  
Sarangapani Jagannathan, USA  
Jun Ji, USA  
Emilio Jiménez Macías, Spain  
Nikos I. Karachalios, Greece  
Eric R. Kaufmann, USA  
Nickolai Kosmatov, USA  
Victor S. Kozyakin, Russia  
Mustafa R. S. Kulenovic, USA  
Kousuke Kuto, Japan  
Aihua Li, USA  
Xiaohui Liu, UK  
Francesco Longo, Italy  
Miguel Ángel López, Spain  
Ricardo López-Ruiz, Spain  
Rodica Luca, Romania  
Agustin Martin, Spain  
Akio Matsumoto, Japan  
Rigoberto Medina, Chile  
Driss Mehdi, France  
Vicenc Méndez, Spain  
Dorota Mozyrska, Poland  
Yukihiko Nakata, Japan  
Luca Pancioni, Italy  
G. Papashinopoulos, Greece  
Juan Pavón, Spain  
Ewa Pawluszewicz, Poland  
Alfred Peris, Spain  
Allan C. Peterson, USA  
Adrian Petrusel, Romania  
Andrew Pickering, Spain

Chuanxi Qian, USA  
Morteza Rafei, Netherlands  
Youssef N. Raffoul, USA  
Maria Alessandra Ragusa, Italy  
Mustapha A. Rami, Spain  
Aura Reggiani, Italy  
Pavel Rehak, Czech Republic  
Paolo Renna, Italy  
Marko Robnik, Slovenia  
Yuriy Rogovchenko, Norway  
Silvia Romanelli, Italy  
Christos J. Schinas, Greece  
Daniel Sevcovic, Slovakia  
Leonid Shaikhet, Israel  
Seenith Sivasundaram, USA  
Charalampos Skokos, South Africa  
Piergiulio Tempesta, Spain  
Tetsuji Tokihiro, Japan  
J. R. Torregrosa, Spain  
Delfim F. M. Torres, Portugal  
Fabio Tramontana, Italy  
Firdaus E. Udwardia, USA  
Jose C. Valverde, Spain  
Antonia Vecchio, Italy  
Rafael J. Villanueva, Spain  
Francisco R. Villatoro, Spain  
Hubertus Von Bremen, USA  
Abdul-Aziz Yakubu, USA  
Bo Yang, USA  
Guang Zhang, China  
Zhengqiu Zhang, China  
Lu Zhen, China  
Yong Zhou, China  
Zuonong Zhu, China

# Contents

## **Discrete Optimization for Dynamic Systems of Operations Management in Data-Driven Society**

Lu Zhen , Shuaian Wang, Xiaobo Qu, and Xinchang Wang  
Editorial (5 pages), Article ID 3597314, Volume 2019 (2019)

## **Probability Mechanism Based Particle Swarm Optimization Algorithm and Its Application in Resource-Constrained Project Scheduling Problems**

Shuai Li , Zhicong Zhang , Xiaohui Yan , and Liangwei Zhang  
Research Article (11 pages), Article ID 9085320, Volume 2019 (2019)

## **Omni-Channel Product Distribution Network Design by Using the Improved Particle Swarm Optimization Algorithm**

Si Zhang , Heying Zhu, Xia Li, and Yan Wang  
Research Article (15 pages), Article ID 1520213, Volume 2019 (2019)

## **Analysis of Signal Game for Supply Chain Finance (SCF) of MSEs and Banks Based on Incomplete Information Model**

Zhang Tao, Xin Li , Xinquan Liu , and Nana Feng   
Research Article (6 pages), Article ID 3646097, Volume 2019 (2019)

## **Bilevel Programming Model of Urban Public Transport Network under Fairness Constraints**

Jingjing Hao , Xinquan Liu , Xiaojing Shen , and Nana Feng   
Research Article (10 pages), Article ID 2930502, Volume 2019 (2019)

## **Incorporating a Bayesian Network into Two-Stage Stochastic Programming for Blood Bank Location-Inventory Problem in Case of Disasters**

Shutong Chen  and Changjun Wang   
Research Article (28 pages), Article ID 7214907, Volume 2019 (2019)

## **Evaluating the Spatial Deprivation of Public Transportation Resources in Areas of Rapid Urbanization: Accessibility and Social Equity**

Chunyang Han , Xinquan Liu , Xiaojing Shen , Ling Zhang , and Nana Feng  
Research Article (11 pages), Article ID 6890362, Volume 2019 (2019)

## **Discrete Optimization Model and Algorithm for Driver Planning in Periodic Driver Routing Problem**

Lin Huang, Wenya Lv, Qian Sun, and Chengle Ma   
Research Article (10 pages), Article ID 9476362, Volume 2019 (2019)

## **Research on the Scheduling Problem of Movie Scenes**

Yulian Liu, Qiuji Sun , Xiaotian Zhang, and Yiwei Wu   
Research Article (8 pages), Article ID 3737105, Volume 2019 (2019)

## **Simulation Optimization of Discrete Logistics Processes: A Case Study on Logistics of an E-Commerce Enterprise in Shanghai**

Xiang Xu , Yue Li, and Ruiyun Tang   
Research Article (11 pages), Article ID 2493638, Volume 2019 (2019)

**Pricing Decisions on Reward-Based Crowdfunding with Bayesian Review System Facing Strategic Consumers**

Yuting Chen, Rong Zhang, and Bin Liu 

Research Article (14 pages), Article ID 4592690, Volume 2019 (2019)

**Research on Supply Chain Coordination Based on Block Chain Technology and Customer Random Demand**

Yongfei Li, Bill Wang, and Dong Yang 

Research Article (10 pages), Article ID 4769870, Volume 2019 (2019)

**Adopting a QCA Approach to Investigating the Risks Involved in Megaprojects from Auditing Perspective**

Liang Ma, Bing Zhang , Miao Cui, and Ruoyu Jin

Research Article (9 pages), Article ID 4037859, Volume 2019 (2019)

**A Metaheuristic Algorithm to Transporter Scheduling for Assembly Blocks in a Shipyard considering Precedence and Cooperating Constraints**

Ning-Rong Tao, Zu-Hua Jiang, Jian-Feng Liu, Bei-Xin Xia , and Bai-He Li

Research Article (14 pages), Article ID 2615154, Volume 2019 (2019)

**Distance-Based Congestion Pricing with Day-to-Day Dynamic Traffic Flow Evolution Process**

Qixiu Cheng, Jiping Xing, Wendy Yi, Zhiyuan Liu , and Xiao Fu 

Research Article (7 pages), Article ID 7438147, Volume 2019 (2019)

**Applying Big Data Analytics to Monitor Tourist Flow for the Scenic Area Operation Management**

Siyang Qin , Jie Man , Xuzhao Wang, Can Li, Honghui Dong, and Xinquan Ge

Research Article (11 pages), Article ID 8239047, Volume 2019 (2019)

**An Improved Genetic Algorithm Based Robust Approach for Stochastic Dynamic Facility Layout Problem**

Yunfang Peng, Tian Zeng, Lingzhi Fan, Yajuan Han, and Beixin Xia 

Research Article (8 pages), Article ID 1529058, Volume 2018 (2019)

**Optimizing the Composite Cost Involved in Road Motor-Transporting Trucks by Taking into Account Traffic Condition**

Hongying Fei  and Chengyi Zhang

Research Article (11 pages), Article ID 8703852, Volume 2018 (2019)

**Optimal Scheme for Process Quality and Cost Control by Integrating a Continuous Sampling Plan and the Process Yield Index**

Chunzhi Li, Shurong Tong , and Keqin Wang

Research Article (14 pages), Article ID 1917252, Volume 2018 (2019)

**Data-Driven Recovery Potential Analysis and Modeling for Batteries Recovery Operations in Electric Bicycle Industry**

Ping Zhang  and Guangfu Liu

Research Article (18 pages), Article ID 6783190, Volume 2018 (2019)

**Implementation Flexibility of Multiperiod Rail Line Design with Consideration of Uncertainties in Population Distribution**

Kunpeng Zhang, Binglei Xie, and Ding Liu 

Research Article (17 pages), Article ID 5715304, Volume 2018 (2019)

**Optimization of Vehicle Routing with Pickup Based on Multibatch Production**

Hongtao Hu, Jiao Mo , and Chengle Ma 

Research Article (9 pages), Article ID 2804589, Volume 2018 (2019)

**Integrated Optimization on Assortment Packing and Collaborative Shipping for Fashion Clothing**

Qianying Wang, Yiping Jiang , and Yang Liu

Research Article (10 pages), Article ID 3973801, Volume 2018 (2019)

**Optimal Strategies for Manufacturers with the Reference Effect under Carbon Emissions-Sensitive Random Demand**

Baojun Zhang , Shaojian Qu , Panpan Li, and Ripeng Huang

Research Article (15 pages), Article ID 2452406, Volume 2018 (2019)

**A Multiobjective Route Robust Optimization Model and Algorithm for Hazmat Transportation**

Changxi Ma , Wei Hao , Ruichun He , and Bahman Moghimi

Research Article (12 pages), Article ID 2916391, Volume 2018 (2019)

**Hydrological Layered Dialysis Research on Supply Chain Financial Risk Prediction under Big Data Scenario**

Jia Liu, Shiyong Li , and Xiaoxia Zhu 

Research Article (9 pages), Article ID 3259858, Volume 2018 (2019)

**A Metaheuristic Method for the Task Assignment Problem in Continuous-Casting Production**

Hongtao Hu , Yiwei Wu, and Tingsong Wang 

Research Article (12 pages), Article ID 8073648, Volume 2018 (2019)

**Inventory and Production Dynamics in a Discrete-Time Vendor-Managed Inventory Supply Chain System**

Yongchang Wei , Fangyu Chen , and Hongwei Wang 

Research Article (15 pages), Article ID 6091946, Volume 2018 (2019)

**Dynamic Strategies on Firm Production and Platform Advertisement in Crowdfunding considering Investor's Perception**

Ying Ji , Ju Wei , Zhong Wu, Shaojian Qu , and Baojun Zhang 

Research Article (12 pages), Article ID 3726080, Volume 2018 (2019)

**Goodwill and System Dynamics Modeling for Film Investment Decision by Interactive Efforts**

Jian Feng and Bin Liu 

Research Article (10 pages), Article ID 7521689, Volume 2018 (2019)

---

**An Incentive Dynamic Programming Method for the Optimization of Scholarship Assignment**

Di Huang, Yu Gu , Hans Wang, Zhiyuan Liu , and Jun Chen 

Research Article (7 pages), Article ID 5206131, Volume 2018 (2019)

**Optimizing Price of Credit Default Swaps for Dynamic Project System of Public-Private Partnership**

Ming Wu, Wenya Lv, and Qiuji Sun 

Research Article (10 pages), Article ID 7280974, Volume 2018 (2019)

**Complexity Analysis of Dynamic Cooperative Game Models for Supply Chain with the Remanufactured Products**

Jianwei Chang and Liuwei Zhao 

Research Article (10 pages), Article ID 9246934, Volume 2018 (2019)

## Editorial

# Discrete Optimization for Dynamic Systems of Operations Management in Data-Driven Society

Lu Zhen <sup>1</sup>, Shuaian Wang,<sup>2</sup> Xiaobo Qu,<sup>3</sup> and Xinchang Wang<sup>4</sup>

<sup>1</sup>*School of Management, Shanghai University, Shanghai, China*

<sup>2</sup>*The Hong Kong Polytechnic University, Hung Hom, Hong Kong*

<sup>3</sup>*Chalmers University of Technology, Gothenburg, Sweden*

<sup>4</sup>*Mississippi State University, Starkville, USA*

Correspondence should be addressed to Lu Zhen; [lzhen@shu.edu.cn](mailto:lzhen@shu.edu.cn)

Received 7 May 2019; Accepted 7 May 2019; Published 15 May 2019

Copyright © 2019 Lu Zhen 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.

Society is being reshaped by large volumes of data generated from various operations, and research in handling dynamic issues related to data-driven systems has greatly increased in the past decades, where most rely upon discrete optimization models for handling dynamic features contained in operations management (OM) activities. With the rapid growth in computational technologies, including data mining technologies, discrete event simulation techniques, and intelligence algorithms, OM relies more and more on optimal solutions (or their approximates) based on high-performance models and algorithms.

This special issue collects 32 original research contributions that present recent advances in models and algorithms concerning discrete optimizations on dynamic OM systems relevant to the data-driven society. In particular, the papers address the topics related to discrete optimization methodologies for stochastic OM problems, dynamic programming based exact methods for stochastic OM problems, system dynamics in behavior OM for E-commerce, data-driven risk analysis and modeling for OM decisions in dynamic contexts, data-driven models for dynamic supply chain management, etc. It is hoped that this special issue will help the integration of the latest research achievements in the relevant field.

The paper by S. Li et al. entitled “Probability Mechanism Based Particle Swarm Optimization Algorithm and Its Application in Resource-Constrained Project Scheduling Problems” proposes a new probability mechanism based particle swarm optimization (PMPSO) algorithm to solve

combinatorial optimization problems by introducing new particles based on the optimal particles in the population and the historical optimal particles in the individual changes. This method is applied to solve resource-constrained project scheduling problems and the experimental results are quite encouraging.

S. Zhang et al. in the paper entitled “Omni-Channel Product Distribution Network Design by Using the Improved Particle Swarm Optimization Algorithm” construct the joint randomization planning model of location and routing for the minimization of total cost of a supply chain distribution network under uncertain customer demand.

The paper entitled “Analysis of Signal Game for Supply Chain Finance (SCF) of MSEs and Banks Based on Incomplete Information Model” by Z. Tao et al. uses a signal gaming model based on incomplete information to analyze the decisions of commercial banks and medium-size and small enterprises (SMEs) in supply chain finance business, and concludes that the returns of banks are closely relied on the probability of good SMEs join in supply chain finance business and the default cost is an important constraint for determining the strategies adopted by both SMEs and banks.

The paper by J. Hao et al. entitled “Bilevel Programming Model of Urban Public Transport Network under Fairness Constraints” presents a bilevel programming model of the public transport network considering factors such as the per capita occupancy area and travel cost of different groups to alleviate the urban transportation equity and optimize the

urban public transport network under fairness constraints. The upper layer minimizes the travel cost deprivation coefficient and the road area Gini coefficient by taking into account fairness constraints; the lower layer models the selection behavior of different groups in the optimal scheme obtained at the upper layer as a stochastic equilibrium traffic assignment problem with multimode and multiuser. An improved genetic algorithm is developed to solve the constructed model and is validated via a simple network. The experimental results show that the current model can improve significantly the transportation equity feeling of low-income groups, reduce the overall travel time via public transport network, and thus attract more high and middle income users who owned cars to use the public transport network.

S. Chen and C. Wang in the paper entitled “Incorporating a Bayesian Network into Two-Stage Stochastic Programming for Blood Bank Location-Inventory Problem in Case of Disasters” construct a Bayesian Network to describe a blood logistics network considering uncertainties and interdependences caused by earthquakes. A two-stage multiperiod stochastic programming model is developed where the first stage focuses on blood bank location and inventory planning problem and the second stage is composed of multiple periods, some of which may suffer disasters and initiate corresponding rescue operations.

The paper entitled “Evaluating the Spatial Deprivation of Public Transportation Resources in Areas of Rapid Urbanization: Accessibility and Social Equity” by C. Han et al. introduces an evaluation system with six indexes to evaluate the spatial differentiation of public transportation resources and services. SPSS is applied to analyze the data collected from a typical rapid urbanization area in China and two main factors are shown to have significant impact on the spatial deprivation associated with public transportation resources and services. The authors conclude that the public transportation situation in rapid urbanization areas is consistent with the local land-use context and the methods applied in this study are suitable for extracting spatial public transportation characteristics.

The paper by L. Huang et al. entitled “Discrete Optimization Model and Algorithm for Driver Planning in Periodic Driver Routing Problem” presents a mixed-integer linear programming model for a periodic driver routing problem with an objective of minimizing the total workload by taking into account the relationship between workload differential among drivers and total workload. A local branching based method is developed to solve large instances of the problem and numerical experiments are conducted to validate the effectiveness and efficiency of the proposed model and solution method, as well as the effect of small workload differential among drivers on the total workload.

Y. Liu et al. in the paper entitled “Research on the Scheduling Problem of Movie Scenes” consider the factors affecting the cost of movie scenes shooting in the real world and construct an integer linear programming model for the scheduling of the movie sense with an objective of minimizing the total cost. The constructed model is solved by a Tabu search based method (TSBM) and a particle swarm

optimization based method (PSOBM). Both of these methods are verified to be suitable for solving small-scale problems while the former is shown to be more efficient in solving the large-scale problem according to the experimental results.

The paper entitled “Simulation Optimization of Discrete Logistics Processes: A Case Study on Logistics of an E-Commerce Enterprise in Shanghai” by X. Xu et al. introduces the simulation of the logistics distribution process of an E-commerce enterprise in Shanghai with AnyLogic Software to optimize the targeted system from three aspects, including routes selection, warehouses quantity, and warehouses layout. The results of this study can provide valuable references for practical logistics of similar E-commerce enterprise.

Y. Chen et al. in the paper entitled “Pricing Decisions on Reward-Based Crowdfunding with Bayesian Review System Facing Strategic Consumers” extend the research on the optimal pricing decision with review system for the reward-based crowdfunding. Firstly, a Bayesian analysis is established to construct consumers’ belief update process in presence of review system. Secondly, the strategy without the review system is taken as a benchmark to explore the impacts of review system under preannounced pricing and responsive pricing. It is found, through the equilibrium analysis, that the review system has positive impact on the creator under responsive pricing policy and the fraction of favorable review has a large effect on the profit of preannounced pricing.

The paper entitled “Research on Supply Chain Coordination Based on Block Chain Technology and Customer Random Demand” by Y. Li et al. focuses on supply chain coordination under the combined effects of block chain technology and random demand. Firstly, both a decentralized and a centralized supply chain decision model are built in a single-cycle newsvendor random demand situation. Then, through revenue sharing contract the study designs a brand-new supply chain coordination model which is Del trust, decentralized, and traded anonymously. According to the numerical comparative analysis on the optimal decision and supply chain coordination, it is found that the whole supply chain revenue can achieve and even exceed the performance level of the centralized supply chain with effectively expanding sales market and reducing supply chain risk. When the retail price is stable and supply chain is coordinated with revenue sharing mechanism, decentralized supply chain can achieve minimum optimal revenue. Coordination results have effect on short-term revenues of block chain members only.

The paper by L. Ma et al. entitled “Adopting a QCA Approach to Investigating the Risks Involved in Megaprojects from Auditing Perspective” introduces the microscopic empirical analysis on twenty-two typical cases by adopting the quality comparative analysis (QCA) from the auditing perspective. The results reveal that there is complex multiple concurrent causation among eight conditions, and the configuration of those conditions can be divided into six types, three of which, namely, the project management risk, preliminary and construction risk, and tendering and contract management related risk, are almost eighty percent. The analysis further reveals that megaproject risks in China are caused by complicated and changeable combination

conditions, providing a new breakthrough for the researchers and practitioners to control the megaproject risks from a more systematic way.

The paper entitled “A Metaheuristic Algorithm to Transporter Scheduling for Assembly Blocks in a Shipyard considering Precedence and Cooperating Constraints” by N.-R. Tao et al. considers an optimization transporter scheduling problem for assembly blocks in shipyards with an objective of minimizing logistics time, which includes empty travel time of transporters and waiting time and delay time of block tasks. A mathematical model is constructed by considering the time windows of ship blocks, carrying capacity of transporters, and precedence relationships of tasks and is solved by a metaheuristic algorithm based on the hybrid topological graph, genetic algorithm, and Tabu search. The results show the efficiency and effectiveness of the proposed algorithm compared to the optimal results in small-size instances and several strategies in large-scale instances.

The paper entitled “Distance-Based Congestion Pricing with Day-to-Day Dynamic Traffic Flow Evolution Process” by Q. Cheng et al. describes the distance-based congestion pricing in a network considering the day-to-day dynamic traffic flow evolution process with a mini-max regret model to optimize the worst condition among the whole planning period and ameliorate severe traffic congestions in some bad days. Firstly, a piecewise linear function is adopted to formulate the nonlinear distance toll, which can be encapsulated to a day-to-day dynamics context. Then, a logit-type Markov adaptive learning model is proposed to depict commuters’ day-to-day route choice behaviors. Finally, a robust optimization model which minimizes the maximum total travel cost among the whole planning horizon is formulated and a modified artificial bee colony algorithm is developed for the robust optimization model.

S. Qin et al. in the paper entitled “Applying Big Data Analytics to Monitor Tourist Flow for the Scenic Area Operation Management” use the big data technology and Call Detail Record (CDR) data with the mobile phone real-time location information to monitor the tourist flow and analyze the travel behavior of tourists in scenic areas. By collecting CDR data and implementing a modeling analysis of the data to simultaneously reflect the distribution of tourist hot spots in Beijing, tourist locations, tourist origins, tourist movements, resident information, and other data, the results provide big data support for alleviating traffic pressure at tourist attractions and tourist routes in the city and rationally allocating traffic resources. The analysis shows that the big data analysis method based on the CDR data of mobile phones can provide real-time information about tourist behaviors in a timely and effective manner. This information can be applied for the operation management of scenic areas and can provide real-time big data support for “smart tourism”.

Y. Peng et al. in the paper entitled “An Improved Genetic Algorithm Based Robust Approach for Stochastic Dynamic Facility Layout Problem” deal with stochastic dynamic facility layout problem under demand uncertainty in terms of material flow between facilities with the consideration of transport device management. An improved adaptive

genetic algorithm with population initialization strategy is developed to reduce the search space and improve the solving efficiency. The effectiveness of the proposed algorithm is verified by comparing it with particle swarm optimization (PSO) algorithm, and the experimental results show the good performance of the robust layout compared to the expected layout.

The paper by H. Fei and C. Zhang entitled “Optimizing the Composite Cost Involved in Road Motor-Transporting Trucks by Taking into Account Traffic Condition” focuses on the planning of road motor-transporting services by taking into account road traffic condition, especially for urban areas with an objective of minimizing the composite cost, including both the economic cost related to the driver cost and fuel consumption, and the social cost related to the vehicle emissions. The dynamic road traffic condition is imitated dynamically with a discretization technique. A metaheuristic is applied with data collected from a dense district in a huge city. Experimental results show that the proposed approach can always converge quickly to the best solution and the solution with minimal composite cost can always dominate the other solutions with classic route optimization goals.

The paper entitled “Optimal Scheme for Process Quality and Cost Control by Integrating a Continuous Sampling Plan and the Process Yield Index” by C. Li et al. presents the development of an optimal scheme for process quality and cost control to monitor the process cost and improve the process quality by taking into account four parameters: clearance number, inspecting fraction, sample size, and critical value. The Continuous Sampling Plan and the Process Yield Index are integrated to improve the output of the scheme and a case study is illustrated to validate the effectiveness and practicality of the proposed scheme.

P. Zhang and G. Liu in the paper entitled “Data-Driven Recovery Potential Analysis and Modeling for Batteries Recovery Operations in Electric Bicycle Industry” estimate the annual waste quantity of lead-acid batteries used in electric bicycles in 2000-2022 using the “market supply A model” and the “Stanford Model”, respectively. Based on the proportion of raw materials contained in lead-acid batteries and the proportion between reclaimed and discarded lead-acid batteries, the authors estimate the recovery potential of such batteries in 2000-2022. The research data and results can help decision-makers make more effective and more accurate management measures and policies.

The paper entitled “Implementation Flexibility of Multi-period Rail Line Design with Consideration of Uncertainties in Population Distribution” by K. Zhang et al. introduces the investigation results related to the implementation flexibility of multiperiod rail line design in a linear monocentric city. Three alternatives (fast-tracking, deferring, and do-nothing-alternative (DNA) of a candidate rail line project) are examined, based on an in-depth uncertainties analysis of the demand side for this candidate rail line project. Conditions for the three alternatives of fast-tracking, deferring, and DNA are analytically explored and an illustrative example is given to demonstrate the application of the proposed models. Insightful findings are reported on the interrelationship between the rail line length and spatial and temporal

correlation of population distribution as well as the implication of the correlation in practice. Sensitivity analyses are carried out in several scenarios in another numerical example to show the proposed conditions of the three alternatives.

The paper by H. Hu et al. entitled “Optimization of Vehicle Routing with Pickup Based on Multibatch Production” suggests a mixed-integer programming model for the joint optimization of multibatch production and vehicle routing problems involving a pickup to reduce both transportation and inventory cost. The targeted problem is solved within two stages. In the first stage, an integrated algorithm, combining the Clarke-Wright (CW) algorithm and the Record to Record (RTR) travel algorithm, is developed to solve vehicle routing problem. In the second stage, the particle swarm optimization (PSO) algorithm is proposed to allocate vehicles to each production batch. Multiple sets of numerical experiments are performed to validate the effectiveness of the proposed model and the performance efficiency of this two-stage hybrid algorithm.

Q. Wang et al. in the paper entitled “Integrated Optimization on Assortment Packing and Collaborative Shipping for Fashion Clothing” construct an integrated optimization problem that combines fashion clothing assortment packing with collaborative shipping simultaneously. A simplified model is then derived from the original model and solved by a commercial programming solver. Numerical results show that the proposed model is beneficial to the fashion clothing assortment packing and collaborative shipping planning.

The paper entitled “Optimal Strategies for Manufacturers with the Reference Effect under Carbon Emissions-Sensitive Random Demand” by B. Zhang et al. proposes the optimal strategies for a newsvendor system with joint reference effect, carbon emissions-sensitive random demand, and strategic customers’ behavior, with the purpose of determining the selling price, production quantity, and carbon emissions under exogenous and endogenous price cases, respectively. How the loss aversions affect the newsvendor’s decisions is explored and the study concludes that the newsvendor has a uniquely optimal policy. Furthermore, the results show that the influence of the reference effect makes the final decisions deviate from the optimal solutions of the classical model and that the loss aversions have a great impact on the newsvendor’s decisions.

The paper by C. Ma et al. entitled “A Multiobjective Route Robust Optimization Model and Algorithm for Hazmat Transportation” introduces a multiobjective robust optimization model for the routing optimization problem of hazardous materials transportation in uncertain environment by applying the Bertsimas-Sim robust optimization theory. A fuzzy C-means clustering-particle swarm optimization (FCMC-PSO) algorithm is designed, where the FCMC algorithm is used to cluster the demand points and the PSO algorithm with the adaptive archives grid is applied to calculate the robust optimization route of hazmat transportation. The numerical results verify the effectiveness of the proposed method and this study can provide basic theory support for hazmat transportation safeguarding.

J. Liu et al. in the paper entitled “Hydrological Layered Dyalysis Research on Supply Chain Financial Risk Prediction

under Big Data Scenario” construct a supply chain financial risk prediction method under big data by drawing on the risk management theory in economics and the distributed hydrological model in hydrology. First, a “hydrological database” is built for the risk analysis of supply chain financing under big data. Second, the risk identification models of “water circle model”, “surface runoff model”, and “underground runoff model” are applied to carry on the risk prediction from the overall level (water circle). Finally, the supply chain financial risk analysis from breadth level (surface runoff) and depth level (underground runoff) is performed. The results of this study can enrich the research on risk management of supply chain finance and provide feasible and effective risk prediction methods and suggestions for financial institutions.

H. Hu et al., in the paper entitled “A Metaheuristic Method for the Task Assignment Problem in Continuous-Casting Production”, focus on the task assignment problem in the downstream stage within the given information resulting from the upstream stage involved in the steelmaking and continuous-casting process in integrated iron and steel enterprises. A nonlinear mixed-integer programming model is constructed, with an objective of minimizing the total tardiness within the resource constraints and time windows constraints for the tasks, and is solved by an improved solution algorithm based on particle swarm optimization.

The paper by Y. Wei et al. entitled “Inventory and Production Dynamics in a Discrete-Time Vendor-Managed Inventory Supply Chain System” analyzes production and order dynamics in the context of a discrete-time VMI supply chain system composed of one retailer and one manufacturer. The authors firstly derive the lower bound and upper bound on the range of inventory fluctuations for the retailer under unknown demand and prove that the production fluctuations can be interestingly smoothed and stabilized independently of the delivery frequency of the manufacturer used to satisfy the retailer’s demand, even if the retailer subsystem is unstable. The sufficient and necessary stability condition for the whole supply chain system is obtained, and the bullwhip effect under unknown demand is further explored based on a transfer function model with the purpose of disclosing the influences of parameters on production fluctuations. All the theoretical results derived with respect to inventory and production fluctuations are validated by simulation experiments.

The paper entitled “Dynamic Strategies on Firm Production and Platform Advertisement in Crowdfunding considering Investor’s Perception” by Y. Ji et al. introduces a dynamic decision model for film investment by taking into account the effects of information about product quality and platform advertisement on the investor’s perception. Firstly, investment desire and reference price of the investor are introduced in two dynamic settings to describe investor’s perception. Then, the optimal decisions about the product quality and platform advertisement are formulated under two circumstances: the sponsor and the platform make decisions independently and they cooperate as a system. Finally, the influences of reference price and cost-sharing ratio on the optimal results are compared and the data simulation experiment verifies the necessity of the study.

J. Feng and B. Liu, in the paper entitled “Goodwill and System Dynamics Modeling for Film Investment Decision by Interactive Efforts”, develop a goodwill model and system dynamic (SD) model to optimize the film investment decision considering advertising, film-making, and power of stars. The results show that advertising has great impact on absorbing moviegoers’ attention, investing in film-making should be emphasized when film quality has a great impact on the movie’s reputation and audience’s viewing decision, and the film producer should pay more attention to the higher cost-performance stars who have more reasonable remuneration, better acting skills, and bigger box-office guarantee. It can also be concluded that rational audience contribute more than fans to a movie’s box-office, bankable stars contribute more than high-profile stars to a movie’s returns, and the film series yields higher profits than new theme movies although the cost of investment is the same.

The paper by D. Huang et al. entitled “An Incentive Dynamic Programming Method for the Optimization of Scholarship Assignment” proposes an incentive method inspired by dynamic programming to find the optimal scholarship assignment scheme with the highest equity considering both the practical constraints and the equity requirement. In order to assign the scholarship avoiding time- and energy-consuming application processes conducted by students, feasible assignment schemes are generated by iteratively solving a series of knapsack subproblems based on dynamic programming and adjusting the monetary value of a unit score, and then the optimal solution will be screened out by applying the Gini coefficient for quantifying the equity of each feasible scheme. The numerical results indicate that the proposed method is an efficient tool to assign scholarships to students with consideration of the equity.

The paper entitled “Optimizing Price of Credit Default Swaps for Dynamic Project System of Public-Private Partnership” by M. Wu et al. introduces a method for optimizing the price of credit default swaps (CDS) for the dynamic PPP system. This study investigates the credit risk measurement of PPP project financing and the pricing of risk mitigation instruments which are widely used in the case of immature markets in the early stage of China’s PPP development. Based on the credit risk measurement theory of the corporate and debt ratings, this paper considers the differences in various credit enhancement methods in the equity-like debt agreement and determines the credit rating of the equity-like debt in PPP projects. Some optimization methods are also proposed to derive the probability of default, so as to determine the price of the credit risk mitigation instrument of CDS which is based on the equity-like debt.

The paper entitled “Complexity Analysis of Dynamic Cooperative Game Models for Supply Chain with the Remanufactured Products” by J. Chang and L. Zhao suggests coupling dynamics of the forward supply chain of Stackelberg game model constructed for the supply chain within remanufacturing production system composed of one manufacturer and one retailer, where the manufacturer is responsible for the production of both new products and remanufactured products. After performing experiments designed by taking into account some dynamic phenomena such as bifurcation

and chaos, the authors conclude that the equilibrium of the system can lose stability via flip bifurcation or Neimark-Sacker bifurcation and that time-delayed feedback control is appropriate for stabilizing the chaotic behaviors of the system.

We believe that the special issue will provide useful references for the researchers and practitioners working in the discrete optimization for dynamic systems of operations management in data-driven society.

## Conflicts of Interest

The editors declare that they have no conflicts of interest regarding the publication of the special issue.

## Acknowledgments

The guest editorial team would like to express gratitude to all the authors for their interest in selecting this special issue as a venue for disseminating their scholarly work. The editors also wish to thank the anonymous reviewers for their careful reading of the manuscripts submitted to this special issue and their many insightful comments and suggestions.

*Lu Zhen  
Shuaian Wang  
Xiaobo Qu  
Xinchang Wang*

## Research Article

# Probability Mechanism Based Particle Swarm Optimization Algorithm and Its Application in Resource-Constrained Project Scheduling Problems

Shuai Li , Zhicong Zhang , Xiaohui Yan , and Liangwei Zhang

*Department of Industrial Engineering, Dongguan University of Technology, Songshan Lake District, Dongguan 523808, Guangdong Province, China*

Correspondence should be addressed to Shuai Li; [lishuai@dgut.edu.cn](mailto:lishuai@dgut.edu.cn)

Received 1 November 2018; Accepted 15 April 2019; Published 2 May 2019

Academic Editor: Florentino Borondo

Copyright © 2019 Shuai 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 this paper, a new probability mechanism based particle swarm optimization (PMPSO) algorithm is proposed to solve combinatorial optimization problems. Based on the idea of traditional PSO, the algorithm generates new particles based on the optimal particles in the population and the historical optimal particles in the individual changes. In our algorithm, new particles are generated by a specially designed probability selection mechanism. We adjust the probability of each child element in the new particle generation based on the difference between the best particles and the elements of each particle. To this end, we redefine the speed, position, and arithmetic symbols in the PMPSO algorithm. To test the performance of PMPSO, we used PMPSO to solve resource-constrained project scheduling problems. Experimental results validated the efficacy of the algorithm.

## 1. Introduction

Particle Swarm Optimization (PSO) is an evolutionary computational technique proposed by Kennedy and Eberhart [1] in 1995 for continuous optimization problems. Since the PSO algorithm only needs to adjust a small number of parameters and it is easy to implement, theories related to PSO algorithm has been greatly developed in the last two decades [2]. In the literature, there are many successful PSO applications for single- or multiobjective optimization [3, 4].

Although PSO was developed for continuous optimization problems initially, some work focused on its discrete version (discrete PSO and DPSO). Kennedy and Eberhart proposed a discrete binary version of particle swarm algorithm in 1997 [5], where the trajectories of the particles are changes in the probability that a coordinate will take on a zero or one value. Clerc [6] gave out a brief outline of the PSO method for solving Traveling Salesman Problems (TSP). Hendtlass used the PSO algorithm to solve small-size TSP problems and improved its performance by adding a memory capacity to each particle [7]. Pang et al. [8] used

fuzzy matrixes to represent the position and velocity of the particles in the PSO and the operators in the original PSO formulas were redefined. Wang et al. [9] redefined the PSO operators by introducing the concepts of “Swap operator” and “Swap sequence”. They solved the TSP problems in another way. We refer to all the DPSO algorithms above as actual DPSO algorithms, since they are from the original PSO. These studies have chosen some of the benchmark questions in the Traveling Salesman’s Problems Library (TSPLIB), where the city size does not exceed 52.

Some researchers combined the DPSO with other algorithms to obtain better results or enlarge the problems scale. Zhong et al. introduced a new parameter named mutation factor to the DPSO for solving the TSP [10], and the particle is not a permutation of numbers but a set of edges. Yuan et al. [11] proposed a novel algorithm based on particle optimization algorithm (PSO) and Chaotic Optimization Algorithm (COA) to solve TSP problems. Shi et al. used an uncertain searching strategy and a crossover eliminated technique to design a novel particle swarm optimization algorithm for the traveling salesman problem [12]. Machado

et al. [13] presented a new hybrid model, based on particle swarm optimization, genetic algorithms, and fast local search, for solving traveling salesman problem. Fang et al. [14] combined particle swarm optimization with simulated annealing for TSP. Except traveling salesman problem, other problems were taken into account by many scholars. Anghinol and Paolucci [15] presented a new discrete particle swarm optimization approach for the single-machine total weighted tardiness scheduling problem. Jia et al. [16] improved particle swarm optimization (PSO) algorithm with rank-priority-based representation employing a double justification skill for solving the resource-constrained project scheduling problem (RCPSP). Although most of the above hybrid algorithms have achieved good results, it is difficult to say that the discretization of PSO is successful because they all combine other algorithms.

For problems such as RCPSP and TSP, an excellent solution corresponds to a good arrangement consisting of good elements. In order to obtain these excellent elements, we propose a new DPSO algorithm PMPSO (probability-based particle swarm optimization algorithm). PMPSO is effective in jumping out of local optimum, computationally efficient, and easy to implement. We use this algorithm to solve the RCPSP problems, and the results proved the efficacy of the method.

## 2. Probability Mechanism Based Particle Swarm Optimization (PMPSO) Algorithm

In the PSO algorithm, a particle swarm is composed of a specific number of particles. At each iteration, all the particles move in the  $N$ -dimensional problem space with a certain velocity to find the global optima. The position of particle  $i$  ( $i = 1 \dots N$ ) is composed of  $D$  vectors denoted  $x_{id}$ , where  $x_{id}^k$  denotes vector at iteration  $k$ . The velocity of particle  $i$  is denoted as of a  $D$ -dimensional vectors  $v_{id}$ , where  $v_{id}^k$  denotes the velocity vector at iteration  $k$ . The velocity and position of each particle is adjusted according the following formulas:

$$v_{id}^{k+1} = wv_{id}^k + c_1 r_1 (P_{id}^k - x_{id}^k) + c_2 r_2 (P_{gd}^k - x_{id}^k) \quad (1)$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \quad (2)$$

where  $P_{id}^k$  is the local best position that the  $i$ -th particle had reached at iteration  $k$  and  $P_{gd}^k$  is the global best position that all the particles had reached at iteration  $k$ .  $r_1$  and  $r_2$  are random numbers between 0 and 1.  $w$  is the inertia weight.  $c_1$  and  $c_2$  are cognitive confidence coefficients, which are all constant numbers.

**2.1. The Representation of Particles.** When solving continuous optimization problems, the viable solution is usually directly viewed as a particle  $x_{id}$  which can be updated according to (1) and (2). However, in discrete problems, the update operation of a viable solution is mostly meaningless unless we can represent a viable solution in a feasible way. Therefore, when the PSO algorithm is applied to the discrete domain, the representation of the particle becomes a key issue. In other

words, the key point is to establish a link between particles and the solution of the problem. Most discrete problems can be regarded as sequencing problems. For example, the scheduling problem to a project comprised of five activities  $\{1, 2, 3, 4, 5\}$  can be considered as a sequencing problem to the five elements. Usually, the solution is in a form of permutation, e.g.,  $(2, 3, 1, 4, 5)$ . A permutation is formed by some fundamental elements, for example, the permutation of  $(1, 2, 3, 4, 5)$  is formed by the fundamental elements of  $(1, 2)$ ,  $(2, 3)$ ,  $(3, 4)$ , and  $(4, 5)$ . We call each fundamental element as subpermutation. A superior permutation must be consisted of some reasonably good subpermutations. To find these good subpermutations according to (1), we describe the permutation by a two-dimensional array called adjacency matrix, in which the elements correspond to the subpermutations are set to 1, the other elements are set to 0. Here are two examples of adjacency matrixes representing permutations:  $(1, 2, 3, 4, 5)$  can be represented by

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (3)$$

and  $(2, 3, 1, 4, 5)$  can be described as

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (4)$$

In the adjacency matrix, when the element  $a_{i,j}=1$ , it states that the number "1" is in front of the number "j" and they are adjacent in the permutation. In our algorithm, the position of a particle is a  $(0, 1)$  adjacency matrix, which corresponds to a permutation combination.

**2.2. Notation and Definitions of Operators.** The operators in (1) and (2) are redefined because the position and velocity are represented in the form of a matrix in discrete problems.

$$V_{id}^{k+1} = wV_{id}^k \oplus c_1 \otimes R_1 \times (P_{id}^k \ominus X_{id}^k) \oplus c_2 \otimes R_2 \times (P_{gd}^k \ominus X_{id}^k) \quad (5)$$

$$PX_{id}^k = f_p (X_{id}^k) \quad (6)$$

$$PX_{id}^{k+1} = PX_{id}^k \oplus V_{id}^{k+1} \quad (7)$$

$$X_{id}^{k+1} = f_s (PX_{id}^{k+1}) \quad (8)$$

where  $X_{id}^k$  denotes the position of a particle at iteration  $k$  and  $V_{id}^k$  denotes the transition probability velocity vector at iteration  $k$ .  $P_{id}^k$  is the local best position that the  $i$ -th particle

had reached at iteration  $k$ , and  $P_{gd}^k$  is the global best position that all the particles had reached at iteration  $k$ .  $X_{id}^k$ ,  $P_{id}^k$  and  $P_{gd}^k$  are all (0,1) adjacency matrixes corresponding to their permutations.  $R_1$  and  $R_2$  are matrixes with the same dimensions as  $X_{id}$ , whose elements are random numbers between 0 and 1.  $w$  is called inertia weight and its value is between 0 and 1.  $c_1$  and  $c_2$  are cognitive confidence coefficients, which are constant numbers. The symbol “ $\ominus$ ” denotes subtraction between matrixes and the negative elements in the subtraction result are set to 0. The symbol “ $\oplus$ ” denotes addition between matrixes and all elements greater than 1 in the addition result are set to 1. The symbol “ $\times$ ” denotes element-wise multiplication between matrixes. Suppose  $A$  and  $B$  are two matrixes with the same dimension, then the result of  $A \times B$  is the same dimension matrix, where the elements are the products of corresponding elements in  $A$  and  $B$ . The symbol “ $\otimes$ ” is used to denote the modified multiplication. Let  $c$  be a real number, then  $c \otimes A$  means all the elements of the matrix  $A$  are multiplied by  $c$ , and the elements with a value greater than 1 are set to 1.

The reason why one permutation combination is distinguished from another is its subpermutation. Therefore, a good permutation combination must contain subpermutations that are good and not available in other permutations. Based on the above ideas, we have designed a probability coefficient selection method. The basic idea of the method is that new permutations are generated based on probability selection, and those good subpermutations should have high probabilities of being selected. We try to find these reasonably good subpermutations through the PSO algorithm and make them have higher probabilities of being selected when generating new permutations. According to (5), the transition probability velocity vector  $V_{id}^{k+1}$  is combined by three parts:  $wV_{id}^k$ ,  $c_1 \otimes R_1 \times (P_{id}^k \ominus X_{id}^k)$ , and  $c_2 \otimes R_2 \times (P_{gd}^k \ominus X_{id}^k)$ .  $wV_{id}^k$  is used to govern the extent to which the old velocity  $V_{id}^k$  determines the new velocity  $V_{id}^{k+1}$ . Conventionally,  $w$  is assigned a value of 0.9. Moreover,  $c_1$  and  $c_2$  are the learning factors used to control the effects of the local experience and global experience on the new velocity, respectively. Conventionally, they are set to 1.  $P_{id}^k \ominus X_{id}^k$  is used to find the unique subpermutations in the local best position  $P_{id}^k$  but not in the individual  $X_{id}^k$ .  $P_{gd}^k \ominus X_{id}^k$  is used to find the unique subpermutations in the global best position  $P_{gd}^k$  but not in the individual  $X_{id}^k$ .

In the following, we give an example to illustrate the operators in (2), (5), (6), (7), and (8). Let  $X_{id}^k$  be the matrix of

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (9)$$

representing the permutation (1, 2, 3, 4, 5), and  $P_{id}^k$  be matrix of

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (10)$$

representing the permutation (2, 3, 1, 4, 5). Then  $P_{id}^k \ominus X_{id}^k$  is

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (11)$$

and this result denotes that the subpermutations (3, 1) and (1, 4) are unique for  $P_{id}^k$  comparing with  $X_{id}^k$ . If the random matrix  $R_1$  is

$$\begin{pmatrix} 0.2 & 0.5 & 0.8 & 0.7 & 0.2 \\ 0.4 & 0.6 & 0.9 & 0.3 & 0.7 \\ 0.3 & 0.1 & 0.8 & 0.9 & 0.2 \\ 0.4 & 0.5 & 0.9 & 0.4 & \mathbf{0.6} \\ 0.9 & 0.2 & 0.1 & 0.8 & 0.7 \end{pmatrix}, \quad (12)$$

then the 0-1 matrix is transformed to

$$\begin{pmatrix} 0 & 0 & 0 & 0.7 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (13)$$

by  $c_1 \otimes R_1 \times (P_{id}^k \ominus X_{id}^k)$  and in  $V_{id}^{k+1}$ . The elements in the matrix corresponding to the subpermutations (3,1) and (1,4) are highlighted, which increases the probabilities that they are selected in the new permutation.

New particles are obtained by selection based on probability coefficients. According the main idea of PSO, we need to update the selection probability coefficients considering the current position and velocity. Therefore, we must convert  $X_{id}^k$  to the same form as probability coefficients matrix. In (6),  $PX_{id}^k$  is defined as probability matrix corresponding to  $X_{id}^k$ , which stems from  $X_{id}^k$  according to  $f_p$  function. The function of  $f_p$  can be freely constructed according to the following rules: ensure the elements in  $PX_{id}^k$  corresponding to the elements of 1 in  $X_{id}^k$  have big probabilities and the value of

the elements in  $PX_{id}^k$  that cannot be used to make a selection is 0. Here we give an example. Let  $X_{id}^k$  be

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (14)$$

Firstly we can generate a random matrix and we suppose that it is

$$\begin{pmatrix} 0.4897 & \mathbf{0.4984} & 0.8234 & 0.1456 & 0.8564 \\ 0.7567 & 0.2345 & \mathbf{0.2551} & 0.5987 & 0.3334 \\ 0.7234 & 0.3213 & 0.9876 & \mathbf{0.1386} & 0.8876 \\ 0.2223 & 0.6456 & 0.7986 & 0.1234 & \mathbf{0.2435} \\ 0.1987 & 0.2234 & 0.9876 & 0.3235 & 0.9856 \end{pmatrix}. \quad (15)$$

Secondly, we divide all the elements by 1000, except for the elements corresponding to the elements of one in  $X_{id}^k$ . Finally, we set the diagonal elements to be 0, because the diagonal elements in  $X_{id}^k$  cannot equal one and the corresponding elements in  $PX_{id}^k$  cannot be used for the selection operation. Then,  $PX_{id}^k$  is

$$\begin{pmatrix} 0 & \mathbf{0.4984} & 0.0008 & 0.0001 & 0.0008 \\ 0.0007 & 0 & \mathbf{0.2551} & 0.0005 & 0.0003 \\ 0.0007 & 0.0003 & 0 & \mathbf{0.1386} & 0.0008 \\ 0.0002 & 0.0006 & 0.0007 & 0 & \mathbf{0.2435} \\ 0.0001 & 0.0002 & 0.0009 & 0.0003 & 0 \end{pmatrix}. \quad (16)$$

In (7), the selection probability coefficients matrix is updated considering the current position and the velocity. Suppose  $V_{id}^{k+1}$  is

$$\begin{pmatrix} 0.0 & 0.0 & 0.0 & 0.7 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.8 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{pmatrix}, \quad (17)$$

then

$$\begin{aligned} PX_{id}^{k+1} &= PX_{id}^k \oplus V_{id}^{k+1} \\ &= \begin{pmatrix} 0 & 0.4984 & 0.0008 & \mathbf{0.7010} & 0.0008 \\ 0.0007 & 0 & 0.2551 & 0.0005 & 0.0003 \\ \mathbf{0.8007} & 0.0003 & 0 & 0.1386 & 0.0008 \\ 0.0002 & 0.0006 & 0.0007 & 0 & 0.2435 \\ 0.0001 & 0.0002 & 0.0009 & 0.0003 & 0 \end{pmatrix}. \end{aligned} \quad (18)$$

In (8),  $f_s(PX_{id}^{k+1})$  was defined as probability selecting operation, which selects the right elements and set them to one in  $X_{id}^{k+1}$  according to the given probability coefficients in  $PX_{id}^{k+1}$ . The selection operation must obey the rule that if an element in  $PX_{id}^{k+1}$  has a larger probability coefficient value, the element with the same index in  $X_{id}^{k+1}$  has a greater probability of being selected and set to 1. Then we get the new adjacency matrix and the new permutation is obtained by decoding it.

### 2.3. Selection Mechanism Based on Probability Coefficients.

A solution formed by means of probability coefficients selection can also break the local optimum and without losing the opportunity to find better solutions. We illustrate the selection process as follows.

*Process 1* (process of probability coefficients selection).

*Step 1.* Set all the elements of adjacency matrix  $X_{id}^{k+1}$  to zero.

*Step 2.* Select a row according to the row number of the row of the maximum element in the probability matrix of  $PX_{id}^{k+1}$ .

*Step 3.* The corresponding element in  $PX_{id}^{k+1}$  is used as the probability selection coefficient, and one element in the row is randomly selected accordingly. The selection rule is that elements with larger probability coefficients have a greater probability of being selected.

*Step 4.* Let the selected element be equal to 1, and set the corresponding so-called contradictory elements in  $PX_{id}^{k+1}$  to zero. There are three types of contradictory elements: (1) elements with the same row number as the selected element; (2) elements with the same column number as the selected element; (3) elements which can lead to incompletely permutation.

*Step 5.* If the stopping criterion is not satisfied, go to Step 2.

In the above example in Section 2.2,  $PX_{id}^{k+1}$  is

$$\begin{pmatrix} 0 & 0.4984 & 0.0008 & 0.7010 & 0.0008 \\ 0.0007 & 0 & 0.2551 & 0.0005 & 0.0003 \\ 0.8007 & 0.0003 & 0 & 0.1386 & 0.0008 \\ 0.0002 & 0.0006 & 0.0007 & 0 & 0.2435 \\ 0.0001 & 0.0002 & 0.0009 & 0.0003 & 0 \end{pmatrix}. \quad (19)$$

We then use the same example to explain the implementation of Process 1. Firstly,  $X_{id}^{k+1}$  is set to

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (20)$$

According to Process 1, we select the row of (0.8007,0.0003, 0.0000,0.1386,0.0008) as coefficients, because the biggest

value of the matrix is in this row. We randomly select a column according to the coefficients. According to the selection rule, each element could be selected except the third element, which coefficients equal to zero. However, the elements which have bigger probability coefficients have bigger probabilities to be selected. Hence, we assume that the first element was selected. Then in this example,  $X_{id}^{k+1}$  is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (21)$$

and  $PX_{id}^{k+1}$  is

$$\begin{pmatrix} 0 & 0.4984 & 0 & 0.7010 & 0.0008 \\ 0 & 0 & 0.2551 & 0.0005 & 0.0003 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0.0006 & 0.0007 & 0 & 0.2435 \\ 0 & 0.0002 & 0.0009 & 0.0003 & 0 \end{pmatrix}. \quad (22)$$

The element  $PX_{id}^{k+1}(1,3)$  is set to 0 because the third situation in Step 4. Next, in this example, the first row (0, 0.4984, 0, 0.7010, 0.0008) is selected and it is used as coefficients to randomly select a row. Assume that the forth column is selected. Then  $X_{id}^{k+1}$  is

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (23)$$

and  $PX_{id}^{k+1}$  is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2551 & 0 & 0.0003 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0.0006 & 0 & 0 & 0.2435 \\ 0 & 0.0002 & 0.0009 & 0 & 0 \end{pmatrix}. \quad (24)$$

$PX_{id}^{k+1}(4,3)$  is set to 0, because if the corresponding element in  $X_{id}^{k+1}$  is selected to 1, the circle permutation (3, 1, 4, 3) would exist, which lead to a incompletely permutation. Next, the second row (0, 0, 0.2551, 0, 0.0003) is selected and it is used as

weights to randomly select a column. Assume that the third column is selected. Then  $X_{id}^{k+1}$  is

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (25)$$

and  $PX_{id}^{k+1}$  is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.2435 \\ 0 & 0.0002 & 0 & 0 & 0 \end{pmatrix}. \quad (26)$$

Next, the fourth row (0, 0, 0, 0, 0.2435) is selected and it is used as coefficients to randomly select a column. The fifth column is selected. Then  $X_{id}^{k+1}$  is

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (27)$$

and the whole 0-1 adjacency matrix is generated and the new permutation is (2, 3, 1, 4, 5).

*2.4. Description of the Algorithm.* PMPSO algorithm employs the probability coefficients selection mechanism to improve the exploration of the solution space and it assures the convergence through the PSO evolutionary mechanism. We summarize the procedures of the proposed algorithm as follows.

*Algorithm 2* (procedures of PMPSO algorithm).

*Step 1.* Initialization

*Step 1.1.* Set the number of the particles and the max number of iterations.

*Step 1.2.* Each particle gets a random position matrix  $X_i$  corresponding to a permutation.

*Step 1.3.* Decode  $X_i$  and calculate the fitness of the position.

*Step 1.4.* Initialize the local best  $P_i = X_i^0$  and the global best  $P_g$  is the best of all the  $P_i$ .

*Step 2.* Calculate the position and velocity of each particle.

*Step 2.1.* Calculate new velocities according to (5).

*Step 2.2.* Calculate  $PX_{id}^k$  according to (6).

*Step 2.3.* Update the  $PX_{id}^{k+1}$  according to (7).

*Step 2.4.* Calculate  $X_{id}^{k+1}$  according to (8).

*Step 3.* Calculate the fitness of each position. If the fitness of the new position is better than that of the local best position of the particle, update the best position with the new position. If the fitness of the local best position of some particles is better than that of the global best position, then update the global best with the local best position.

*Step 4.* If the stopping criterion is not satisfied, go to Step 2.

*Step 5.* Output the best solution and the fitness.

### 3. Problem Statement

Resource-constrained project scheduling problems (RCPSp) were investigated in many situations such as building construction, facility and equipment maintenance, and new product development. Due to their academic appeal and wide application, many researchers have always shown interest in these issues. The problem of RCPSp can be stated as follows. A project consists of a set  $N = \{0, 1, \dots, J+1\}$  of activities, where activities 0 and  $J+1$  are dummy activities that represent the events of the start and finish time of the project respectively. The RCPSp is based on the following assumptions: (1) the durations of activities composing the project are deterministic and known; (2) all activities must obey their precedence constraints; (3) resources can be available in limited amounts and renewable from period to period; (4) activities cannot be interrupted when in progress; (5) the goal of optimization is to minimize project duration. The duration of activity  $j$  is denoted  $d_j$ , and resource type  $k$  requested by activity  $j$  is denoted  $r_{jk}$ . For the dummy activities, for any resource type  $k$ , we have  $d_0 = d_{J+1} = 0$ , and  $r_{0,k} = r_{J+1,k} = 0$ . The availability of each unit of resource type  $k$  is  $R_k$  ( $k = 1 \dots K$ ). A project schedule can be represented by the start time of each activity denoted  $(s_0, s_1, \dots, s_{J+1})$  with  $s_0 = 0$  and the finish time of each activity denoted by  $(f_0, f_1, \dots, f_{J+1})$  with  $f_0 = 0$ . Note that  $f_{J+1} = s_{J+1}$  and  $f_j = s_j + d_j$  for all  $j$ .  $I_t$  is the set of tasks being executed in the time period of  $t$ . A schedule is feasible if it satisfies precedence relationships and all these constraints during each execution time period. The mathematical model of the problem can be formulated as follows:

$$\min f_{J+1} \quad (28)$$

$$\text{s.t. } f_j - f_i \geq d_j, \quad j = 1, 2, \dots, J+1; \quad \forall i \in P_j \quad (29)$$

$$\sum_{j \in I_t} r_{jk} \leq R_k, \quad k = 1, 2, \dots, K; \quad t = 0, 1, 2, \dots \quad (30)$$

Equation (28) is the objective function that minimizes the completion time of the last activity of the project, that is, minimizes the span of the project. Equation (29) means each task's start time must be subject to the finish time of its preceding activities. It also can be expressed as activity

$j$  cannot start until all of the predecessor activities have finished, where  $P_j$  stands for the set of preceding activities of activity  $j$ . Equation (30) ensure that the total number of used resources of type  $k$  cannot exceed its available maximum, denoted as  $R_k$ .

It was proved that the optimization of the RCPSp is NP-hard in the strong sense [17]. Mingozzi et al. [18] developed an exact method for solving the problems with sizes up to 30 activities. Donrndorf et al. [19] also reported satisfactory results using a branch-and-bound algorithm for the problems with 30 activities. For large-scale RCPSp instances, most researchers are employing metaheuristics algorithm to solve them, including tabu algorithms [20], genetic algorithms [21, 22], ant colony optimization [23], simulated annealing [24], and variable neighborhood search [25]. However, the heuristic methods based on heuristic rules are often variable effectiveness on different cases and may be trapped within local optima [21–23]. Since the advent of the PSO algorithm, there have been many attempts to apply PSO to solve RCPSp. In contrast to other heuristic methods, PSO has the advantages of computational efficiency, great capability of escaping local optima, and easy implementation [26, 27]. Zhang et al. [28] took the priorities of the scheduling activities as particles to develop a DPSO algorithm for solving the RCPSp and the size of the problem investigated was 25. Xie et al. [29] combined the chaos algorithm with PSO to solve the typical multiple resources-constrained project scheduling problem. Ni et al. [30] proposed the hybrid particle swarm optimization based on differential evolution for solving the RCPSp with fifteen activities. In the following, we use the proposed PMPSO algorithm to solve the RCPSp problems.

### 4. RCPSp Solution Based on PMPSO Algorithm

To solve a RCPSp problem using PMPSO, we must have the information of the project, such as the duration of each activity, each resource required for each activity, the availability of each resource, and precedence relationships of the project activities. Then we can use PMPSO algorithm to solve the problem. The flow chart of the solution is given as Figure 1.

*4.1. Encoding and Decoding of the Particles.* We directly take the set of activity number sequence as the source of particles, which has the advantage of simplicity and effectiveness. Then the permutation is transformed to a 0-1 adjacency matrix, which is the Encoding of particle. The 0-1 adjacency matrix is considered to be the position of the particle. According to the 0-1 adjacency matrix, we can derive the permutation, which called decoding of the particle.

*4.2. Generation of Feasible Sequences.* A feasible sequence  $S = \{s_1, s_2, \dots, s_n\}$  is a sequence in which the predecessor activity of any activity in the sequence must be arranged to the left of the activity. A feasible sequence is the basis of solving the RCPSp, regardless the algorithm. Researchers used various methods to get the feasible sequence. Montoya-Torres et al.

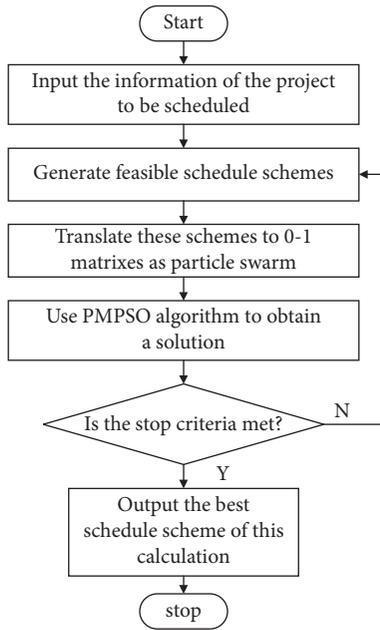


FIGURE 1: Flow chart of the solution.

[31] transformed sequences to feasible sequences using SSGS (Serial Schedule Generation Schemes) when they solved the RCPSP problem with genetic algorithm. Peng and Wang [32] adopted the priority list in the genetic representation combining the SGS to generate the feasible sequence. Xie et al. [33] allotted the priorities to the activities and then applied the bubble method to generate the feasible sequence. Zhou et al. [33] eliminated the improper sequence with a logic judgment procedure to get the feasible sequence. In all these articles, the feasible scheduling sequences of the activities were acquired indirectly. In order to improve the validity of the calculation, we have designed a method called Modulation Weighting Mechanism for Generating feasible scheduling sequences (MWMG), which can directly obtain a feasible solution. First, the precedence relationships of the activities must be given as a 0-1 matrix  $R$ , in which the element  $r_{ij}=1$  means that the activity  $i$  is a predecessor of activity  $j$ . Secondly, an activity must be scheduled after all his predecessor activities. If the activity  $i$  as the predecessor of the activity  $j$  is scheduled, we can deactivate the coupling relationship between the activities  $i$  and  $j$  by setting the element  $r_{ij}=0$ . Finally, when we generate the scheduling scheme according the probability weight matrix, the weight matrix should be modulated by the relation 0-1 matrix  $R$  to ensure that all the predecessor relations are obeyed. Because the activity “0” is dummy activity which represents the event of the project start, we begin to schedule from the activity “0”. The generation process of the scheduling sequence is illustrated as follows.

*Process 3* (process of generating feasible sequence).

*Step 1.* A 0-1 matrix representing the scheduling scheme is initialized by setting all elements to zero. Set the first column of the weight matrix to zero to let the project begin from

activity “0”. Set the first row of the relationship matrix to zero to release subsequent activities as activity “0” so that these activities can be scheduled.

*Step 2.* The value of the element in the modulation vector depends on the value of the above row vector. When the element value of the row vector is zero, the corresponding element value in the modulation vector is 1; otherwise, the element value is 0, where the element in the modulation vector equals to 1 means that the activity denoted by the column number can be scheduled, because all of its predecessor activities have been scheduled.

*Step 3.* A row is selected from the weight matrix as a weight vector and is modulated by multiplying each element in the weight vector by its corresponding element in the modulation vector. At the beginning, the selected row is the first row.

*Step 4.* Select the column number according the modulated weight vector randomly. Then, the corresponding element value in the 0-1 matrix of the scheduling scheme is set to 1 according to the row number and the column number.

*Step 5.* Renew the weight matrix according the rules in Section 2.3.

*Step 6.* Update the row number by the column number.

*Step 7.* Set the row of the relationship matrix to zero to release the subsequent activities of the activity represented as the column number so that they can be scheduled.

*Step 8.* If the stopping criterion is not satisfied, go to Step 3.

*Step 9.* Decode the 0-1 matrix of the scheduling scheme to get the feasible sequence.

Through the above generation mechanism, not only the feasible scheme can be directly obtained, but also the excellent subpermutations have better chance to enhance their probabilities.

*4.3. Generation of a Scheduling Scheme.* We adopt the method of Serial Scheduling Generation Scheme (SSGS) to obtain a feasible scheduling scheme. According to the sequence, we schedule the activities successively. The following gives the generation process of the scheduling scheme.

*Process 4* (process of generating a scheduling scheme).

*Step 1.* Select an activity  $i$  from the feasible sequence in order, which duration is  $d_i$ , and calculate the early start time of the activity.

*Step 2.* Schedule the start time of the activity as early as possible, denoted  $t_{is}$ , and calculate its finish time denoted  $t_{if}$ , where  $t_{if} = t_{is} + d_i$ .

*Step 3.* Calculate the availability of the resources required for the activity  $i$  during the time period in which the activity is performed.

*Step 4.* If the availability of the resources is satisfied, go to Step 5; otherwise, let  $t_{is} = t_{is} + 1$  and go to Step 2.

*Step 5.* If there are still unscheduled activities, go to Step 1; otherwise go to Step 6.

*Step 6.* Output the result.

*4.4. Total Process of Solving a RCPSP Problem Using PMPSO Algorithm.* According to Processes 1–4 we summarize the general process of solving a RCPSP problem using PMPSO algorithm as follows.

*Process 5* (process of solving the RCPSP problem).

*Step 1.* Input the information of the project to be scheduled and initialize the parameters of the PMPSO algorithm.

*Step 2.* Initialize the probability coefficients matrixes according to the rules mentioned in Process 1.

*Step 3.* Generate 0-1 matrixes of the scheduling scheme using the method in Process 3.

*Step 4.* Decode each 0-1 matrix to get a scheduling scheme and calculate the fitness function value according to Process 4.

*Step 5.* Update the best position of the swarm and the best positions of each particle experienced.

*Step 6.* Update the probability coefficients matrixes according to (8).

*Step 7.* If the stopping criterion of the PMPSO algorithm is not satisfied, go to Step 3; otherwise go to Step 8.

*Step 8.* Output the result.

## 5. Computational Experiments

We conduct the experiments for PMPSO in this section. The test set comes from J30, J60, J90 and J120 of the PSPLIB, provided on the website (<http://www.om-db.wi.tum.de/psplib/>) [34]. J30, J60 and J90 all contain 480 instances, while J120 contains 600 instances. Each instance set has four resource types. The currently best known solution in PSPLIB as reported on June 17, 2015. For the J30 set, these solutions are all optimal. The test environment is as follows: CPU: INTEL-2.4G, RAM: 4GB DDR 333, HD: 250G 7200 rpm, OS: Win 10.

We set  $w$  to 0.98, both  $c_1$  and  $c_2$  to 1, and the maximal number of iteration to 5000. Then we randomly select 10 instances from each test set and run the scheduler 100 times on each problem instance. Table 1 shows the statistical results of these solutions.

From Table 1, we observed that PMPSO can get the best known solution for all of the instances randomly selected from J30, and the maximum deviation of the mean of all solutions from the optimal value is only 0.17%. For other test

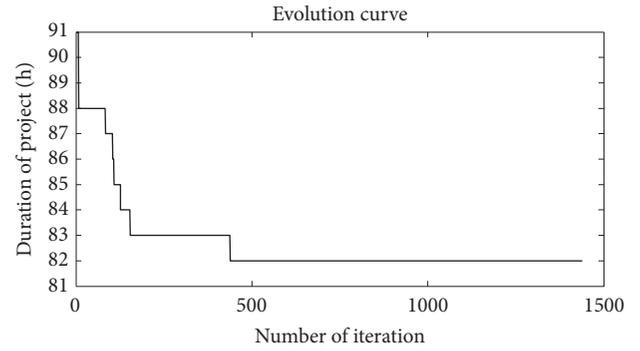


FIGURE 2: Evolution curve of the instance of J305-2.

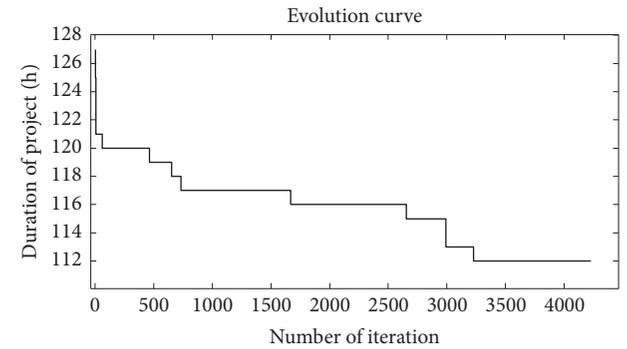


FIGURE 3: Evolution curve of the instance of J1201-5.

sets, PMPSO can get most of the best known solutions, and the max deviation from the best known solution is 1.69%. As we can see from Table 1, the best known solutions of these instances were found by various researchers using various methods. In this paper, we use PMPSO to get some of the best known solutions and some solutions deviating from the best known solutions slightly. We also execute the scheduling program for each instance of each problem set. Table 2 shows the results of relative deviations from the currently best known solution.

Tables 1 and 2 summarize the experimental results of PMPSO. We observed that the larger the problem, the greater the deviation from the best value or the current best known solution, and as the size of the problem increases, the average number of best solutions obtained decreases. For most instances, we can get the best known solutions. Moreover, as it is shown in the Table 2, using the algorithm we can get a very high percentage of hits for the problem size within 30. We note that the best known solutions in PSPLIB are found by different people using kinds of algorithms. Therefore, PMPSO is an algorithm that adapts to problems of various scales and has unique advantages especially for small-scale problems.

To observe the performance of PMPSO, we also illustrate the evolution curves of the instances of J305-2 and J1201-5, which are showed in Figures 2 and 3, respectively.

Figure 2 shows that evolution curve drops sharply and then converges at the best value. This shows that the PMPSO algorithm has the advantage of fast convergence, which is

TABLE 1: Solutions of the selected instances.

Instance	Results in PSPLIB		Results by PMPSO			
	Best known Solution	Best solution	Worst solution	Average	Deviation (%)	
J30	J301-1	43	43	/	43	0
	J302-4	43	43	/	43	0
	J303-9	65	65	/	65	0
	J305-10	70	70	72	70.04	0.06
	J308-1	44	44	/	44	0
	J3010-4	58	58	/	58	0
	J3015-2	47	47	48	47	0
	J3026-6	53	53	/	53	0
	J3030-1	47	47	49	47.08	0.17
	J3040-8	57	57	/	57	0
J60	J602-5	54	54	56	54.15	0.27
	J603-9	67	67	68	67.05	0.07
	J605-1	76	76	77	76.51	0.67
	J608-8	66	66	69	66.44	0.66
	J6020-4	86	86	88	86.13	0.15
	J6025-2	98	98	100	98.02	0.02
	J6026-9	65	65	67	65.07	0.1
	J6030-2	70	70	71	70.05	0.07
	J6040-3	70	70	71	70.42	0.6
	J6048-8	88	88	89	88.01	0.01
J90	J903-5	75	75	76	75.51	0.68
	J904-4	92	92	94	93.03	1.12
	J908-9	97	97	99	98.04	1.07
	J9010-1	77	77	78	77.11	0.14
	J9018-8	92	92	93	92.34	0.37
	J9021-4	106	106	107	106.5	0.47
	J9022-2	85	85	86	85.23	0.27
	J9036-9	102	103	105	103.12	1.09
	J9038-1	85	85	87	85.82	0.91
	J9041-2	168	168	170	169.12	0.67
J120	J1203-5	84	85	87	85.21	1.44
	J1204-6	90	91	92	91.5	1.67
	J1208-9	94	94	95	95.55	1.65
	J12010-1	111	112	114	112.88	1.69
	J12020-8	107	107	109	108.56	1.46
	J12021-4	135	135	137	136.24	0.92
	J12032-2	131	131	131	132.23	0.94
	J12046-9	166	167	168	167.88	1.13
	J12048-2	113	113	114	114.55	1.37
	J12058-5	120	120	122	121.98	1.65

TABLE 2: Experimental results by PMPSO on each instance.

	Average deviation (%)	Best deviation (%)	Worst deviation (%)	no. best/no. instance	Percentage of hits (%)
J30	0.01	0	0.20	465/480*	96.88
J60	0.22	0	0.67	385/480*	80.21
J90	0.31	0	1.1	370/480	77.08
J120	0.98	0	2.5	251/600	41.83

\*hit means finding the best known solution

the most important advantage of the original PSO algorithm. It also demonstrates the outstanding performance of solving small-scale problems.

For large-scale problems, PMPSO is doing well too, which can be observed from Figure 3. For the instance of J1201-5, which has 120 activities, PMPSO can get its best known solution at 3200 iterations.

In Figure 3, before the curve levels off to the minimal value, it experiences several platforms. This shows that the PMPSO algorithm can effectively avoid converging on local optimum and reduce solution time greatly.

## 6. Conclusions

This paper proposes a probability mechanism based particle swarm optimization algorithm. The design of the algorithm includes the following features: a special decoding scheme, a mechanism for finding good subpermutations, a new particle swarm iteration formula, and a selection mechanism based on probability coefficients. When applying the PMPSO algorithm to discrete optimization problems, we only need to convert the solution to a 0-1 matrix form, and then use our mutated PSO formula. The experimental results showed that state-of-the-art solutions can be obtained using the PMPSO algorithm to solve some benchmark problems. In our future work, we will establish a better probability selection mechanism to improve the selection process. We will implement dynamic adjustment of parameters to improve the performance of PMPSO. In addition, we can combine PMPSO with other algorithms, such as local search, metaheuristic, to get better results. In the experiments of this paper, PMPSO is capable of solving the RCPSP efficiently. Although these experiments are made based on PSPLIB, PMPSO should perform equally well for the other types of sequencing problems, such as production scheduling, route planning, and the assignment problem.

## Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest in the article.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 71201026; Grant no. 71801045). This work was partially supported by Research Start-Up Funds of DGUT (GC300502-46) and the Youth Innovative Talent Project (2017KQNCX191) from the Department of Education of Guangdong Province, China.

## References

- [1] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, vol. 4, pp. 1942–1948, IEEE, Perth, Australia, 1995.
- [2] J. F. Chang, S. C. Chu, J. F. Roddick, and J. S. Pan, "A parallel particle swarm optimization algorithm with communication strategies," *Journal of Information Science and Engineering*, vol. 4, no. 21, pp. 809–818, 2005.
- [3] R. C. Eberhart and Y. Shi, "Comparing inertia weights and constriction factors in particle swarm optimization," in *Proceedings of the 2000 Congress on Evolutionary Computation*, vol. 1, pp. 84–88, IEEE, La Jolla, Calif, USA, July 2000.
- [4] C. A. C. Coello, G. T. Pulido, and M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 256–279, 2004.
- [5] J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, vol. 5, pp. 4104–4109, IEEE, Orlando, Fla, USA, October 1997.
- [6] M. Clerc, "Discrete particle swarm optimization, illustrated by the traveling salesman problem," in *New Optimization Techniques in Engineering*, B. V. Babu and G. C. Onwubolu, Eds., vol. 141 of *Studies in Fuzziness and Soft Computing*, pp. 219–239, Springer, Berlin, Germany, 2004.
- [7] T. Hendtlass, "Preserving diversity in particle swarm optimization," in *Lecture Notes in Computer Science*, vol. 2718, pp. 4104–4108, Springer, 2003.
- [8] W. Pang, K. Wang, C. Zhou, and L. Dong, "Fuzzy discrete particle swarm optimization for solving traveling salesman problem," in *Proceedings of the Fourth International Conference on Computer and Information Technology*, pp. 796–800, Wuhan, China, September 2004.
- [9] K.-P. Wang, L. Huang, C.-G. Zhou, and W. Pang, "Particle swarm optimization for traveling salesman problem," in *Proceedings of the International Conference on Machine Learning and Cybernetics*, pp. 1583–1585, IEEE, November 2003.
- [10] W. Zhong, J. Zhang, and W. Che, "A novel discrete particle swarm optimization to solve traveling salesman problem," in *Proceedings of the 2007 IEEE Congress on Evolutionary Computation*, pp. 3283–3287, Singapore, September 2007.
- [11] Z. Yuan, L. Yang, Y. Wu, L. Liao, and G. Li, "Chaotic particle swarm optimization algorithm for traveling salesman problem," in *Proceedings of the IEEE International Conference on Automation and Logistics*, pp. 1121–1124, Jinan, China, 2007.
- [12] X. H. Shi, Y. C. Liang, H. P. Lee, C. Lu, and Q. X. Wang, "Particle swarm optimization based algorithms for TSP and generalized TSP," *Information Processing Letters*, vol. 103, pp. 169–176, 2007.
- [13] T. R. Machado and H. S. Lopes, "A hybrid particle swarm optimization model for the traveling salesman problem," in *Natural Computing Algorithms*, H. Ribeiro, R. F. Albrecht, and A. Dobnikar, Eds., pp. 255–258, Springer, 2005.
- [14] L. Fang, P. Chen, and S. Liu, "Particle swarm optimization with simulated annealing for TSP," in *Proceedings of the 6th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases*, C. A. Long, V. M. Mladenov, and Z. Bojkovic, Eds., pp. 206–210, Corfu Island, Greece, February 2007.
- [15] D. Anghinolfi and M. Paolucci, "A new discrete particle swarm optimization approach for the single-machine total weighted tardiness scheduling problem with sequence-dependent setup

- times," *European Journal of Operational Research*, vol. 193, no. 1, pp. 73–85, 2009.
- [16] Q. Jia and Y. Seo, "An improved particle swarm optimization for the resource-constrained project scheduling problem," *The International Journal of Advanced Manufacturing Technology*, vol. 67, no. 9, pp. 2627–2638, 2013.
- [17] J. Blazewicz, J. K. Lenstra, and A. H. Rinnooy Kan, "Scheduling subject to resource constraints: classification and complexity," *Discrete Applied Mathematics*, vol. 5, no. 1, pp. 11–24, 1983.
- [18] A. Mingozzi, V. Maniezzo, S. Ricciardelli, and L. Bianco, "An exact algorithm for the resource-constrained project scheduling problem based on a new mathematical formulation," *Management Science*, vol. 44, no. 5, pp. 714–729, 1998.
- [19] U. Dorndorf, E. Pesch, and T. Phan-Huy, "A branch-and-bound algorithm for the resource constrained project scheduling problem," *Mathematical Methods of Operations Research*, vol. 52, no. 3, pp. 413–439, 2000.
- [20] T. Baar, P. Brucker, and S. Knust, "Tabu search algorithms and lower bounds for the resource-constrained project scheduling problem," in *Meta-Heuristics: Advances and Trends in Local Search Paradigms for Optimization*, S. Voss, S. Martello, I. Osman, and C. Roucarion, Eds., pp. 1–18, Kluwer, Boston, Mass, USA, 1998.
- [21] S. Hartmann, "A competitive genetic algorithm for resource-constrained project scheduling," *Naval Research Logistics (NRL)*, vol. 45, no. 7, pp. 733–750, 1998.
- [22] S. Hartmann, "A self-adapting genetic algorithm for project scheduling under resource constraints," *Naval Research Logistics (NRL)*, vol. 49, no. 5, pp. 433–448, 2002.
- [23] D. Merkle, M. Middendorf, and H. Schmeck, "Ant colony optimization for resource-constrained project scheduling," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 4, pp. 333–346, 2002.
- [24] K. Bouleimen and H. Lecocq, "A new efficient simulated annealing algorithm for the resource-constrained project scheduling problem and its multiple mode version," *European Journal of Operational Research*, vol. 149, no. 2, pp. 268–281, 2003.
- [25] K. Fleszar and K. S. Hindi, "Solving the resource-constrained project scheduling problem by a variable neighbourhood search," *European Journal of Operational Research*, vol. 155, no. 2, pp. 402–413, 2004.
- [26] R. C. Eberhart and Y. Shi, "Comparison between genetic algorithms and particle swarm optimization," in *Evolutionary Programming VII*, vol. 1447 of *Lecture Notes in Computer Science*, pp. 611–616, Springer, Berlin, Germany, 1998.
- [27] J. Robinson, S. Sinton, and Y. R. Samii, "Particle swarm, genetic algorithm, and their hybrids: optimization of a profiled corrugated horn antenna," in *Proceedings of the IEEE International Symposium in Antennas and Propagation Society International symposium and URSI*, vol. 1, pp. 168–175, IEEE, San Antonio, Tex, USA, 2002.
- [28] H. Zhang, H. Li, and C. M. Tam, "Particle swarm optimization for resource-constrained project scheduling," *International Journal of Project Management*, vol. 24, no. 1, pp. 83–92, 2006.
- [29] X. Yang, Y. Chun-ming, C. Jun-lan, and Z. Rong, "Resource-constrained project scheduling based on chaos particle swarm optimization," *Industrial Engineering Journal*, vol. 15, no. 3, pp. 57–61, 2012.
- [30] L. Ni, C. Duan, and C.-l. Jia, "Hybrid particle swarm optimization algorithm based on differential evolution for project scheduling problems," *Application Research of Computers*, vol. 28, no. 4, 2011.
- [31] J. R. Montoya-Torres, E. Gutierrez-Franco, and C. Pirachicán-Mayorga, "Project scheduling with limited resources using a genetic algorithm," *International Journal of Project Management*, vol. 28, no. 6, pp. 619–628, 2010.
- [32] P. Wuliang and W. Chengen, "A multi-mode resource-constrained discrete time-cost tradeoff problem and its genetic algorithm based solution," *International Journal of Project Management*, vol. 27, no. 6, pp. 600–609, 2009.
- [33] X. Yang, Y. Chun-ming, C. Jun-lan, and Z. Rong, "Resource-constrained project scheduling based on chaos particle swarm optimization," *Industrial Engineering Journal*, vol. 15, no. 3, 2012.
- [34] D. Debels, B. de Reyck, R. Leus, and M. Vanhoucke, "A hybrid scatter search/electromagnetism meta-heuristic for project scheduling," *European Journal of Operational Research*, vol. 169, no. 2, pp. 638–653, 2006.

## Research Article

# Omni-Channel Product Distribution Network Design by Using the Improved Particle Swarm Optimization Algorithm

Si Zhang , Heying Zhu, Xia Li, and Yan Wang

School of Management, Shanghai University, Shanghai 200444, China

Correspondence should be addressed to Si Zhang; zhangsi817@sina.com

Received 30 November 2018; Accepted 13 February 2019; Published 4 April 2019

Guest Editor: Xiaobo Qu

Copyright © 2019 Si 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.

Distribution network under Omni-Channel integration contains many levels. There are one or more dealers at each level which forms a many-to-many distribution network. Consumers purchase a wide variety of products and their demands are uncertain, which constitutes a complex demand network and increases the complexity of the supply chain network. This paper focuses on the integrated optimization of supply chain distribution network and demand network and constructs the joint randomization planning model of location and routing. The goal is to minimize the total costs of the supply chain network under uncertain customer demands. Based on the traditional particle swarm optimization (PSO), this study introduces the collaborative idea to reduce the coding dimension, improves the boundary processing strategy, and adopts the mutation operator to expand the search space. A case study of distribution under Omni-Channel integration in a large enterprise was done. The validity of the model and the effectiveness of the proposed method were verified by numerical experiments.

## 1. Introduction

Over the last two decades, the product transportation of the retail industry was largely driven by the Internet Technology development. More and more retailers not only have the traditional offline stores but also start to have the online stores to satisfy the demands of customers from the Internet. As the emergence of new online channels, retailers begin to transform from traditionally storefront-based retailers to multi-channel retailers ([1–3]) and consider the operation management on two channels. This gave birth to the concept of Omni-Channel integration.

Omni-Channel integration is the deep integration of online services, offline experience, and modern logistics in new retail era. It makes the supply chain more complex. The operation goal of Omni-Channel integration supply chain is to meet customer demands with the lowest costs. Saghiri et al. [4] proposed that it is a challenging task to analyze the distribution strategy of Omni-Channel integration. The supply chain usually contains many levels and each level has many sellers, which leads to the complexity of the supply chain network. Besides, the geographical locations of the sellers are spread all over the country. These factors have

a significant impact on transportation costs. In addition, customer demands are often uncertain which will increase the costs of supply chain. Therefore the optimization for the supply chain with Omni-Channel is a challenge task and it is necessary to have some efficient approaches to help retailers reduce their costs. For example, Su'ning, one of the largest retailers with Omni-Channel supply chain in China, adopted a distribution method which is city-centered distributed to the county and rural areas step by step and constructed a three-level logistics distribution system. This system is “National Logistics Center—Regional Logistics Center—City Logistics Center”. As of June 2017, it has covered 297 prefecture-level cities and above, with about 2079 authorized chain stores and 833 authorized service outlets. It launched “100-City Half-Day” service to require products that are delivered to the customer within half day after the order is placed, which poses greater challenges to the distribution of logistics.

Therefore, facing such a large supply chain network, building an efficient, high-quality, and low-cost distribution network is the key to improve competitiveness and sustainable development of enterprises. The main aim of this

paper is to construct a highly efficient supply chain system to minimize costs and improve service quality.

This paper is organized as follows. Section 1 introduces the background of Omni-Channel distribution. The related literature is reviewed in Section 2 from the perspective of LRP, uncertain demand, and solving methods, respectively. In Section 3, a distribution network for Omni-Channel supply chain is proposed and a model for this problem is built. To solve this model, one algorithm based on the traditional PSO is developed in Section 4 by introducing the collaborative idea to reduce the coding dimension, improving the boundary processing strategy, and introducing the mutation operator to expand the search space. Sections 5 and 6 are the numerical results and sensitivity analysis based on one case study. Section 7 concludes the whole paper.

## 2. Literature Review

This paper mainly focuses on the location and routing problem (LRP) for Omni-Channel supply chain. Early LRP issues were limited to the location of a single facility. Maranzana [5] studied the problem of supplier location with the goal of minimizing transportation costs. Subsequently, many researchers conducted further research on the systematic and networking nature of transportation and location decisions. Govindan et al. [6] introduced a two-tier supply chain with a time window based on the supply chain network distribution of perishable food, which determined the number of distribution centers and facility locations. Abdulrazik et al. [7] optimized the tertiary supply chain under the background of producing oil palm in Malaysia. In addition, from the types of product distribution in the LRP, products transportation has expanded from single-product distribution to multi-product distribution. Geoffrion and Graves [8] developed a multi-product distribution model using Mixed-Integer Programming (MIP). Tiwari et al. [9] proposed an algorithm for solving multi-stage, multi-product supply chain network design. Viswanathan and Mathur [10] studied a distribution system consisting of a warehouse multiple products and multiple retailers. Arntzen [11] et al. used a mixed integer programming to establish a multi-product and multi-stage supply chain model for large multinational companies. Although research on the LRP changes from the initial single location planning to more consideration of the systemic and networking feature of the supply chain, the current researches on supply chain network are mainly focused on two-level or three-level system, and the distribution route planning is mostly one-to-many or many-to-one delivery. For the retail industry, it usually has many supply chain levels and many types of products in the distribution. Therefore, it is necessary to study the supply chain with more levels and more products. This paper is to design a five-level supply chain network for the LRP with multi-product distribution in retail industry.

One another factor that cannot be ignored in Omni-Channel retail supply chain study is the uncertainties of customer demands. Different scholars used different methods to research uncertainty. Wang and He [12] introduced the uncertainty of demand and studied the problem of

distribution route and warehouse location under the condition of uncertain demand. Zhen et al. [13] used the scenario method to solve the problem of uncertainty in the production process of automobiles. Tapiero and Soliman [14] applied optimal control theory to solve the problem of multi-product transportation, multi-regional production, and cross-time inventory planning under demand uncertainty. Baghalian et al. [15] used multi-objective optimization to model the location and routing problem in low-carbon supply chain and explored the impact of demand uncertainty. In general, customers have great differences in the demand for different products. It is necessary to form a personalized demand network. To describe customer demands more accurately, we set each customer's uncertain demand for different kinds of products to follow different normal distributions instead of letting all customers uniformly meet a normal distribution for all products uniformly.

The LRP with multi-level, multi-product, and personalized demand is a very complex optimization problem. There are two main streams solving the large-scale optimization problem. The first stream is to use new evolutionary algorithms or add local search strategies (e.g. tabu search) to original algorithm. Zhang et al. [16] adopted tabu search algorithm to improve the accuracy and efficiency of cold chain product distribution. Javid and Azad [17] proposed a method of synchronous optimization of location and routing. Two-stage heuristic algorithm based on tabu search and simulated annealing was used to improve the solution space. Xia et al. [18] established the dual-objective programming model and designed an adaptive tabu search algorithm. The second stream of literature studies is to decompose the high-dimensional of large-scale complex problems into low-dimensional simple problems and optimize them separately. Hu et al. [19] studied the large-scale refrigerated truck scheduling problem and combined a variable neighborhood search with a particle swarm optimization to develop a two-stage decomposition algorithm. In 2009, Li and Yao [20] proposed a new co-evolutionary particle swarm optimization algorithm (CPSO), which added random grouping and adaptive weighting strategy to prove that it is effective in high-dimensional indivisible problems. The optimized 30-dimensional CPSO algorithm which has been proposed earlier can be extended to 1000-dimensional problems now. Compared with manufacture industry, distribution in retail industry usually have more supply chain levels, which results in the high dimensions of decision variables. We develop an improved PSO algorithm based on CPSO to efficiently solve the optimization problem.

Therefore, in our paper, we aim to contribute to the current studies on this area from the following three aspects. Firstly, under the Omni-Channel distribution, the supply chain network level has been increased, and a many-to-many distribution network has been formed, which makes the supply chain network distribution more in line with the actual situation. Secondly, we allow each customer's demands to follow different normal distributions for each product and form a demand network accordingly. The demand network can better predict customer demands. Thirdly, based on the traditional PSO, the algorithm introduces the collaborative

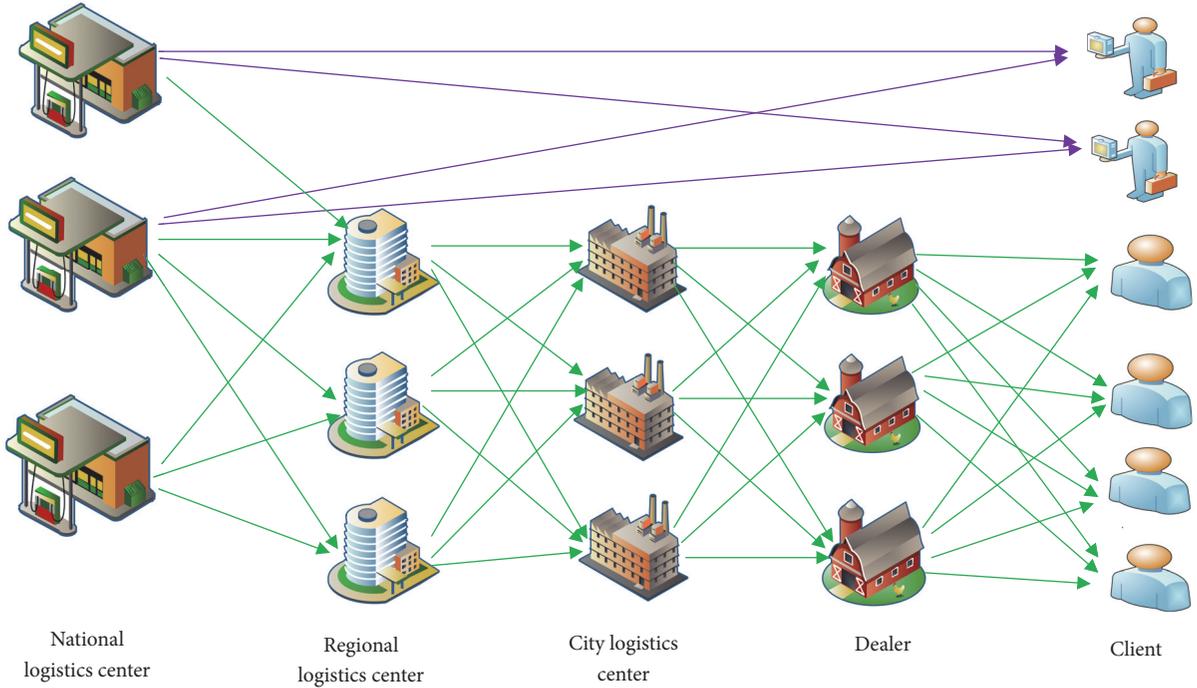


FIGURE 1: Supply chain distribution network.

idea to reduce the coding dimension, improves the boundary processing strategy, and introduces the mutation operator to expand the search space. This method improves the efficiency of the algorithm and provides a solution for solving high dimensional problems.

### 3. Model

**3.1. Model Description.** This model includes some common characteristics of distribution under Omni-Channel integration, so the conclusions of the model can be useful for many enterprises of supply-chain distribution under Omni-Channel integration. There are five levels in the entire supply chain. These levels include National Logistics Center A, Regional Logistics Center B, City Logistics Center C, Dealer D, and Customer E. Distribution network shall distribute  $m$ -type products to two types of customers. The first type of customers  $E_1$  is large customers such as companies; the demand of such customers is delivered directly by the national logistics center A. The second type of customers  $E_2$  is small supermarkets, like stores. The orders of such customers will be delivered through the regional logistics center B, the city logistics center C, and the dealer D. We can intuitively see the model in Figure 1.

According to the supply chain network, we built a mixed integer programming model to minimize the total supply chain costs. The total costs include three items: transportation costs of the products, fixed operating costs of logistics center, and carbon treatment costs which is to deal with carbon emission generated from the product transportation process.

### 3.2. The Index of the Model

- a: The number of national logistics centers
- b: The number of regional logistics centers
- c: The number of city logistics centers
- d: The number of dealers
- $e_1$ : The number of first type of customers
- $e_2$ : The number of second type of customers
- m: The number of the category of products

### 3.3. Parameter Symbols

- $G_{ae1}^m$ : The unit transportation cost of product  $m$  delivered to customer  $e_1$  through national logistics center a
- $G_{ab}^m$ : The unit transportation cost of product  $m$  delivered to regional logistics center b through national logistics center a
- $G_{bc}^m$ : The unit transportation cost of product  $m$  delivered to city logistics center c through regional logistics center b
- $G_{cd}^m$ : The unit transportation cost of product  $m$  delivered to dealer d through city logistics center c
- $G_{de2}^m$ : The unit transportation cost of product  $m$  delivered to customer  $e_2$  through dealer d
- $dis_{ae1}$ : The distance between national logistics center a and customer  $e_1$

$dis_{ab}$ : The distance between national logistics center a and regional logistics center b

$dis_{bc}$ : The distance between regional logistics center b and city logistics center c

$dis_{cd}$ : The distance between city logistics center c and dealer d

$dis_{de2}$ : The distance between dealer d and customer  $e_2$

$F_a$ : The fixed operating costs of national logistics center located at point a

$F_b$ : The fixed operating costs of regional logistics center located at point b

$F_c$ : The fixed operating costs of city logistics center located at point c

$F_d$ : The fixed operating costs of dealer located at point d

$O_a^m$ : The average inventory of product m in national logistics center located at point a

$R_{e1}^m$ : The demand of customer  $e_1$  for product m

$R_{e2}^m$ : The demand of customer  $e_2$  for product m

T: Unit carbon treatment cost

N: A large enough positive number

### 3.4. Decision Variables

$Q_{ae1}^m$ : The quantity of product m from national logistics center a to customer  $e_1$

$Q_{ab}^m$ : The quantity of product m from national logistics center a to regional logistics center b

$Q_{bc}^m$ : The quantity of product m from regional logistics center b to city logistics center c

$Q_{cd}^m$ : The quantity of product m from city logistics center c to dealer d

$Q_{de2}^m$ : The quantity of product m from dealer d to customer  $e_2$

$y_a$ : Whether the national logistics center a operates, if it operates, then  $y_a$  equals 1; otherwise the value is 0

$y_b$ : Whether the regional logistics center b operates, if it operates, then  $y_b$  equals 1; otherwise the value is 0

$y_c$ : Whether the city logistics center c operates, if it operates, then  $y_c$  equals 1; otherwise the value is 0

$y_d$ : Whether the dealer d operates, if it operates, then  $y_d$  equals 1; otherwise the value is 0

### 3.5. Data Model

$$\min f(x) \quad (1)$$

$$\begin{aligned} f(x) = & \sum_A \sum_{E1} \sum_M G_{ae1}^m Q_{ae1}^m dis_{ae1} + \sum_A \sum_B \sum_M G_{ab}^m Q_{ab}^m dis_{ab} \\ & + \sum_B \sum_C \sum_M G_{bc}^m Q_{bc}^m dis_{bc} + \sum_C \sum_D \sum_M G_{cd}^m Q_{cd}^m dis_{cd} \\ & + \sum_D \sum_{E2} \sum_M G_{de2}^m Q_{de2}^m dis_{de2} + \sum_A F_a y_a + \sum_B F_b y_b \\ & + \sum_C F_c y_c + \sum_D F_d y_d + \sum_A \sum_{E1} \sum_M T Q_{ae1}^m dis_{ae1} \\ & + \sum_A \sum_B \sum_M T Q_{ab}^m dis_{ab} + \sum_B \sum_C \sum_M T Q_{bc}^m dis_{bc} \\ & + \sum_C \sum_D \sum_M T Q_{cd}^m dis_{cd} + \sum_D \sum_{E2} \sum_M T Q_{de2}^m dis_{de2} \end{aligned} \quad (2)$$

$$\sum_A Q_{ae1}^m = R_{e1}^m \quad \forall e_1 \in E_1; \forall m \in M \quad (3)$$

$$\sum_D Q_{de2}^m = R_{e2}^m \quad \forall e_2 \in E_2; \forall m \in M \quad (4)$$

$$\sum_A Q_{ab}^m \geq \sum_C Q_{bc}^m \quad \forall b \in B; \forall m \in M \quad (5)$$

$$\sum_B Q_{bc}^m \geq \sum_D Q_{cd}^m \quad \forall c \in C; \forall m \in M \quad (6)$$

$$\sum_C Q_{cd}^m \geq \sum_{E2} Q_{de2}^m \quad \forall d \in D; \forall m \in M \quad (7)$$

$$\sum_{E1} Q_{ae1}^m + \sum_B Q_{ab}^m \leq O_a^m * y_a \quad \forall a \in A; \forall m \in M \quad (8)$$

$$\sum_A Q_{ab}^m - N \times y_b \leq 0 \quad \forall b \in B; \forall m \in M \quad (9)$$

$$\sum_B Q_{bc}^m - N \times y_c \leq 0 \quad \forall c \in C; \forall m \in M \quad (10)$$

$$\sum_C Q_{cd}^m - N \times y_d \leq 0 \quad \forall d \in D; \forall m \in M \quad (11)$$

$$\begin{aligned} y_a, y_b, y_c, y_d \in \{0, 1\} \\ \forall a \in A; \forall b \in B; \forall c \in C; \forall d \in D \end{aligned} \quad (12)$$

$$\begin{aligned} Q_{ae1}^m, Q_{ab}^m, Q_{bc}^m, Q_{cd}^m, Q_{de2}^m \geq 0 \\ \forall a \in A; \forall b \in B; \forall c \in C; \forall d \in D; \forall e_1 \in E_1; \forall e_2 \in E_2 \end{aligned} \quad (13)$$

The objective function (2) aims to minimize the total costs of the supply chain network. The constraint (3) requires that the demands of the customer  $e_1$  can only be directly delivered by the national logistics center a, and the constraint

(4) states that the product can only be provided from the dealer  $d$  to the customer  $e_2$ . Equations (5)-(7) specify that the volume of each logistics center is greater than or equal to the shipment at that point. Constraint (8) ensures that the inventory from the national logistics center  $a$  is greater than the total volume of shipped products, and the operation of the national logistics center  $a$  is represented by the 0-1 variable. Constraint (9) donates that if the regional logistics center  $b$  is not in operation, the quantity of product shipped to the regional logistics center  $b$  is zero. While constraint (10) and constraint (11) indicate that if city logistics center  $c$  and the dealer  $d$  do not operate, the volume of product shipped to the city logistics center  $c$  and the dealer  $d$  is zero. Constraint (12) and constraint (13) are the value constraints of the decision variable.

#### 4. Proposed Heuristic Method

The model of LRP presented in Section 3 aims to minimize the total costs in product transportation process. In a large-scale transportation network, the code dimension is high, making it difficult to find an optimal solution in a limited time. To solve the problem, we resort to heuristic methods, such as particle swarm optimization (PSO). PSO is a stochastic, parallel optimization algorithm that does not require the properties of the optimized function to be divisible, steerable, and continuous. In addition, to further improve the accuracy and efficiency of the algorithm, the improved PSO is proposed.

*4.1. Particle Swarm Optimization.* PSO was proposed by Kennedy and Eberhart [21] and Lian [22] pointed out that this method was a stochastic optimization method based on swarm intelligence and PSO was inspired by the social behavior of bird flocking and their means of information exchange. Due to its easy implementation and fast convergence, PSO has been successfully applied to solve nonlinear optimization problems. Ai and Kachitvichyanukul [23], Goksal et al. [24], and Alinaghian et al. [25] designed an improved random topology particle swarm optimization algorithm (RT-PSO) for time-dependent vehicle routing problems. Hu et al. [19] developed a two-stage decomposition algorithm by combining variable neighborhood search algorithm and PSO to solve the large-scale refrigerated truck scheduling problem.

In PSO, each particle represents a solution to the optimization problem. Particle initialization is generated in a random way, and each particle iteratively updates its position through personal best and global best. Each particle searches for the optimal fitness value in the  $D$ -dimensional feasible solution space. There are  $N$  particles in a group. In iteration  $t$ , the position of particle  $i$  is represented as a  $D$ -dimensional vector  $X_{id}^t = (x_{i1}^t, x_{i2}^t, \dots, x_{iD}^t)$ ,  $i = 1, 2, \dots, N$ . At the same time, the velocity of the particle  $i$  is also a  $D$ -dimensional vector, denoted as  $V_{id}^t = (v_{i1}^t, v_{i2}^t, \dots, v_{iD}^t)$ ,  $i = 1, 2, \dots, N$ . The optimal position searched by the particle  $i$  in iteration  $t$  is called the personal best, which is denoted as  $P_{best} = (p_{i1}, p_{i2}, \dots, p_{iD})$ ,  $i = 1, 2, \dots, N$ . The optimal position searched by the group in iteration  $t$  is the global best,

denoted as  $G_{best} = (p_{g1}, p_{g2}, \dots, p_{gD})$ . After finding the two optimal values, each particle updates its velocity and position according to the following formula:

$$v_{id}^{t+1} = w * v_{id}^t + c_1 * r_1 * (P_{best}_{id}^t - x_{id}^t) + c_2 * r_2 * (G_{best}_{id}^t - x_{id}^t) \quad (14)$$

$$x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1} \quad (15)$$

In the standard PSO, there are some shortcomings; for example, (1) when solving the high dimensional optimization problem, the increase of the dimension makes the search space expand into an exponential form, which is prone to premature convergence. That is, the particle group aggregates prematurely to the local optimal solution. (2) Particles may aggregate to local optimum too early in the process of gathering. If there is no escaping mechanism, the optimal solution may not be obtained. Therefore, we proposed improved strategy accordingly.

*4.2. Strategies for Improving PSO.* The improved PSO focuses on three aspects: (1) facing the ‘‘dimensional disaster’’ in large-scale problems, cooperative particle swarm optimization (CPSO) is proposed to transform large-scale complex problems into low-dimensional simple problems. (2) When particle flies out of the boundary, the boundary processing strategy is adopted to change the flight trajectory of the particles. (3) When the particles are aggregated, in order to prevent premature convergence to the local optimum, the variation factor is introduced to further explore the search space.

*4.2.1. CPSO for Dimensionality Reduction.* As the dimension of the optimization problem increases, the performance of all algorithm drops dramatically. An effective solution is the cooperative evolution strategy. Bergh and Engelbrecht [26] combined cooperative evolution strategy with PSO and proposed cooperative particle swarm optimization (CPSO). The basic idea of CPSO is using  $K$  independent particle groups to search different dimension directions in the  $D$ -dimensional search space. Each particle swarm is updated independently, and no information is shared between the particle swarms. When calculating the fitness value, the position vectors of the optimal particles in the particle swarm are combined together to form a  $D$ -dimensional vector to calculate the fitness value.

In this paper, to calculate the traffic volume between supply chain levels, the particle coding in standard PSO is  $11$ -dimensional  $X(A, B, C, D, E_1, E_2, m, A', B', C', D')$ . The first 6 dimensions are the supply chain routing between neighboring level, the 7th dimension demonstrates different products, and the last 4 dimensions represent the node operation decision. According to CPSO, decision variable is classified by supply chain levels. Decision variable can be split into five types of 3-dimensional routing decision  $Q1(A, E_1, m)$ ,  $Q2(A, B, m)$ ,  $Q3(B, C, m)$ ,  $Q4(C, D, m)$ ,  $Q5(D, E_2, m)$  and four types of 1-dimensional node operation decision  $Y1(A)$ ,  $Y2(B)$ ,  $Y3(C)$ ,  $Y4(D)$ . The process is shown

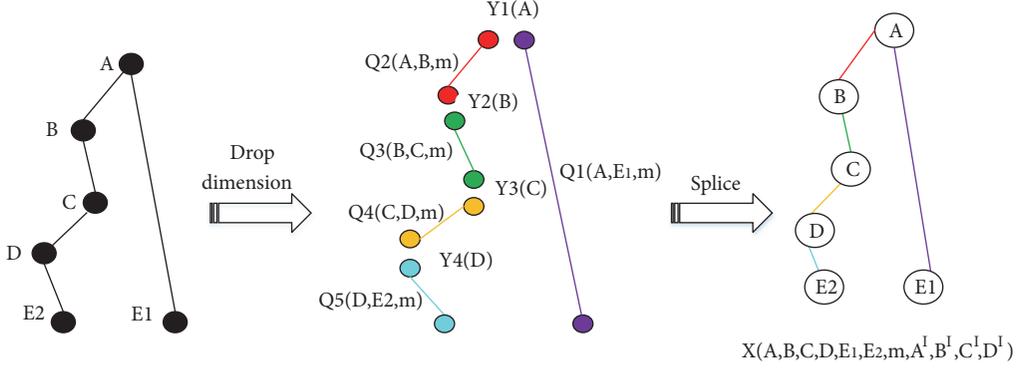


FIGURE 2: Collaborative dimension reduction.

in Figure 2. Coding dimension is reduced and the algorithm speed is improved through CPSO.

**4.2.2. Improved Boundary Processing Strategy.** In the search process, in order to improve the convergence efficiency and reduce the invalid search, the boundary conditions are usually set for the position and velocity to narrow the search range. In the traditional PSO, when a particle goes beyond the range, the general processing method is to set the particle on the boundary. That is,

$$\text{If } X_{id} < X_{min} \text{ then } X_{id} = X_{min} \quad (16)$$

$$\text{Else if } X_{id} > X_{max} \text{ then } X_{id} = X_{max} \quad (17)$$

Through this method, after several iterations, a plurality of particles will aggregate toward the boundary in a plurality of dimensions and the flight path of the particles will tend to be the same. So the efficiency of the particle group is reduced. Thus, the following improvement strategies have been proposed:

$$\begin{aligned} \text{If } X_{id} < X_{min} \\ \text{then } X_{id} = X_{min} + c_3 * r_3 * X_{min} \end{aligned} \quad (18)$$

$$\begin{aligned} \text{Else if } X_{id} > X_{max} \\ \text{then } X_{id} = X_{max} - c_3 * r_3 * X_{max} \end{aligned} \quad (19)$$

$r_3$  is a random number distributed between  $[0, 1)$ , and  $c_3$  is a constant on  $[0.01, 0.5)$ . The value of  $c_3$  will change according to the algorithm, the objective function, etc. This paper sets it to 0.05 to enhance the ability of the algorithm to jump out of the boundary region, so that the particles return to the feasible search space. After using the boundary strategy, on the one hand, it can ensure that the particles are located in the feasible search space. On the other hand, it can prevent the particles from accumulating too much to the boundary, resulting in local optimization on the boundary of the search space. It can maximize the search space and improve the quality of the solution. A comparison of the two strategies is shown in Figure 3.

**4.2.3. Mutation Operators to Expand Search Space.** When PSO appears as local convergence or global convergence, the particles will have an aggregation phenomenon; that is, the particles have the same fitness. The fitness variance is the judgment of the convergence degree of the whole group. The larger the fitness variance is, the more likely the particle swarm is in the search state. When the fitness variance becomes smaller, the particle swarm is tending to converge; when the fitness variance is 0, PSO achieves local optimization or global optimization. The formula for calculating the fitness variance is as follows:

$$\sigma^2 = \sum_{i=1}^m \left( \frac{f_i - f_{av}}{f} \right)^2 \quad (20)$$

And,  $f = \max\{1, \max_{1 \leq i \leq N} \{|f_i - f_{av}|\}\}$ , the function of  $f$  is to limit  $\sigma^2$ .

$N$  is the total number of particles in the particle group.

$f_i$  is the fitness value of particle  $i$ .

$f_{av} = (1/N) \sum_{i=1}^N f$  is average fitness function value of each particle.

If it is local optimum, the algorithm is premature. When the algorithm is premature, it will affect the accuracy of the algorithm. So,  $Gbest_d^t$  will be the personal best. The variation factor  $Pr$  is introduced, thereby changing the forward direction of the particle. So  $Gbest_d^t$  and  $Pbest_{id}^t$  will be updated:

$$\text{If } Pr > \text{rand}(d) \text{ then update particle} \quad (21)$$

The probability of mutation is relatively small;  $Pr$  generally takes a number between  $[0, 1)$ ; in this algorithm,  $Pr = 0.1$ ;  $\text{rand}(d)$  is a random number distributed between  $[0, 1)$ . The specific approach is to mutate the particles when they are within 10% of the probability of mutation. Firstly, all the particles are sorted according to the fitness value; secondly,  $m$  particles with the best fitness value ( $m$  is usually half of the total number of particles) are obtained; then, the former  $Pr * m$  particles are mutated as follows:

$$x_{id}^{t+1} = x_{id}^t * (1 + 0.5r\eta) \quad (22)$$

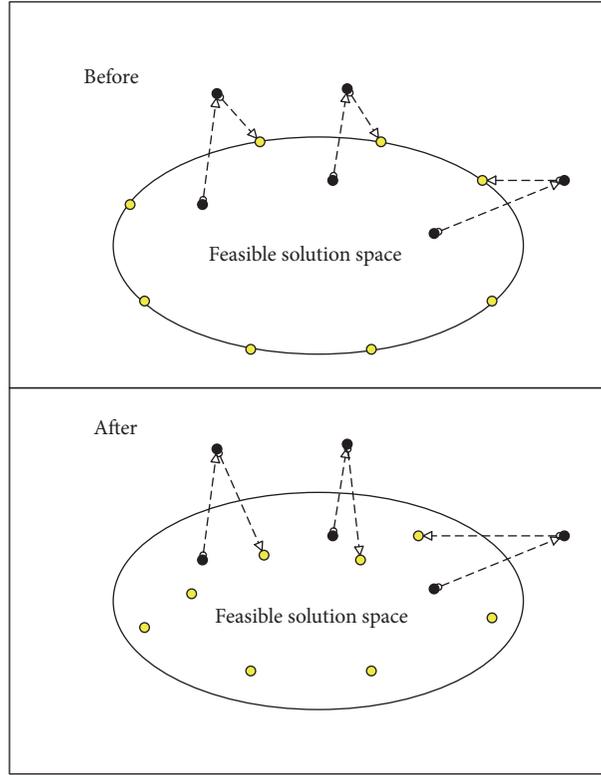


FIGURE 3: Improved boundary processing strategy.

TABLE 1: Parameters setting for the experiments.

Parameter	Distribution type	Distributions	Unit
$G_{ae1}^m, G_{ab}^m, G_{bc}^m, G_{cd}^m, G_{de2}^m$	Uniform distribution	$U(190, 250)$	RMB/ton/km
$F_a, F_b, F_c, F_d$	Uniform distribution	$U(150000, 350000)$	RMB/year
$O_a^m$	Uniform distribution	$U(150000, 200000)$	ton
$R$	Normal distribution	$N(\mu_j^p, \sigma^2)$	ton/year

where  $\eta$  is a random variable obeying the standard normal distribution  $N(0, 1)$ . The purpose of using the mutation factor is to avoid the particle group falling into local optimum and better approaching the global optimum. When the algorithm appears premature, the mutation can change the direction of the entire particle swarm, and the particle swarm reaches a new field for search. By introducing the mutation operator, the algorithm can expand the search space when the search is stagnant. The improved PSO flow is shown in Figure 4. The pseudocode of improved PSO is presented in Box 1.

### 5. Numerical Experiments

In this section, some numerical experiments are performed to verify the effectiveness of the proposed model of multi-product distribution and location problem based on uncertain demand networks. The programming language is C#, and the experiments are completed on a PC (Intel Core i7, 2.6 GHz; Memory, 8G).

**5.1. Parameter Setting.** According to the background of the distribution network under Omni-Channel integration, the parameters of this experiment are set. The location and routing process is the same as the model description in Section 3. To demonstrate the broad applicability of the model, some of the parameters of the experiment were generated based on a random probability distribution shown in Table 1.

The unit transportation cost  $G$  is uniformly generated according to different road conditions. The operating costs  $F$  of dealers for different levels are uniformly generated according to the land price and scale. The capacity  $O$  of the distribution center is uniformly distributed according to the product category and scale. This article deals with the uncertain customer demand as the following two steps.

*Step 1.* Each customer has different demands for different products. A set of scenarios with different demands are set according to the characteristics of the normal distribution function.

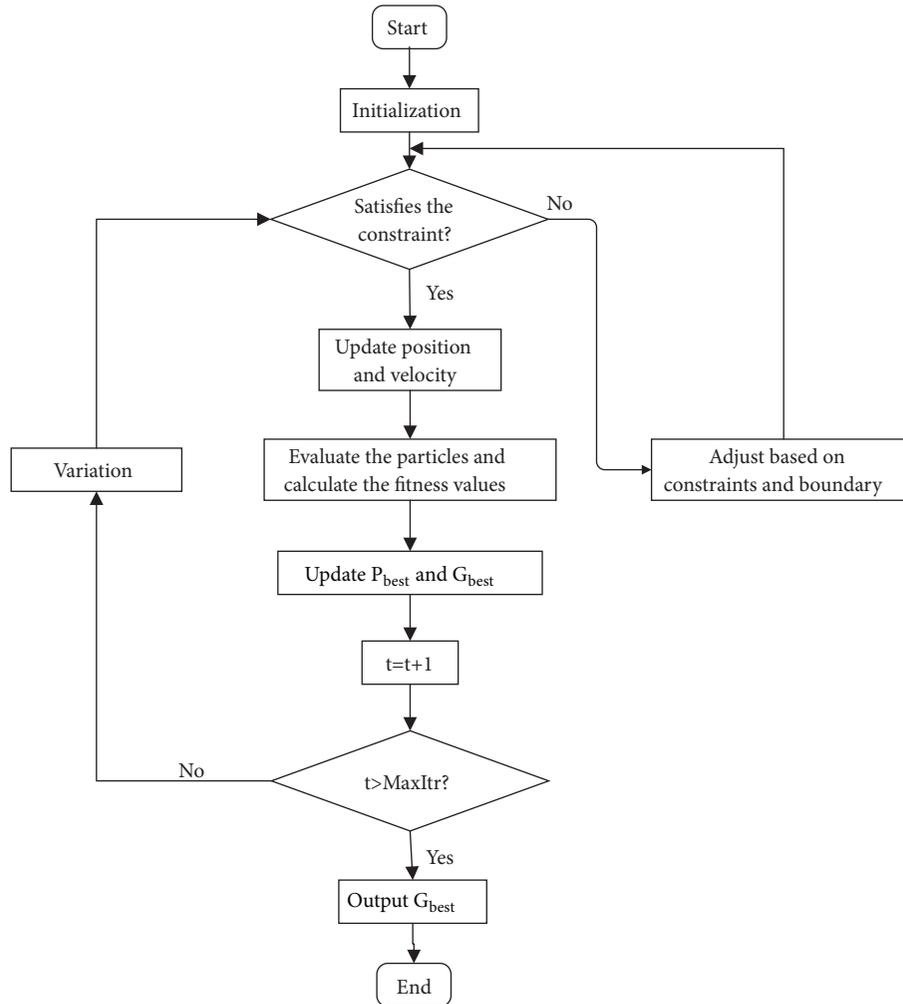


FIGURE 4: Particle swarm optimization flowchart.

*Step 2.* Let different customer demands for different products meet different normal distributions. This creates a network of demand of each customer for each product, rather than allowing all customers to meet only one normal distribution demand for each product, which could predict customer demands more accurately.

*5.2. Experimental Results.* This paper studies multi-product distribution under the uncertainty of customer demand, so the experiment is divided into four groups according to the changes of customer demand. Each group of experiments is divided into five different situations according to the number of suppliers and customers. The results in Table 2 show 5 experimental scales with different changes of demand. The objective function value and running time are calculated by the improved PSO and CPLEX solver. “1-1-2-4-10 (2+8)” represents one national logistics center, one regional logistics center, two city logistics centers, four dealers, and ten customers (two of them are first-type customers and eight of them are second-type customers).

In small-scale cases, the running time of CPLEX is less than the improved PSO. As the scale increases, the running

time of CPLEX increases significantly. For example, in the first three experiments, improved PSO and CPLEX have similar effects. When the scale increased to 85 customers, the running time of CPLEX increased quickly, while the running time of improved PSO is less increased. In general, the results of the CPLEX solver are considered to be the optimal solution. Therefore, the CPLEX solver solves the model faster than the improved PSO method on a small scale. However, when the network exceeds a certain scale, the improved PSO gets results faster than the CPLEX solver. In addition, the average deviation between the improved PSO and the optimal result is about 0.53%, which indicates that the result is accurate and the proposed algorithm is suitable for solving the model.

## 6. Case Analysis

Based on the background of the distribution of three types of products in Chengdu, China, the scale is “2-4-12-22-85(7+78)”. M Company is a typical retail enterprise and products transportation is under Omni-Channel integration.

TABLE 2: Comparison of different kinds of algorithms.

	Scale of the study (1)	CPLEX		Improved PSO		Gap [(4)-(2)]/(2) (6)
		Obj. (2)	CPU(s) (3)	Obj. (4)	CPU(s) (5)	
Change in demand =0	1-1-2-4-10 (2+8)	155118311.9	9.95	155118311.9	13.31	0.00%
	2-2-4-10-35 (4+31)	4850885922	26.8	4861980932	29.03	0.23%
	2-2-6-14-55 (5+50)	5181467988	45	5213955899	42	0.63%
	2-4-12-22-85 (7+78)	5318725918	304.1	5363403216	77.06	0.84%
	3-5-18-70-205 (10+195)	5619609011	>3600	5672995297	503	0.95%
Change in demand =0.01	1-1-2-4-10 (2+8)	155478693.3	9.95	155478693.3	11.71	0.00%
	2-2-4-10-35 (4+31)	4905668616	23.45	4913027119	28.57	0.15%
	2-2-6-14-55 (5+50)	5193458934	49	5224634879	44.54	0.60%
	2-4-12-22-85 (7+78)	5304694312	325	5350314683	83.86	0.86%
	3-5-18-70-205 (10+195)	5608465011	>3600	5666793047	524	1.04%
Change in demand =0.05	1-1-2-4-10 (2+8)	155446843.3	9.9	155446843.3	8.63	0.00%
	2-2-4-10-35 (4+31)	4902814227	23.75	4913110137	24.77	0.21%
	2-2-6-14-55 (5+50)	5220507224	45.73	5231902083	38.29	0.22%
	2-4-12-22-85 (7+78)	5342902440	341.6	5392591433	76.8	0.93%
	3-5-18-70-205 (10+195)	5621433450	>3600	5685517791	537	1.14%
Change in demand =0.1	1-1-2-4-10 (2+8)	155867650.3	9.95	155867650.3	12.71	0.00%
	2-2-4-10-35 (4+31)	4928341567	29.9	4941648089	27.94	0.27%
	2-2-6-14-55 (5+50)	5210996332	57.91	5233955899	48.36	0.44%
	2-4-12-22-85 (7+78)	5339691338	364.7	5389350467	74.74	0.93%
	3-5-18-70-205 (10+195)	5620433450	>3600	5691250911	556	1.26%
Average value						0.53%

TABLE 3: Annual average demand of 3 products for 85 customers.

No.	D1	D2	D3	No.	D1	D2	D3	No.	D1	D2	D3	No.	D1	D2	D3
1	99	9	12	23	41	27	3	45	24	4	5	67	29	5	1
2	21	7	1	24	44	1	1	46	14	4	4	68	10	2	2
3	68	18	3	25	6	5	2	47	4	2	2	69	33	8	1
4	4	2	2	26	2	0	1	48	15	12	2	70	56	9	2
5	55	2	1	27	104	23	1	49	19	7	4	71	4	1	3
6	16	6	3	28	12	4	2	50	2	1	0	72	43	23	2
7	14	5	1	29	1	1	1	51	9	7	0	73	16	5	2
8	3	1	0	30	1	1	1	52	4	1	1	74	9	1	1
9	17	3	1	31	5	2	1	53	16	7	7	75	24	2	3
10	1	0	0	32	8	1	1	54	14	8	1	76	2	1	1
11	4	1	7	33	48	9	7	55	22	7	2	77	7	3	2
12	18	1	6	34	23	15	1	56	4	3	1	78	23	5	2
13	67	14	7	35	65	8	4	57	6	2	2	79	163	19	1
14	3	1	2	36	31	15	1	58	1	1	2	80	429	4	1
15	9	5	1	37	31	3	4	59	4	3	0	81	417	54	1
16	7	3	1	38	9	1	1	60	13	7	3	82	501	25	1
17	43	3	15	39	16	2	1	61	13	3	1	83	286	22	4
18	1	0	7	40	45	11	2	62	12	2	2	84	455	25	2
19	29	2	3	41	86	8	1	63	45	8	3	85	393	45	2
20	28	8	4	42	39	5	4	64	39	14	3				
21	38	6	1	43	16	3	1	65	23	6	1				
22	5	2	3	44	6	2	1	66	67	7	1				

No. represents the number of customers; D1-D3 represent demand of products 1-3.

**Initialization**

Set the parameters  $\text{MaxItr}$ ,  $N$ ,  $w$ ,  $c_1$ ,  $c_2$ ,  $r_1$ ,  $r_2$ ,  $c_3$ ,  $r_3$ ,  $\text{Pr}$ ;

Initialize 9 independent particle swarms:

$Q_{ae1}^m(a, e_1, m)$ ,  $Q_{ab}^m(a, b, m)$ ,  $Q_{bc}^m(b, c, m)$ ,  $Q_{cd}^m(c, d, m)$ ,  $Q_{de2}^m(d, e_2, m)$

$y_a(a)$ ,  $y_b(b)$ ,  $y_c(c)$ ,  $y_d(d)$

which are in the 11-dimensional solution space  $X(a, b, c, d, e_1, e_2, m, a', b', c', d')$ ;

Define position ( $x$ ) and velocity ( $v$ ) for each particle;

Assume individual best ( $Pbest_{id}^t$ ) equal to the initialized particles;

Select the best particle as global best ( $Gbest_d^t$ );

**Update**

Let 9 particle swarms be updated independently as follows:

**For t = 1 to MaxItr**

Update position ( $x$ ) and velocity ( $v$ );

$$v_{id}^{t+1} = w * v_{id}^t + c_1 * r_1 * (Pbest_{id}^t - x_{id}^t) + c_2 * r_2 * (Gbest_d^t - x_{id}^t)$$

$$x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1}$$

**If**  $X_{id} < X_{min}$  **then**  $X_{id} = X_{min} + c_3 * r_3 * X_{min}$

**Else if**  $X_{id} > X_{max}$  **then**  $X_{id} = X_{max} - c_3 * r_3 * X_{max}$

**Respectively;**

**End**

Evaluate the particles and calculate the fitness values;

Sort the fitness values;

**If**  $\text{Pr} > \text{rand}(d)$

$$x_{id}^{t+1} = x_{id}^t * (1 + 0.5\eta)$$

**End**

Update personal best ( $Pbest_{id}^t$ ) of 9 independent particle swarms;

Update global best ( $Gbest_d^t$ ) of 9 independent particle swarms;

**End**

**Report**

Global best ( $Gbest_d^t$ ) of 9 independent particle swarms;

$Gbest_d^t$  found by splicing 9 independent particle swarms;

Fitness values calculated by  $Gbest_d^t$ .

Box 1: The improved PSO.

Products include electrical appliances, household products, and daily chemical products. The average annual demand of three products for 85 customers is shown in Table 3. Their location is shown in Figure 5.

**6.1. Location Decision Results.** The locations of the logistics center and customers (blue icon) are shown in Figure 5. After the location decision, three regional logistics centers, eight city logistics centers, nine dealers may not operate. The non-operating logistics centers are identified by red rectangle. Reasons for not operating are analyzed.

Two national logistics centers (numbered  $A_i$ , yellow icon): Since the national distribution center needs to distribute to regional logistics centers and large customers, one national distribution center's capacity is not enough, both of them need to operate.

Four regional logistics centers (numbered  $B_i$ , purple icon): Regional logistics center B2 needs to operate. B2 is closest to the two national logistics centers and is closer to the city logistics center compared with B1, B3, and B4. In the vicinity of B1, B3, and B4, the city logistics centers (numbered  $C_i$ ) are less distributed and dispersed. Considering the operating costs, B1, B3, and B4 may not operate.

Twelve city logistics centers (numbered  $C_i$ , green icon): City logistics centers C1, C4, C8, and C11 need to operate. C2, C9, and C10 are far away from B2 and have high transportation costs compared with C4 and C8, so they are not operated. C3, C5, C6, C7, and C12 are remote and scattered, and the demand of dealer is small. It does not operate due to transportation and operating costs.

Twenty-two dealers (numbered  $D_i$ , red icon): Through the calculation of location and routing, 9 dealers do not need to operate. The remaining 13 dealers can fulfill the product distribution demands of 78 second-class customers.

The above analyses show that M company's "2-4-12-22" supply chain network is too large. The current demand can be satisfied by the "2-1-4-13" scale network. The location-routing decision can reduce the operations of three regional distribution centers, eight city distribution centers, and nine dealers.

**6.2. Transportation Route Decision Results.** Distribution route of the three products for 85 customers is shown in Table 4.

If distribution routes of the three products are the same, the route decision follows the customer number, and the first

TABLE 4: Delivery route of 3 products for 85 customers (Ton).

No.	Delivery path	No	Delivery path	No.	Delivery path	No.	Delivery path
1	A1-B2-C1-D8-E1	23	A1-B2-C1-D7-E23	43	A1-B2-C11-D19-E43	67	A1-B2-C1-D1-E67
2	A1-B2-C1-D1-E2	24	A1-B2-C1-D2-E24	44	A1-B2-C1-D1-E44	68	A1-B2-C8-D21-E68
3	A1-B2-C1-D1-E3	25	A1-B2-C11-D2-E25	45	A1-B2-C11-D2-E45	69	A1-B2-C1-D1-E69
4	A1-B2-C11-D19-E4	26	A1-B2-C1-D1-E26	46	A1-B2-C1-D8-E46	70	A1-B2-C11-D19-E70
5	A1-B2-C1-D1-E5	27	A1-B2-C11-D2-E27	47	A1-B2-C11-D19-E47	71	A1-B2-C1-D1-E71
6	A1-B2-C11-D19-E6	28	A1-B2-C11-D19-E28	48	A1-B2-C11-D2-E48	72	A1-B2-C1-D1-E72
7	A1-B2-C1-D10-E7	29	A1-B2-C1-D8-E29	49	A1-B2-C1-D1-E49	73	A1-B2-C1-D1-E73
9	A1-B2-C1-D1-E9	30	A1-B2-C1-D1-E30	51	A1-B2-C8-D21-E51	74	A1-B2-C1-D1-E74
10	A1-B2-C1-D1-E10	31	A1-B2-C1-D1-E31	52	A1-B2-C1-D7-E52	75	A1-B2-C1-D1-E75
11	A1-B2-C1-D22-E11	32	A1-B2-C1-D1-E32	53	A1-B2-C1-D1-E53	76	A1-B2-C1-D1-E76
12	A1-B2-C1-D1-E12	33	A1-B2-C11-D2-E33	54	A1-B2-C1-D1-E54	77	A1-B2-C1-D1-E77
13	A1-B2-C8-D21-E13	34	A1-B2-C11-D19-E34	55	A1-B2-C1-D1-E55	78	A1-B2-C1-D1-E78
14	A1-B2-C1-D1-E14	35	A1-B2-C1-D22-E35	56	A1-B2-C1-D1-E56		A1-E79
15	A1-B2-C1-D1-E15	36	A1-B2-C1-D16-E36	57	A1-B2-C11-D17-E57	80	A2-E80
16	A1-B2-C11-D2-E16	37	A1-B2-C11-D2-E37	58	A1-B2-C1-D1-E58	81	A2-E81
17	A1-B2-C1-D22-E17	38	A1-B2-C11-D2-E38	60	A1-B2-C1-D8-E60	82	A2-E82
19	A1-B2-C1-D10-E19	39	A1-B2-C11-D19-E39	61	A1-B2-C1-D1-E61	83	A1-E83
20	A1-B2-C8-D21-E20	40	A1-B2-C4-D18-E40	62	A1-B2-C8-D21-E62	84	A2-E84
21	A1-B2-C8-D21-E21	41	A1-B2-C4-D18-E41	63	A1-B2-C1-D1-E63	85	A2-E85
22	A1-B2-C1-D8-E22	42	A1-B2-C1-D1-E42	64	A1-B2-C1-D4-E64		
8	A1-B2-C1-D8-E8	8	A1-B2-C1-D8-E8	8	/		
18	/	18	A1-B2-C1-D8-E18	18	A1-B2-C1-D8-E18		
50	A1-B2-C8-D21-E50	50	A1-B2-C8-D21-E50	50	/		
59	A1-B2-C4-D12-E59	59	A1-B2-C4-D12-E59	59	/		
65	A1-B2-C1-D4-E65	65	A1-B2-C1-D4-E65	65	/		
66	A1-B2-C4-D18-E66	66	A1-B2-C4-D18-E66	66	/		



TABLE 6: Comparison of non-site location and site location.

	Total costs	Total costs of customer <sub>1</sub>	Total costs of customer <sub>2</sub>	Product1 Volume of Customer <sub>1</sub>	Product2 Volume of Customer <sub>1</sub>	Product3 Volume of Customer <sub>1</sub>	Product1 Volume of Customer <sub>2</sub>	Product2 Volume of Customer <sub>2</sub>	Product3 Volume of Customer <sub>2</sub>	Time
Non-site location	5367103216	3422230239	1944872977	2644	194	12	1815	427	192	68
Site location	5363403216	3422230239	1941172977	2644	194	12	1815	427	192	75
Difference	3700000	0	3700000	0	0	0	0	0	0	7

achieved between the distribution network levels, and the demand network considers the different demands of customers for different products. Less attention is concentrated on integration optimization issues of distribution network and demand network in existing research.

(2) Through three improvements of particle swarm optimization, such as collaborative dimension reduction, boundary processing, and introduction of mutation factors, the performance of the algorithm is improved. The paper provides an algorithmic solution to solve high-dimensional problems.

In future study, the research results of this paper can be extended to the scale economy of the distribution network, while increasing the uncertainty factors and solving larger scale examples.

## Data Availability

The data used to support this study have been uploaded to Baidu Cloud Disk. Readers can access the data by the following link: <https://pan.baidu.com/s/1FFECqz8yP7y13j6cj-W4mA>.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China [Grant no. 71701123].

## References

- [1] P. C. Verhoef, P. Kannan, and J. J. Inman, "From multi-channel retailing to omni-channel retailing," *Journal of Retailing*, vol. 91, no. 2, pp. 174–181, 2015.
- [2] S. Min and M. Wolfınbarger, "Market share, profit margin, and marketing efficiency of early movers, bricks and clicks, and specialists in e-commerce," *Journal of Business Research*, vol. 58, no. 8, pp. 1030–1039, 2005.
- [3] J. Zhang, P. W. Farris, J. W. Irvin, T. Kushwaha, T. J. Steenburgh, and B. A. Weitz, "Crafting integrated multichannel retailing strategies," *Journal of Interactive Marketing*, vol. 24, no. 2, pp. 168–180, 2010.
- [4] S. Saghiri, R. Wilding, C. Mena, and M. Bourlakis, "Toward a three-dimensional framework for omni-channel," *Journal of Business Research*, vol. 77, pp. 53–67, 2017.
- [5] F. E. Maranzana, "On the location of supply points to minimize transportation costs," *Journal of the Operational Research Society*, vol. 15, no. 3, pp. 261–270, 1964.
- [6] K. Govindan, A. Jafarian, R. Khodaverdi, and K. Devika, "Two-echelon multiple-vehicle location-routing problem with time windows for optimization of sustainable supply chain network of perishable food," *International Journal of Production Economics*, vol. 152, no. 2, pp. 9–28, 2014.
- [7] A. Abdulrazik, M. Elsholkami, A. Elkamel, and L. Simon, "Multi-products productions from Malaysian oil palm empty fruit bunch (EFB): analyzing economic potentials from the optimal biomass supply chain," *Journal of Cleaner Production*, vol. 168, no. 2, pp. 131–148, 2017.
- [8] A. M. Geoffrion and G. W. Graves, "Multicommodity distribution system design by benders decomposition," *Management Science*, vol. 26, no. 8, pp. 855–856, 1980.
- [9] A. Tiwari, P. C. Chang, and M. K. Tiwari, "A highly optimised tolerance-based approach for multi-stage, multi-product supply chain network design," *International Journal of Production Research*, vol. 50, no. 19, pp. 5430–5444, 2012.
- [10] S. Viswanathan and K. Mathur, "Integrating routing and inventory decisions in one-warehouse multiretailer multiproduct distribution systems," *Management Science*, vol. 43, no. 3, pp. 294–312, 1997.
- [11] B. C. Arntzen, G. G. Brown, T. P. Harrison, and L. L. Trafton, "Global supply chain management at digital equipment corporation," *Interfaces*, vol. 25, no. 1, pp. 69–93, 1995.
- [12] B. H. Wang and S. W. He, "Robust optimization model and algorithm for logistics center location and allocation under uncertain environment," *Journal of Transportation Systems Engineering and Information Technology*, vol. 9, no. 2, pp. 69–74, 2009.
- [13] L. Zhen, D. Zhuge, and J. Lei, "Supply chain optimization in context of production flow network," *Journal of Systems Science and Systems Engineering*, vol. 25, no. 3, pp. 351–369, 2016.
- [14] C. S. Tapiero and M. A. Soliman, "Multi-commodities transportation schedules over time," *Networks*, vol. 2, pp. 311–327, 1972.
- [15] A. Baghalian, S. Rezapour, and R. Z. Farahani, "Robust supply chain network design with service level against disruptions and demand uncertainties: a real-life case," *European Journal of Operational Research*, vol. 227, no. 1, pp. 199–215, 2013.
- [16] G. Zhang, W. Habenicht, and W. E. L. Spieß, "Improving the structure of deep frozen and chilled food chain with tabu search procedure," *Journal of Food Engineering*, vol. 60, no. 1, pp. 67–79, 2003.
- [17] A. Ahmadi Javid and N. Azad, "Incorporating location, routing and inventory decisions in supply chain network design," *Transportation Research Part E: Logistics and Transportation Review*, vol. 46, no. 5, pp. 582–597, 2010.
- [18] Y. Xia, Z. Fu, L. Pan, and F. Duan, "Tabu search algorithm for the distance-constrained vehicle routing problem with split deliveries by order," *PLoS ONE*, vol. 13, no. 5, Article ID e0195457, 2018.
- [19] H. Hu, Y. Zhang, and L. Zhen, "A two-stage decomposition method on fresh product distribution problem," *International Journal of Production Research*, vol. 55, no. 16, pp. 4729–4752, 2017.
- [20] X. Li and X. Yao, "Cooperatively coevolving particle swarms for large scale optimization," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 2, pp. 210–224, 2012.
- [21] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, pp. 39–43, Nagoya, Japan, 1995.
- [22] Z. Lian, "A united search particle swarm optimization algorithm for multi-objective scheduling problem," *Applied Mathematical Modelling*, vol. 34, no. 11, pp. 3518–3526, 2010.
- [23] T. J. Ai and V. Kachitvichyanukul, "A particle swarm optimization for the vehicle routing problem with simultaneous pickup and delivery," *Computers & Operations Research*, vol. 36, no. 5, pp. 1693–1702, 2009.
- [24] F. P. Goksal, I. Karaoglan, and F. Altıparmak, "A hybrid discrete particle swarm optimization for vehicle routing problem with simultaneous pickup and delivery," *Computers & Industrial Engineering*, vol. 65, no. 1, pp. 39–53, 2013.

- [25] M. Alinaghian, M. Ghazanfari, N. Norouzi, and H. Nouralizadeh, "A novel model for the time dependent competitive vehicle routing problem: modified random topology particle swarm optimization," *Networks and Spatial Economics*, vol. 17, no. 4, pp. 1–27, 2017.
- [26] F. van den Bergh and A. P. Engelbrecht, "A cooperative approach to particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 225–239, 2004.

## Research Article

# Analysis of Signal Game for Supply Chain Finance (SCF) of MSEs and Banks Based on Incomplete Information Model

Zhang Tao,<sup>1</sup> Xin Li ,<sup>2</sup> Xinquan Liu ,<sup>1</sup> and Nana Feng <sup>1</sup>

<sup>1</sup>Nanning Normal University, China

<sup>2</sup>Macau University of Science and Technology, Macau

Correspondence should be addressed to Xin Li; [xli@must.edu.mo](mailto:xli@must.edu.mo) and Xinquan Liu; [gxsfx80@163.com](mailto:gxsfx80@163.com)

Received 30 November 2018; Accepted 3 February 2019; Published 26 March 2019

Academic Editor: Lu Zhen

Copyright © 2019 Zhang Tao 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 signal gaming model based on incomplete information is used to analyze the decisions of commercial banks and medium-sized and small enterprises (SMEs) in supply chain finance business. It is found that the returns of banks are closely related to the probability of good SMEs joining, which is proportional to  $\theta$  (the probability of “good” SMEs in the market) and  $p$  (the probability of “good” SMEs chosen to join the supply chain finance) in supply chain finance business, and the default cost is an important constraint for determining the strategies adopted by the SMEs and the banks. To achieve higher returns, SMEs and banks should make efforts to create a better supply chain finance business environment to achieve the separation equilibrium.

## 1. Introduction

As a new financing mode, supply chain finance has played an increasingly important role in supply chain operational and financial practices and greatly improved the current situation of financing difficulties for medium-sized and small enterprises (SMEs) [1, 2]. For banks, supply chain finance has reduced the cost of evaluating enterprises, effectively circumvented some of the credit risks, and established close ties with supply chain companies to increase the business volume. However, the issue of “information asymmetry” has also brought certain obstacles to MSE and banks.

A signaling game model based on incomplete information is established to analyze the game revenue and game process between banks and enterprises. Then, numerical simulation is utilized to analyze the interaction between variables and the influence of variables on the profits of the bank. Harsanyi [3] defined the incomplete information situation as “each player will receive only partial information about the outcome and the values of these parameters”. Some participants may find in his self-interest to impair the interests of other supply chain members in an incomplete information status. Claude d’Aspremont et al. [4] imposed compatibility condition to solve the incentive problem.

Stemmler [5] finds that the key characteristic of Supply Chain Finance (SCF) is the integration of financial flows into the physical supply chain, and SCF can be characterized as an essential part of supply chain management. Hofmann [6] describes SCF as located at the intersection of logistics, supply chain management, and finance and defines it as an approach for two or more organizations in a supply chain, including external service providers, to jointly create value by planning, steering, and controlling the flow of financial resources on an interorganizational level. The financing model of supply chain finance in practice mainly includes accounts receivable financing mode, prepaid account financing mode, and inventory financing mode. This paper analyzes the receivable category which refers to the financing mode in which the enterprise pledges to the commercial bank with its own accounts receivable and serves as the source of the first repayment.

Game analysis method used in many supply chain finance research [7], Cai, Zhang, & Zhang [8], evaluated the impact of price discount contracts and pricing schemes on the dual-channel supply chain competition. Esmaeili, Aryanezhad, and Zeephongsekul [9] proposed a Stackelberg model to analyze the relationships between seller and buyer based on noncooperative and cooperative games scenes.

TABLE 1: Bank-MSE game matrix under complete information.

Bank	MSE	
	Pay back	not pay back
Loan	$\frac{\alpha Lr - c}{1 - \delta}, \frac{\alpha\beta L - \alpha Lr}{1 - \delta}$	$-L - c, \alpha\beta L + L$
No loan	0,0	0,0

Milgrom and Roberts [10] researched the influence of pre-entry prices limit on established firm base on signal game; Friedman [11] defines the basic concept of “fitness” in evolutionary game theory and analyzed the relationship between evolutionary stability state and static equilibrium state of fitness function. Lee and Rhee [12] considered the forms of supply chain financing mechanisms like supply chain contracts.

## 2. Model

*2.1. Bank-MSE Gaming Based on Complete Information.* We assume that a MSE (labeled m) needs to loan L for a project, the success rate of the project is  $\alpha$  ( $0 < \alpha < 1$ ), and the return rate of project is  $\beta$ . The MSE borrows from the bank and the expected project revenue is  $\alpha\beta L$ . Assume that the bank’s credit cost (including evaluation costs, supervision costs, etc.) is c, the bank’s loan interest rate is r, and the bank’s loan interest income is  $\alpha Lr$ . Assume that the transaction cost between banks and enterprises is 0, and the discount rate is  $\delta$  ( $0 < \delta < 1$ ), and  $\delta$  is constant. The game matrix is as Table 1.

When  $\Pi_{c1} \geq \Pi_{c2}$ , that is  $(\alpha\beta L - \alpha Lr)/(1 - \delta) \geq \alpha\beta L + L$ , the MSE will choose to “pay back” the loan, under this situation,  $\alpha \geq (1 - \delta)/(\beta\delta - r)$ , which can be interpreted as when the project success rate  $\alpha$  is greater than or equal to  $(1 - \delta)/(\beta\delta - r)$ , the MSE will choose to “pay-back” the loan; otherwise it will “not pay back” the loan. Based on completed information, the bank can get the success rate of the project and use this as a basis to decide whether to lend to the MSE. But in real life, information is asymmetrical, banks cannot get all needed information, and such Nash equilibrium is hard to produce. In this reality, the emergence of supply chain finance has provided new ideas for SME loans.

*2.2. Supply Chain Finance and Bank-MSE Signal Game Based on Incomplete Information.* Managers often need to make decisions under uncertainty [13]. It is found that the project success rate  $\alpha$  will affect the MSEs pay back strategy, so we divide the MSEs into two types according to the difference in project success rate  $\alpha$ : if  $\alpha \geq (1 - \delta)/(\beta\delta - r)$ , the company will pay back a loan is defined as a “good” company; if  $\alpha < (1 - \delta)/(\beta\delta - r)$ , the company will not pay back a loan is defined as a “bad” company. On the other hand, all the MSEs can decide to “join” or “not join” in a supply chain finance business. The core enterprise of a supply chain is equivalent to providing guarantees for MSEs in the supply chain finance business. Therefore, if the MSE do not repay the loans, the core enterprises will have the responsibility to compensate for the losses of the banks, and the MSE will be punished if it does not pay back the loan. In the supply chain finance mode, the

signal game model is used to analyze the game between banks and MSE.

For simplicity, we given some necessary notations as follows.

(1) “Join” enterprises participate in supply chain finance, and the cost of obtaining core enterprise guarantee is S, and the default cost is F (such as the penalty imposed by core enterprises and banks, the credit loss of non-repayable loans, etc.). In the “join” mode of supply chain finance, the bank obtains the enterprise information set as f; if it does not participate in the supply chain finance, that is, the “no join” mode, the bank obtains the enterprise information set as i. The success rate of a good enterprise project is  $\alpha_1$ , and the success rate of a bad enterprise project is  $\alpha_2$ . Good companies are marked as C1; bad companies are marked as C2; banks are marked as b.

(2) Priori probability:  $P(\text{join}|C1) = p$ ;  $P(\text{not join} | C1) = 1 - p$ .

(3) Priori probability:  $P(\text{join} | C2) = q$ ;  $P(\text{not join} | C2) = 1 - q$ .

(4) The probability of being a good company in all enterprises is  $\theta$ , and the probability of a bad business is  $1 - \theta$ .

(5) Posterior probability:  $P(C1 | \text{join}) = m_1$ ;  $P(C2 | \text{join}) = 1 - m_1$ ; according to Bayes’ rule, you can get the following:

$$\begin{aligned} P(C1 | \text{join}) &= m_1 = \frac{P(\text{join} | C1) P(C1)}{P(\text{join})} \\ &= \frac{p \times \theta}{p \times \theta + (1 - \theta) \times q} \end{aligned} \quad (1)$$

(6) Posterior probability:  $P(C1 | \text{not join}) = m_2$ ;  $P(C2 | \text{not join}) = 1 - m_2$ . According to Bayes’ rule, you can get the following:

$$\begin{aligned} P(C1 | \text{not join}) &= m_2 = \frac{P(\text{not join} | C1) P(C1)}{P(\text{not join})} \\ &= \frac{(1 - p) \times \theta}{(1 - p) \times \theta + (1 - q) \times (1 - \theta)} \end{aligned} \quad (2)$$

(7) The cost of joining the supply chain finance is small and set to be negligible. At the same time, for good companies, when the benefits of “joining” and “not joining” are very similar, good companies will prefer to “join” in order to maintain stable trade income and maintain stable trade targets.

## 3. Equilibrium Analysis

*3.1. Mixed Equalization Analysis.* Under the definition of mixed equilibrium, both good and bad SMEs choose the same strategy as (join, join) or (not join, not join). There are two strategies that banks can choose from “loans” or “no loans”.

(1) *SMEs Choose (Join, Join) Strategy.* When a bank chooses to lend a loan, the proceeds are as follows:

$$\Pi_{b1}^f = m_1 \times \frac{\alpha_1 Lr - c}{1 - \delta} + (1 - m_1) \times (\alpha_2 Lr - c) \quad (3)$$

TABLE 2: Summary of mixed equalization.

Mixed equalization strategy	Strategy of SMEs	(join, join)
		(loan, no loan)
	Strategy of Bank	When MSE joins supply chain finance, the bank chooses the loan; when MSE does not join the supply chain finance, the bank chooses not to lend.
Equilibrium condition	(1) $P(C1   \text{not join}) = m_2, m_2 \leq \frac{L + c - \delta L - \delta c}{\alpha_1 Lr + L}$ .	
	(2) Default cost $F < \alpha_2 \beta L + L - S$ .	
Returns	Returns of bank	$m_1 \times \frac{\alpha_1 Lr}{1 - \delta} + (1 - m_1) \times (\alpha_2 Lr - c)$
	Returns of good MSE	$\frac{\alpha_1 \beta L - \alpha_1 Lr}{1 - \delta} - S$
	Returns of bad MSE	$\alpha_2 \beta L + L - S - F$
Description	Under the condition of mixed equilibrium, the income of enterprises “joining” is higher. Both good and bad companies will choose to join supply chain finance. Banks will be willing to believe that the enterprises in supply chain finance are probable of good enterprises and lend to supply chain finance. SMEs. When the company joins, the bank is willing to lend to the enterprises in the supply chain finance because the guarantee of the core enterprise can share the risk of the non-repayable loan of some bad enterprises.	

When a bank chooses a “no loan” strategy, the revenue of bank is  $\Pi_{b2}^f = 0$ .  $(\alpha_1 Lr - c)/(1 - \delta)$  and  $\alpha_2 Lr - c$  are obviously bigger than 0, then  $m_1 \times (\alpha_1 Lr - c)/(1 - \delta) + (1 - m_1) \times (\alpha_2 Lr - c) > 0$ , so  $\Pi_{b1}^f > \Pi_{b2}^f$ . That means the bank will get a better return if it chooses a “loan” strategy whether the MSE is a good company or a bad company. Under this situation, the return of good company which chooses to “join” is  $(\alpha_1 \beta L - \alpha_1 Lr)/(1 - \delta) - S$ , and the return of “bad” company which chooses to “join” is  $\alpha_2 \beta L + L - S - F$ ; the return of good company which chooses to “not join” is  $(\alpha_1 \beta L - \alpha_1 Lr)/(1 - \delta)$ , and the return of bad company which chooses to “not join” is  $\alpha_2 \beta L + L$ .

(2) *SMEs Choose (Not Join, Not Join) Strategy.* The analysis process is similar to the (join, join) situation. The summary of mixed equalization is shown in Table 2.

3.2. *Separation Equilibrium Analysis.* In the context of separation equilibrium, there may be two situations: good companies that choose to “join” and bad companies to choose “do not join”; good companies choose “not to join” and bad companies to choose “join”.

(1) *Good SME and Bad SME Choose (Join, Not Join) Strategy.* “Good” companies choose to “join”, while bad companies choose “not to join”. At this time,  $p=1, q=0$ . The two sets of information for the bank are both on the equilibrium path. The optimal strategy of bank is “loans” to good SME, and the return is  $(\alpha_1 Lr - c)/(1 - \delta)$ , the optimal strategy of bank is “no loans” to “bad” SME, and the optimal income of banks is 0. In summary, the optimal response of banks is (loans, no loans). The return of “good” SME is  $\Pi_{c1}^f = (\alpha_1 \beta L - \alpha_1 Lr)/(1 - \delta) - S$ ; the return of “bad” SME is  $\Pi_{c2}^i = 0$ .

Based on the assumptions, good SME will prefer to “join”. The bad SME’s income at this time is  $\Pi_{c2}^f = \alpha_2 \beta L + L - S - F$ ; at this time, we need to compare  $\Pi_{c2}^f$  and  $\Pi_{c2}^i$  size; when  $\Pi_{c2}^f < \Pi_{c2}^i$ , bad SMEs will choose “do join” strategy; otherwise they will choose “join” strategy that means  $\alpha_2 \beta L + L - S - F < 0, F > \alpha_2 \beta L + L - S$ . In this case, the bad SME “not join” will get a higher income, which constitutes separation equilibrium of [(join, not join), (loan, no loan)]. When  $F < \alpha_2 \beta L + L - S$ , bad SME “join” gains higher income, and bad companies choose “join”. At this context, the separation equalization does not exist.

(2) *Good SME and Bad SME Choose (Not Join, Join) Strategy.* The analysis process is similar to the (join, not join) situation; the conclusion is that when SME strategy is (not join, join), the separation equalization does not exist.

3.3. *Quasi-Separation Equilibrium Analysis.* Quasi-separation equilibrium refers to that when one type of signal issuer chooses a strategy, another signal sender randomly chooses between two strategies. In our research, we assume that “good” SME chooses “join”, and then the “bad” SME can choose “join” or “not join” randomly.

Based on definition of Quasi-separation equilibrium, we set the probability of good SME in which “join” to supply chain is  $P=1$  and the probability of bad SME is  $q$  (choose “join” strategy) or  $1-q$  (choose “not join” strategy). According to Bayes’ rule, the following is available:

$$P(C1 | \text{join}) = m_1 = \frac{P(\text{join} | C1) P(C1)}{P(\text{join})}$$

$$= \frac{p \times \theta}{p \times \theta + (1 - \theta) \times q}$$

$$\begin{aligned}
P(C1 | \text{not join}) &= m_2 = \frac{P(\text{not join} | C1)P(C1)}{P(\text{not join})} \\
&= \frac{(1-p) \times \theta}{(1-p) \times \theta + (1-q) \times (1-\theta)} \\
m_2(\text{good SME, not join}) &= 0.
\end{aligned} \tag{4}$$

When  $q$  tends to 0, the probability of bad SME choosing to join tends to 0, bad SME will hardly be confused with good SME, and banks infer  $P(C1 | \text{not joined}) = m_2 = 0$ . When a bad SME chooses “not to join” and thus can be separated from a good SME, the probability of good SME is equal to zero in “not join” group.

When the bad SME choose to “join”, and mix with good SME,  $P(\text{join} | C2) = q$ ,  $q$  tends to 1;  $P(\text{not join} | C1) = p = 0$ . When  $q$  tends to 1,  $p = 0$ , so the bank infers that  $m_1$  tends to a priori infer  $\theta$ .

To a “bad” SME, the gain that will be obtained in the “not joined” state is  $\Pi_{c2}^i = 0$ , in the context of mixed equilibrium,  $\Pi_{c2}^f = \alpha_2\beta L + L - S - F$ , to make  $\Pi_{c2}^f = \Pi_{c2}^i$ , then  $\Pi_{c2}^f = \alpha_2\beta L + L - S - F = 0$ ;  $F = \alpha_2\beta L + L - S$ . Therefore, if there is a quasi-separation equilibrium, the bad SME will have a return of 0 regardless of whether it chooses to “join” or “not join” strategy.

To a bank, when receiving the signal of “not joining”, the probability that the bank thinks it is a good SME is equal to 0 in the “not join” group. The good SME is naturally separated from the bad SME. If the bank’s loan income is  $(-L-c)$ , it must be negative. Therefore, the bank’s strategy is not to lend to the “not join” SME with an income 0. When the bank faces the “join” situation of the bad SME, the probability of the good SME in the “join” group enterprise is inferred to be  $m_1$ , because the hypothesis is  $p = 1$ ; then  $m_1$  tends to the prior probability  $\theta$ ; at this time, the bank “loan” income is

$$\Pi = \theta \times \frac{\alpha_1 L r - c}{1 - \delta} + (1 - \theta) \times (\alpha_2 L r - c) \tag{5}$$

At this situation, bank will make a “loan” decision because  $\Pi$  must bigger than 0.

### 3.4. Comparison among Mixed Equalization, Separation Equilibrium, and Quasi-Separation Equilibrium Analysis

(1) *Equilibrium Constrains.* The summary of equilibrium constrains is shown in Table 3.

The value of  $m_1$  is needs to be stint to meet the mixed equilibrium condition when the numerical simulation of bank income is in the next step. We let  $L=100$ ,  $r=2.5\%$ ,  $c=1$ ,  $\alpha_1=0.9$ ,  $\delta=0.71$ , then  $m_2 \leq 0.29$ , that is, the probability of good SMEs which “not join” in supply chain finance in the market is less than 0.29; the Mixed equalization is established.  $m_1$  is Posterior probability:

$$m_1 = \frac{p \times \theta}{p \times \theta + (1 - \theta) \times q} \tag{6}$$

TABLE 3: Equilibrium constrains.

	Equilibrium constrains
Mixed equalization	$m_2 \leq \frac{L + c - \delta L - \delta c}{\alpha_1 L r + L}$ $F < \alpha_2 \beta L + L - S$
Separation equilibrium	$F > \alpha_2 \beta L + L - S$
Quasi-separation equilibrium	$F = \alpha_2 \beta L + L - S$

TABLE 4: The effect of  $\theta$  on  $m_1$ .

$\theta$	p	q	$m_1$
0	0.8	0.2	0
0.1	0.8	0.2	0.307692308
0.2	0.8	0.2	0.5
0.3	0.8	0.2	0.631578947
0.4	0.8	0.2	0.727272727
0.5	0.8	0.2	0.8
0.6	0.8	0.2	0.857142857
0.7	0.8	0.2	0.903225806
0.8	0.8	0.2	0.941176471
0.9	0.8	0.2	0.972972973
1	0.8	0.2	1

We first analyze the value of  $m_1$ , then adjust the value of  $p$ ,  $\theta$  and  $q$ , and observe the impact on  $m_1$ . We let  $p=0.8$ ,  $q=0.2$ , adjusting the value of  $\theta$  from 0 to 1, the effect of  $\theta$  on  $m_1$  is as follows. The effect of  $\theta$  on  $m_1$  is shown in Table 4.

We found that the values of  $\theta$  and  $m_1$  are positively correlated which means that the more the good SME in the market, the greater the probability of good SME in the “joined” SME group. Similarly, the values of  $p$  and  $q$  can be analyzed, values of  $p$  and  $m_1$  are positively correlated, and  $q$  and  $m_1$  are negatively correlated. The establishment of the mixed equalization requires  $m_2 \leq 0.29$ . Since  $m_1, m_2$  are affected by value of  $p, q, \theta$ ,  $m_2 \leq 0.29$  is required. In order to compare the value range of  $m_1$  with  $m_2$ , we set the same group of  $p, q, \theta$ . The effect of  $\theta$  on  $m_2$  is as follows. The effect of  $\theta$  on  $m_2$  is shown in Table 5.

To make  $m_2 \leq 0.29$ , when  $p=0.8$ ,  $q=0.2$ ,  $\theta \leq 0.6$ , then  $m_1 \leq 0.857142857$ , which means that the value of  $m_1$  is 0 to 0.857142857 in this case.

(2) *Returns of Bank.* Summary of returns of bank is shown in Table 6.

We can find that the difference of returns is based on the posterior probability  $m_1$  (separation equilibrium  $m_1=1$ ; quasi-separation equilibrium  $m_1=\theta$ ); what follows, we use a few sets of data for a simple comparison.

In the previous section, we have conducted some relation of  $m_1$  and  $\theta$  using numerical simulation in which  $p=0.7, q=0.5$ , and the mixed equalization model requires  $m_1 \leq 0.8$ . Now, we let  $L=100$ ,  $r=2.5\%$ ,  $\beta=0.5$ ,  $c=1$ ,  $\alpha_1=0.9$ ,  $\alpha_2=0.2$ ,  $\delta=0.71$ ;  $m_1$  is adjusted between 0 and 1; the change of returns of banks is shown in Table 7.

TABLE 5: The effect of  $\theta$  on  $m_2$ .

$\theta$	p	q	$m_2$
0	0.8	0.2	0
0.1	0.8	0.2	0.027027027
0.2	0.8	0.2	0.058823529
0.3	0.8	0.2	0.096774194
0.4	0.8	0.2	0.142857143
0.5	0.8	0.2	0.2
0.6	0.8	0.2	0.272727273
0.7	0.8	0.2	0.368421053
0.8	0.8	0.2	0.5
0.9	0.8	0.2	0.692307692
1	0.8	0.2	1

TABLE 6: Returns of banks.

Equations of returns	
Mixed equalization	$m_1 \times \frac{\alpha Lr - c}{1 - \delta} + (1 - m_1) \times (\alpha Lr - c)$
Separation equilibrium	$\frac{\alpha_1 r - c}{1 - \delta}$
Quasi-separation equilibrium	$\theta \times \frac{\alpha_1 Lr - c}{1 - \delta} + (1 - \theta) \times (\alpha_2 Lr - c)$

TABLE 7: Simulation of returns of bank.

$m_1$	Mixed equalization	Separation equilibrium ( $m_1=1$ )	Quasi-separation equilibrium
0	-0.5	--	-0.5
0.1	-0.01896552	--	-0.01897
0.2	0.462068966	--	0.462069
0.3	0.943103448	--	0.943103
0.4	1.424137931	--	1.424138
0.5	1.905172414	--	1.905172
0.6	2.386206897	--	2.386207
0.7	2.867241379	--	2.867241
0.8	3.348275862	--	3.348276
0.9	--	--	3.82931
1	--	4.310345	4.310345

It is shown that the returns of bank  $\Pi_b$  increase among the increasing of  $m_1$ . The value of  $m_1$  under the separation equilibrium is 1; the bank's income is the largest in a separation equilibrium (join, not join), (loan, no loan) strategy situation with other variables constant.

(3) *Returns of SMEs.* Summary on returns of “good” SMEs is shown in Table 8.

The returns of good SME are the same among the three equilibrium situations with the strategies of good SME and bank being (“join”, “loan”).

Summary on returns of “bad” SMEs is shown in Table 9.

TABLE 8: Returns of “good” SMEs.

Returns of “good” SME	
Mixed equalization	$\frac{\alpha_1 \beta L - \alpha_1 Lr}{1 - \delta} - S$
Separation equilibrium	$\frac{\alpha_1 \beta L - \alpha_1 Lr}{1 - \delta} - S$
Quasi-separation equilibrium	$\frac{\alpha_1 \beta L - \alpha_1 Lr}{1 - \delta} - S$

TABLE 9: Returns of “bad” SMEs.

Returns of “bad” SME	
Mixed equalization	$\alpha_2 \beta L + L - S - F$
Separation equilibrium	0
Quasi-separation equilibrium	0

In practices, the bad SMEs will automatically choose to leave the supply chain finance business when the cost of default reaches a certain value. The constrains of mixed equalization is  $F < \alpha_2 \beta L + L - S$  for “bad” SMEs, which can get higher returns when they join the supply chain finance under the mixed equilibrium situation.

#### 4. Conclusions

The returns of banks closely refer to the probability of good SMEs join in supply chain finance business, denoted as  $m_1$ , which is proportional to  $\theta$  (the probability good SMEs in the market) and p (the probability good SMEs chooses to join the supply chain finance) and is inversely proportional to q (the probability that bad SME chooses to join supply chain finance). That is, with more good SMEs in the market, and more good SMEs choose to join in the supply chain finance business, SMEs and banks can get higher returns from supply chain finance business.

The default cost F is an important constrain for determining the strategies adopted by the SMEs and the banks. From the separation equilibrium, we can find that the higher default cost F can effectively curb the bad SMEs to join the supply chain finance business and then enhance the stability of the supply chain financial business.

In an incomplete information situation, it is particularly beneficial to the bad SME under the mixed equilibrium situation, and the separation equilibrium is beneficial to the banks. Under the separation equilibrium, good SMEs can obtain more loan resources and get better partners among the living of bad SME, and the core enterprises of the supply chain can reduce some of the guarantee risks. So, we suggest that government departments and enterprises jointly adopt appropriate means to create a good financial environment to promote the emergence of separation equilibrium.

#### 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.

## Authors' Contributions

Zhang Tao contributed in writing the paper and communication. Xin Li provided the original research idea. Xinquan Liu contributed in the modeling stage. Nana Feng revised the paper.

## Acknowledgments

The work was supported by the National Natural Science Foundation of China (no. 71462005) and 2017 Philosophy and Social Sciences Planning Project of Guangxi Zhuang Autonomous Region (No. 17FJY019). The second author was supported in part by the National Natural Science Foundation of China under grant no. 71801233.

## References

- [1] N. Yan, B. Sun, H. Zhang, and C. Liu, "A partial credit guarantee contract in a capital-constrained supply chain: financing equilibrium and coordinating strategy," *International Journal of Production Economics*, vol. 173, pp. 122–133, 2016.
- [2] B. Milder, "Closing the gap: Reaching the missing middle and rural poor through value chain finance," *Enterprise Development and Microfinance*, vol. 19, no. 4, pp. 301–316, 2008.
- [3] J. C. Harsanyi, "Games with incomplete information played by "Bayesian" players, I–III Part I. The basic model," *Management Science*, vol. 14, no. 3, pp. 159–182, 1967.
- [4] C. D'Aspremont and L.-A. Gérard-Varet, "Incentives and incomplete information," *Journal of Public Economics*, vol. 11, no. 1, pp. 25–45, 1979.
- [5] L. Stemmler, "The role of finance in supply chain management," in *Cost Management in Supply Chains*, pp. 165–176, Physica, Heidelberg, Germany, 2002.
- [6] E. Hofmann, *Supply Chain Finance: Some Conceptual Insights*, Beiträge Zu Beschaffung Und Logistik, 2005.
- [7] X. Xu, X. Chen, F. Jia, S. Brown, Y. Gong, and Y. Xu, "Supply chain finance: A systematic literature review and bibliometric analysis," *International Journal of Production Economics*, vol. 204, pp. 160–173, 2018.
- [8] G. G. Cai, Z. G. Zhang, and M. Zhang, "Game theoretical perspectives on dual-channel supply chain competition with price discounts and pricing schemes," *International Journal of Production Economics*, vol. 117, no. 1, pp. 80–96, 2009.
- [9] M. Esmaili, M.-B. Aryanezhad, and P. Zeepongsekul, "A game theory approach in seller-buyer supply chain," *European Journal of Operational Research*, vol. 195, no. 2, pp. 442–448, 2009.
- [10] P. Milgrom and J. Roberts, "Limit pricing and entry under incomplete information: An equilibrium analysis," *Journal of the Econometric Society*, pp. 443–459, 1982.
- [11] D. Friedman, "Evolutionary games in economics," *Journal of the Econometric Society*, vol. 59, no. 3, pp. 637–666, 1991.
- [12] C. H. Lee and B.-D. Rhee, "Coordination contracts in the presence of positive inventory financing costs," *International Journal of Production Economics*, vol. 124, no. 2, pp. 331–339, 2010.
- [13] L. Zhen, "Tactical berth allocation under uncertainty," *European Journal of Operational Research*, vol. 247, no. 3, pp. 928–944, 2015.

## Research Article

# Bilevel Programming Model of Urban Public Transport Network under Fairness Constraints

Jingjing Hao <sup>1,2,3</sup> Xinquan Liu <sup>4</sup> Xiaojing Shen <sup>4</sup> and Nana Feng <sup>4</sup>

<sup>1</sup>School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China

<sup>2</sup>Yunnan Integrated Transport Development and Regional Logistics Management Think Tank, Kunming University of Science and Technology, Kunming 650504, China

<sup>3</sup>Yunnan Engineering Research Center of Modern Logistics, Kunming University of Science and Technology, Kunming 650504, China

<sup>4</sup>School of Logistics Management and Engineering, Nanning Normal University, Nanning 530001, Guangxi, China

Correspondence should be addressed to Xiaojing Shen; 614243786@qq.com

Received 27 November 2018; Accepted 6 February 2019; Published 12 March 2019

Academic Editor: Lu Zhen

Copyright © 2019 Jingjing Hao 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, the bilevel programming model of the public transport network considering factors such as the per capita occupancy area and travel cost of different groups was established, to alleviate the urban transportation equity and optimize the urban public transport network under fairness constraints. The upper layer minimized the travel cost deprivation coefficient and the road area Gini coefficient as the objective function, to solve the optimization scheme of public transport network considering fairness constraints; the lower layer was a stochastic equilibrium traffic assignment model of multimode and multiuser, used to describe the complex selection behavior of different groups for different traffic modes in the bus optimization scheme given by the upper layer. The model in addition utilised the noninferior sorting genetic algorithm II to validate the model via a simple network. The results showed that (1) the travel cost deprivation coefficient of the three groups declined from 33.42 to 26.51, with a decrease of 20.68%; the Gini coefficient of the road area declined from 0.248 to 0.030, with a decrease of 87.76%; it could be seen that the transportation equity feeling of low-income groups and objective resource allocation improved significantly; (2) before the optimization of public transport network, the sharing rate of cars, buses, and bicycles was 42%, 47%, and 11%, respectively; after the optimization, the sharing rate of each mode was 7%, 82%, and 11%, respectively. Some of the high and middle income users who owned the car were transferred to the public transportation. It could be seen that the overall travel time of the optimized public transport network reduced, enhancing the attraction of the public transport network to various travel groups. The model improves the fairness of the urban public transport system effectively while ensuring the travel demand of the residents. It provides theoretical basis and model foundation for the optimization of public transit network, and it is a new attempt to improve the fairness of the traffic planning scheme.

## 1. Introduction

Equity is a topic of concern to the whole society. The promotion of equity is more important than the increase of wealth to a great extent. Although the issue of equity has aroused great concern in society, the transportation equity has not received enough attention. At present, urban traffic planning theory attaches importance to traffic efficiency and neglects transportation equity, making it more difficult to be guaranteed.

In recent years, certain scholars have begun to explore the scheme of transportation equity. For instance, Litman

[1] analyzed the manifestation of transportation equity and its interaction with urban planning, traffic planning, and traffic development strategies systematically. Vasconcellos [2] studied the negative externalities of different social groups due to travel by using disaggregate model and pointed out that low-income groups only obtain lower levels of mobility when they spend more on transportation costs in a selected case of São Paulo. Preston et al. [3] investigated the social exclusion phenomenon related to transportation from the perspective of accessibility and mobility and constructed a social-spatial model that caused social exclusion. Ahmed et al. [4] conducted a comparative analysis of Beijing and

Karachi cities as examples and illustrated the development trend of transportation equity in the urbanization from a macrolevel; Nuworsoo et al. [5] analyzed the impact of different public transport charging schemes on transportation equity through resident surveys and assessed the actual impact on low-income groups specifically; Brocker et al. [6] studied the actual impact of transportation infrastructure investment and construction on regional transportation equity and established a spatial equilibrium model to evaluate the transportation infrastructure investment policy; Ying et al. [7] used the Logit model to describe the competition and development of different traffic modes in urban traffic networks and evaluated the fairness of road resource allocation under various scenarios via the Lotka-Volterra model.

The improvement of transportation equity has also attracted the attention of researchers. Farrington et al. [8] analyzed the impact mechanism of accessibility and transportation system on social exclusion in suburban and rural areas and discussed policy design to improve accessibility from the perspective of social equity; Olvera et al. [9] provided policy recommendations for enhancing residents' travel environment and social equity based on the analysis of trends in different household transportation expenditures. Duvarcia et al. [10] discussed the suppressed travel demand of traffic vulnerable groups and their equality with ordinary groups and put forward corresponding suggestions from the technical aspects of urban traffic policy and traffic planning; Lucas [11] took South Africa as an example to analyze the typical social exclusion phenomenon in urban transportation systems in developing countries and proposed specific suggestions for improving the traffic environment of low-income groups from the perspective of traffic policy; Ferguson et al. [12] exemplified the important impact of the accessibility of public transport services on the employment opportunities for low-income groups and tried to embed the evaluation indicators of transportation equity into the evaluation process of the public transport system.

It is not difficult to find that the current research focuses on equity assessment, mainly on the equity analysis of travel expenses, road resource allocation, and congestion charging policies, and proposes specific countermeasures from the aspects of public participation, traffic demand analysis, and operational design. The fairness of public transport system means that, through the public transit network optimization, more road traffic resources can be provided for bus travel groups. However, little attention is paid to the lack of fairness in urban transportation planning practice, especially in the case of relatively fixed resource elements including urban spatial layout, land use form, and road network structure, which planning pattern can be used to maximize the fairness and efficiency.

As the leading travel mode in urban transportation system, public transport and the optimization of its network guarantees basic accessibility for low-income groups effectively, which directly affects urban transportation equity. Thus, the public transport system is of great significance to improve transportation equity. Its service level is related to the transportation rights of most residents, and it is also the

adjustment lever of transportation equity among different social groups. It thereby improves the overall accessibility of the public transport system and shortens the gap between the bus travel groups and other travel groups. Finally, the bus travel groups and other travel groups have the same opportunities to participate in social activities. Obviously, it is helpful to meet the travel needs of different social groups by improving the fairness of public transport system. In particular, it is of great significance for the improvement of the travel environment of vulnerable groups. However, the traditional public transport network optimization schemes are mostly optimized with the minimum passenger travel time, the highest passenger flow rate, the lowest line overlap factor, and the largest bus economic benefit [13–20], excluding the transportation equity. Therefore, in view of uneven urban transportation resources and differentiated spatial accessibility, this paper intends to use the Gini coefficient of road area and relative deprivation coefficient of travel cost to evaluate the utility of public transit network optimization scheme. The bilevel optimization model of public transport network considering transportation equity constraints is established, and NSGA-II algorithm is designed to solve the problem. The main work in this paper is as follows:

- (i) The complex travel behavior of urban residents is analyzed systematically by constructing multimode and multiuser traffic network.
- (ii) The travel cost relative deprivation coefficient and the road area Gini coefficient are proposed to characterize the fairness of the public transport system, based on travel perception and resource allocation.
- (iii) Under the fairness constraint, the bilevel planning model of urban public transport network is constructed. Moreover, the algorithm is designed, and the validity of the model is verified by examples.

The structure of the paper is arranged as follows: (1) construct multimode and multiuser traffic network and analyze the complex travel behavior of urban residents; (2) construct the bilevel planning model of public transport network considering fairness constraints, and give algorithm for solving the model; (3) design an example to verify the actual calculation and validity of the model; (4) give the main conclusions of this paper.

## 2. Urban Traffic Network of Multimode and Multiuser

In order to facilitate the modeling and analysis of this paper, three travel modes, car, bus and bicycle, are considered in the traffic network, and the travelers are classified into three travel groups, high, medium, and low, according to income levels. A simple transportation network is constructed as shown in Figure 1. It consists of 2 traffic zones, 16 nodes and 24 road segments, where O and D indicate the origin and destination zones, respectively.

After each travel group chooses the traffic modes, the mapping to urban transportation network presents different subnetwork structure, that is, the above networks are divided

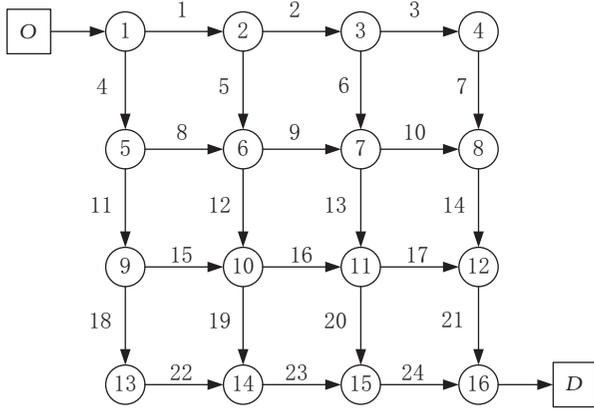


FIGURE 1: Sample of transportation network.

into high income group travel subnets, medium income group travel subnets, and low-income group travel subnets. Also, according to the traffic mode, different groups of travel subnets are subdivided into car subnets, bus subnets, and bicycle subnets. In this study, the initial public transport network among different groups is all the same. There are structural differences between the car and the bicycle subnet, and these differences do not change with the adjustment of the public transit network. Therefore, it can be considered that the multimode and multiuser traffic network mentioned above is composed of nine layers of subnets. Each subnet consists of a single group and a single travel mode. Moreover,  $G_{ik}$  represents the network formed by the  $i$ -th group choosing the  $k$ -th traffic mode. Among them,  $i = 1, 2, 3$  represents three groups of high, medium and low income;  $k = 1, 2, 3$  represents three travel modes of car, bus and bicycle.

Assuming that the travel demand of the multimode and multiuser transportation network is known and invariant, the travel demand of different traffic modes should meet the following constraints:

$$\sum_i q_i^w = q^w \quad \forall i, w \quad (1)$$

$$\sum_k q_{i,k}^w = q_i^w \quad \forall i, k, w \quad (2)$$

where  $q^w$  denotes the total travel amount of OD pair  $w$ ;  $q_i^w$  denotes the travel demand of the  $i$ -th group OD pair  $w$ ;  $q_{i,k}^w$  denotes the travel demand of the  $i$ -th group choosing the traffic mode  $k$  on OD pair  $w$ .

The traveler's choice of transportation is affected by many factors. This paper only considers factors such as transportation cost and travel time. Combining the car ownership ratios of different groups, the Logit model is used to deal with traffic mode choice. After determining the mode, the traveler makes a route selection for the traffic mode subnet. For each type of

traveler, the travel demands of the road segment and the route have the following relationship:

$$x_{i,k}^a = \sum_w \sum_r f_{i,k}^{w,r} \delta_{i,k}^{w,r,a} \quad \forall w, a \quad (3)$$

$$\sum_r f_{i,k}^{w,r} = q_{i,k}^w \quad \forall w, a \quad (4)$$

where  $x_{i,k}^a$  denotes the travel demand on the segment  $a$  in the subnet  $G_{ik}$ ;  $f_{i,k}^{w,r}$  denotes the travel demand of the OD pair  $w$  on the route  $r$  in the subnet  $G_{ik}$ ;  $\delta_{i,k}^{w,r,a}$  denotes the Boolean variable associated with routes and segments. If the segment  $a$  is on the route  $r$  that between OD pair  $w$ , its value is 1; otherwise it is 0.

Considering the difference in the average passenger capacity of each traffic mode, the travel demand of different traffic modes needs to be converted into a unified road flow. The conversion formula is

$$v_{i,k}^a = x_{i,k}^a \frac{E_k}{N_k} \quad \forall k, a \quad (5)$$

where  $v_{i,k}^a$  denotes the road segment traffic on the segment  $a$  in the subnet  $G_{ik}$ ;  $E_k$  denotes the equivalent car conversion coefficient of the traffic mode  $k$ ;  $N_k$  denotes the average number of passengers in the traffic mode  $k$ . In the subnet  $G_{ik}$ ,  $h_{i,k}^{w,r}$  denotes the road traffic on the route  $r$  and satisfies the following path-segment relationship:

$$v_{i,k}^a = \sum_w \sum_r h_{i,k}^{w,r} \delta_{i,k}^{w,r,a} \quad \forall w, a \quad (6)$$

$$h_{i,k}^{w,r} \geq 0 \quad (7)$$

Since the standard BPR function does not consider the interaction of different traffic modes, it is not suitable for multimode traffic networks. Thus, this paper uses the following improvements:

$$t_{i,k}^a = t_{i,k}^{a(0)} \prod_k \left[ 1 + 0.15 \left( \frac{\sum_{i=1} v_{i,k}^a}{C_k^a} \right)^4 \right] \quad \forall k, a \quad (8)$$

where  $t_{i,k}^a$  represents the travel time on the segment  $a$  in the subnet  $G_{ik}$ ;  $t_{i,k}^{a(0)}$  represents the free flow time on the segment  $a$  in the subnet  $G_{ik}$ ;  $C_k^a$  represents the travel capacity of the  $k$ -th travel mode for all groups on the segment  $a$ . In the subnet  $G_{ik}$ , the route travel time and the segment travel time satisfy the following relationship:

$$d_{i,k}^{w,r} = \sum_a t_{i,k}^a \delta_{i,k}^{w,r,a} \quad \forall w, a \quad (9)$$

where  $d_{i,k}^{w,r}$  represents the travel time on the route  $r$  in the subnet  $G_{ik}$ .

In the multimode and multiuser traffic network, assuming that the three types of travel groups select routes according to the Logit mode, the probability that the  $i$ -th travel group selects the path  $r$  in the subnetwork  $G_{ik}$  is

$$P_{i,k}^{w,r} = \frac{\exp(-\theta^i \cdot d_{i,k}^{w,r})}{\sum_l \exp(-\theta^i \cdot d_{i,k}^{w,l})} \quad r \neq l \quad (10)$$

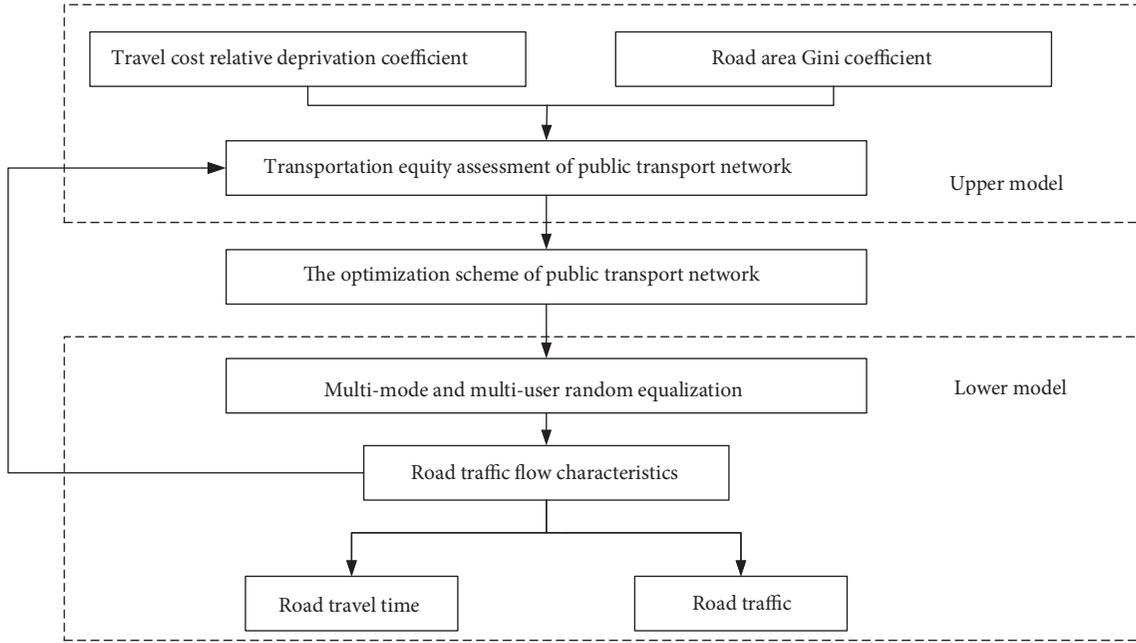


FIGURE 2: Model structure.

where  $P_{i,k}^{u,r}$  represents the probability that the  $i$ -th travel group chooses the route  $r$  in the subnet  $G_{ik}$ ;  $\theta^i$  reflects the familiarity of different travel groups with the road network.

### 3. Model Construction and Solution

The optimization of public transport network considering transportation equity constraint is a special network optimization problem. The optimization of the public transport network leads to the redistribution of traffic resources inevitably, resulting in changes in the traffic mode sharing rate and general travel cost of various travel groups. Moreover, the distribution of bus travel demand of different groups is also affected, which leads to differences in residents' travel costs and per capita road occupancy area. The bilevel optimization model of public transport network considering transportation equity in this paper is an effective improvement on the existing bilevel optimization model of public transport network. In the bilevel optimization model of public transport network, the planners first propose the optimization scheme for the original one, and the lower layer model completes the multimode and multiuser traffic allocation according to the optimization scheme. The two objective functions of the upper model are calculated by the traffic and travel time of each traffic mode to evaluate the adaptability and fairness of the optimization scheme. The model structure is shown in Figure 2.

**3.1. Upper Model.** The assessment of transportation equity is mainly based on the subjective feelings of residents and the allocation of objective resources. The generalized travel cost and the possession of road resources determine the choice of residents' travel modes and routes. Hence, the objective

function of the upper model mainly includes the travel cost relative deprivation coefficient and the road area Gini coefficient. The travel cost relative deprivation coefficient is used to express the subjective feelings of residents, and the road area Gini coefficient is used to express the allocation of objective resources.

**3.1.1. Travel Cost Relative Deprivation Coefficient.** Travel cost relative deprivation coefficient refers to certain travelers who measure the cost of completing their travel behavior and other groups by comparing the generalized travel costs between different groups to reflect the coefficient of deprivation of a certain travel group.

$$\min y_1 = \sum_k \sum_{i,j} |H_{j,k}^w - H_{i,k}^w| \quad i \neq j \quad (11)$$

$$s.t. \quad H_{i,k}^w = \tau_i \cdot T_{i,k}^w + L_k \quad \forall w, i, j \quad (12)$$

where  $y_1$  denotes the relative deprivation of travel costs incurred by all travelers in the multimode transportation network, and the smaller the  $y_1$ , the lower the relative deprivation of all travelers;  $H_{i,k}^w$  denotes the travel cost of selecting the  $k$ -th way for the OD pair  $w$  in the subnet  $G_{ik}$  to the  $i$ -th resident group;  $H_{j,k}^w$  denotes the travel cost of selecting the  $k$ -th way for the OD pair  $w$  in the subnet  $G_{jk}$  to the  $j$ -th group of residents;  $T_{i,k}^w$  denotes the travel time of the  $k$ -th mode in the subnet  $G_{ik}$  to the  $i$ -th resident group for the OD pair  $w$ , and its value can be expressed by the travel time in the subnetwork balance state;  $\tau_i$  denotes the time value of the  $i$ -th resident group;  $L_k$  denotes the cost of vehicle use for selecting the  $k$ -th traffic mode, such as car fuel consumption, bus fare.

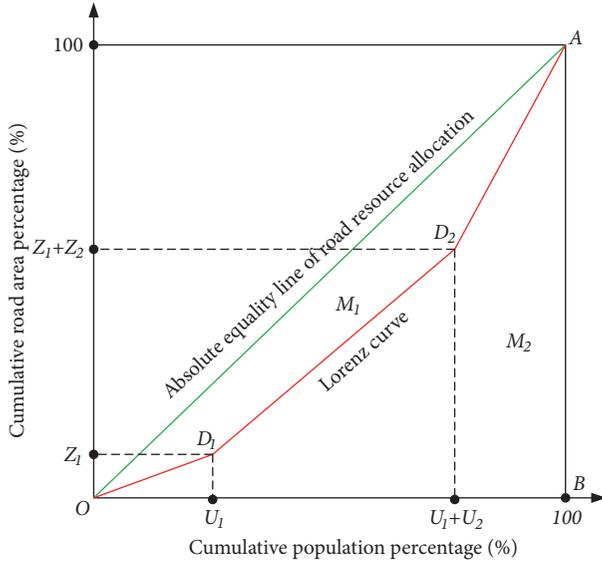


FIGURE 3: Lorenz curve for transport resource allocation.

According to the above discussion, the Logit model was used to describe the random traffic behavior of different traffic modes between the OD pairs based on the principle of the minimum travel cost. The formula is as follows:

$$Q_{i,k}^w = \frac{\exp(-\lambda \cdot H_{i,k}^w)}{\sum_{\pi} \exp(-\lambda \cdot H_{i,\pi}^w)} \quad \forall k, \pi, k \neq \pi \quad (13)$$

where  $Q_{i,k}^w$  denotes the probability that the  $i$ -th group selects the traffic mode  $k$  on OD pair  $w$ ;  $\lambda$  denotes the correction parameter.

**3.1.2. Road Area Gini Coefficient.** In order to quantify the difference of road occupation area per capita among different groups, the Lorenz curve and Gini coefficient were used to measure the fairness of road resource allocation. As shown in Figure 3, the Lorenz curve is  $OD_1D_2A$ , and the absolute fair line for road resource allocation is  $OA$ . The fairness of distribution is determined by the distance between the Lorenz curve and the absolute fair line. That is, the smaller the area enclosed by the Lorenz curve and the absolute fair line, the smaller the Gini coefficient, and the more equitable the actual distribution of traffic resources.

The radians of Lorenz curve are mainly determined by three groups of numerical values: the ratio of low-income group  $U_1$  and the ratio of road resources occupied by the group  $Z_1$ , the ratio of medium income group  $U_2$  and the ratio of road resources occupied by the group  $Z_2$ , and the ratio of high income group  $U_3$  and the ratio of road resources occupied by the group  $Z_3$ . The proportion of road resources occupied by different groups can be expressed as

$$Z_i = \frac{\sum_k (q_{i,k}^w \cdot S_k / N_k)}{\sum_j \sum_k (q_{j,k}^w \cdot S_k / N_k)} \quad \forall i, j, k \quad (14)$$

where  $S_k$  denotes the road area occupied by the single vehicle of the  $k$ -th travel mode;  $N_k$  denotes the standard passenger number of the single vehicle of the  $k$ -th travel mode.

Therefore, it can be determined that the area  $M_2$  enclosed by the line of points  $O, D_1, D_2, A, B$  is

$$M_2 = \frac{1}{2} [Z_1(U_1 + 2U_2 + U_3) + Z_2(U_2 + U_3) + U_3] \quad (15)$$

Then the Gini coefficient evaluation equation is

$$\min y_2 = \frac{M_1}{M_1 + M_2} = 1 - 2M_2 \quad (16)$$

$$\text{s.t. equation (14), (15)} \quad (17)$$

where  $y_2$  represents the Gini coefficient of the road area;  $M_1$  represents the area enclosed by the Lorenz curve and the absolute equality line.

**3.2. Lower Model.** Considering that it is impossible for travelers to master the traffic state of the road network completely, the route selection should be a stochastic process. Therefore, the stochastic user equilibrium assignment model was adopted to characterize the route choice behavior for travelers, of which the travel mode of each group was determined first. Further, the traffic assignment results which solved by the optimal target of minimum travel time in terms of the process of users' route choice behavior were put into the upper model. The optimization objective function of the lower model is minimizing the sum of the travel time of each subnet, and the specific model is as follows:

$$\begin{aligned} \min y_3 &= \\ &= \frac{1}{\theta^i} \sum_i \sum_k \sum_w \sum_r h_{i,k}^{w,r} \ln h_{i,k}^{w,r} \end{aligned} \quad (18)$$

$$+ \sum_i \sum_k \sum_{\alpha} \int_0^{v_{i,k}^{\alpha}} t_{i,k}^{\alpha}(x) dx$$

$$\text{s.t. equation (1) - (4), (6), (7)} \quad (19)$$

where  $y_3$  is the sum of travel time.

**3.3. Model Solution.** Due to the conflict of the objective functions, the travel cost relative deprivation coefficient and the road area Gini coefficient in the upper optimization only obtain multiple sets of noninferior solutions or suboptimal solutions, and it is impossible to obtain the optimal solution simultaneously. Hence, the decision maker chooses a set of noninferior solutions based on the degree of preference. The traditional multiobjective optimization method uses weighted method and transforms multiple targets into single-objective function to solve the problem. The disadvantage of this method is that the weights of each objective are subjective and there is no alternative. The best method is noninferior sorting genetic algorithm II, which gets multiple Pareto optimal solutions after one program runs. With the advantages of speediness, diversity, and uniformity, it has

TABLE 1: Relevant data of the sample of public transport network.

Segment	1	2	3	4	5	6	7	8	9	10	11	12
$t_{i,2}^{a(0)}/h$	0.196	0.186	—	—	0.165	0.164	—	—	—	—	—	0.157
$C_2^a(\text{veh}\cdot h^{-1})$	900	600	—	—	800	700	—	—	—	—	—	700
Segment	13	14	15	16	17	18	19	20	21	22	23	24
$t_{i,2}^{a(0)}/h$	0.196	—	—	—	—	—	0.195	0.206	—	—	0.197	0.187
$C_2^a(\text{veh}\cdot h^{-1})$	900	—	—	—	—	—	600	700	—	—	900	700

TABLE 2: Relevant data of different travel groups.

Travel Groups	Percentage of total travel	Car ownership ratio	Time value $\tau_i$ (yuan/h)
High income group	25%	90%	25
Medium income group	40%	45%	16
Low income group	35%	5%	6

been considered as the most suitable simulation algorithm for solving multiobjective problems [21]. The specific steps of the algorithm are as follows.

*Step 1.* Algorithm parameters such as population size PopSize and evolutionary algebra GenMax were set; the car and bicycle subnet data were initialized; the initial public transport network population was generated randomly by 0-1 coding; the length of bit strings was determined by the number of bus lines searched by depth-first algorithm; for example: 001010001 means that there were 9 bus lines between origin and destination, and the lines 3, 5, and 9 were selected to form the public transport network.

*Step 2.* The users with cars in the three groups were divided into three travel modes, and the users without cars were divided into bus and bicycle travel modes by Logit model.

*Step 3.* For each individual in each generation (public transport network) and other mode subnets to achieve SUE traffic allocation, the MSA algorithm was used to solve 9 mode subnets at the same time. In addition, the upper objective function was calculated by using the distribution result, and the fitness function was the objective function itself.

*Step 4* (NSGA-II algorithm core operation). (1) The number of individuals was selected via the tournament strategy to determine the tournament population; (2) the cross-operation was performed by the OX-like method, and the exchange operation was performed on two variation points in the individual by the exchange method; (3) the upper objective function value of each individual in the new population was calculated after the merging of parent and child populations; (4) the fast undominated sort was made based on the value of the objective function, and the crowding distance of the individual in the population is calculated; (5) the former PopSize individuals were selected to generate a new generation population, according to the frontier order value and the crowding distance of each individual.

*Step 5* (the termination condition was determined). If the maximum number of iterations is reached, the algorithm terminates, producing a Pareto optimal solution set. Otherwise, turn to Step 2.

#### 4. Example Analysis

Taking the simple road network shown in Figure 1 as an example, the optimization model and algorithm of public transport network are verified. Due to space constraints, this paper lists certain initial data of the public transport network, as shown in Table 1. It is assumed that the road network between origin and destination has sufficient capacity to allow multiple travel lines in the network without considering the transfer mode and the interests of the operators. In this paper, the existing two bus lines are optimized, and direct bus lines are rearranged between origin and destination, in order to improve the fairness of the public traffic line network scheme.

Take  $q^w = 8000 \text{ people}\cdot h^{-1}$ ,  $\lambda=0.1$ ,  $\theta^1=1$ ,  $\theta^2=2$ ,  $\theta^3=5$ . The relevant data of different travel groups are shown in Table 2, and the relevant data of traffic modes are shown in Table 3.

In the NSGA-II algorithm, the population size is 100, the crossover rate is 0.8, the mutation rate is 0.1, and the evolutionary algebra is 500. The distributions of the Pareto optimal solution set and other solution set are shown in Figure 4. Table 4 lists the six groups of noninferior solutions obtained by NSGA-II algorithm. Each group of noninferior solution corresponds to the optimization scheme of public transport network under traffic fairness constraints. The decision makers select the best scheme among these six noninferior solutions, considering the construction and operation cost, the service level, and other factors of bus lines. The optimization calculation results are shown in Figures 5 and 6. Through the analysis of six groups of data related to optimization scheme, the following conclusions are drawn:

For the three travel groups, the travel cost relative deprivation coefficient is 33.42, and the road area Gini coefficient is 0.248 before the optimization of the public

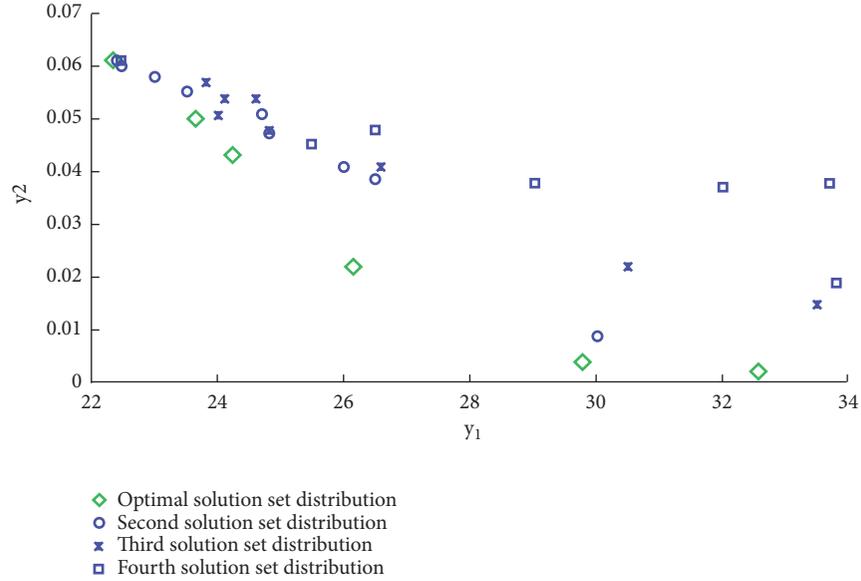


FIGURE 4: Distribution of Pareto optimal solutions and other solutions.

TABLE 3: Relevant data of different traffic modes.

Traffic Modes	Car unit conversion factor $E_k$	Average number of passengers per vehicle $N_k$ (person / vehicle)	Vehicle usage fee $L_k$ (yuan)	Road area occupied by vehicle $S_k$ ( $m^2$ )
Car	1.00	2	12	30
Bus	1.50	30	2	45
Bicycle	0.25	1	0	7.5

Note: due to the particularity of the network structure in the example, the lengths of all travel routes between origin and destination are equal, and the transportation cost of the car is set as a fixed value.

transport network; after optimization, as shown in Table 4, the travel cost relative deprivation coefficient of the six optimized schemes averages 26.51, which is 20.68% lower than that before optimization; the Gini coefficient of road area averages 0.030, which is 87.76% lower than before. Thus, the model optimized the public transit network effectively. While meeting the travel needs of residents, it shortened the difference in travel costs between different groups and improved the fairness of transportation resource allocation.

Before the public transit network optimization, a large number of transportation resources were occupied by high income groups to meet the needs of private car travel. In contrast, low-income groups without car purchasing power only chose buses with less comfort, convenience, and accessibility. After optimization, the preferences of the three groups changed significantly. As can be seen from Figure 5, certain high and medium income groups who used to travel by car begin to choose bus and bicycle modes, while the low-income group choose a higher rate of public transportation. Besides, both car and bicycle usage rates are declining. This trend indicates that the optimized public transport network has significantly increased the attraction for each group. The

reason is known from Figure 6. It is mainly because the travel time of all groups by car has increased, the travel time by bicycle is basically the same, and the travel time by bus has decreased, after the public transit network optimization. The change of travel time not only restrains the demand of each group to choose cars, but also improves the convenience and accessibility of public transportation, which makes more travelers willing to choose bus travel.

### 5. Conclusions

(1) The relative deprivation coefficient of travel cost described the influence of the generalized travel cost difference of different groups on transportation equity, while the Gini coefficient of road area reflected the equilibrium of traffic resources distribution among different groups. In this paper, the bilevel optimization model of public transport network was established under the condition of traffic fairness constraint, and the relationship between the public transit network optimization and travel opportunities of different groups was discussed. The upper level optimization aimed at minimizing the travel cost relative deprivation coefficient and the road area Gini coefficient, and the lower level

TABLE 4: Calculation results.

Optimization	Bus lines	Target value		Optimization	Bus lines	Target value	
		$y_1$	$y_2$			$y_1$	$y_2$
1	1→2→3→4→8→12→16	22.35	0.061	4	1→2→6→7→11→15→16	26.13	0.022
	1→2→6→10→11→15→16				1→5→6→7→8→12→16		
	1→5→9→10→14→15→16				1→5→6→7→11→12→16		
	1→5→9→13→14→15→16						
2	1→2→6→7→11→15→16	23.64	0.050	5	1→2→6→7→8→12→16	29.78	0.004
	1→5→6→10→11→12→16				1→5→6→7→11→12→16		
	1→5→9→10→14→15→16				1→5→6→10→11→15→16		
					1→5→9→13→14→15→16		
3	1→2→6→7→11→12→16	24.24	0.043	6	1→5→6→10→11→12→16	32.70	0.003
	1→5→6→7→8→12→16				1→2→3→4→8→12→16		
	1→5→6→10→11→12→16				1→5→9→13→14→15→16		
	1→5→9→10→11→15→16						
	1→5→9→13→14→15→16						

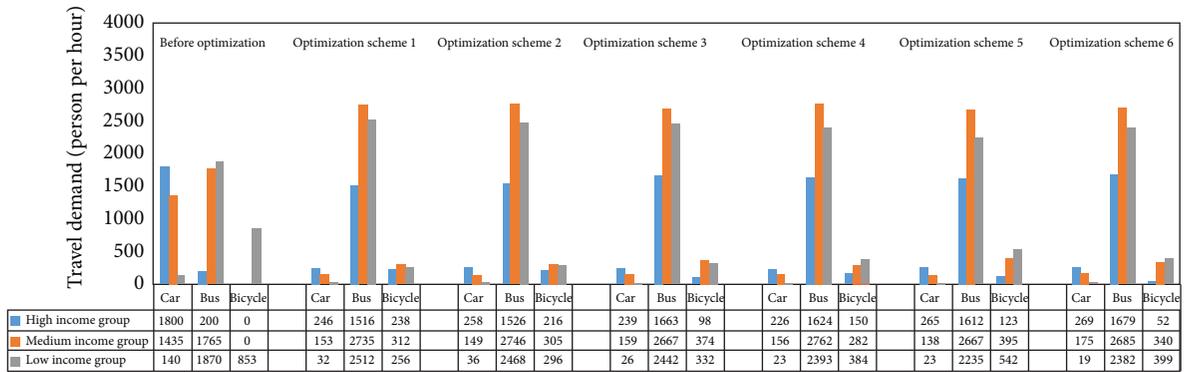


FIGURE 5: Compared the different traffic mode ridership.

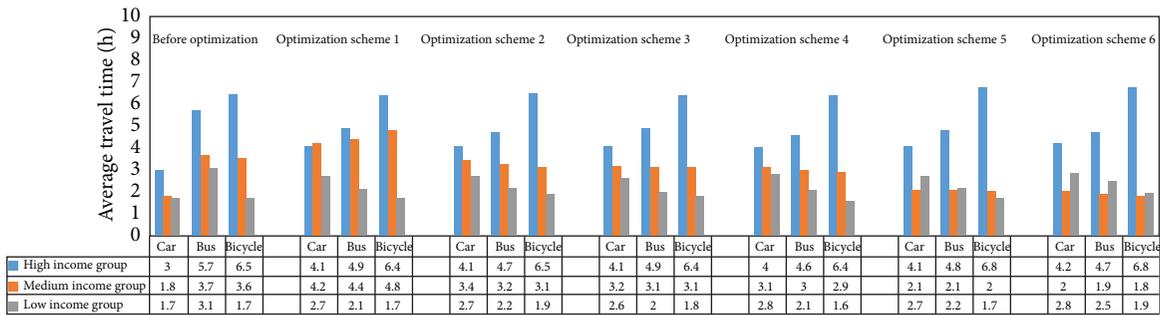


FIGURE 6: Compared the different traffic mode travel time.

optimization was a stochastic equilibrium traffic assignment model of multimode and multiuser. Furthermore, the complex selection behaviors of different groups to different travel modes were analyzed.

(2) The example analysis showed that the public transport network model considering the traffic fairness constraints shortened the travel cost difference among different groups effectively and distribute road resources more equitably on one hand; it guided all travel groups to

choose more bus trips, improve the fairness of the public transit network optimization schemes, and reduced the chances of high income groups choosing cars on the other hand.

### Variables

See Table 5.

TABLE 5: Variable list description.

Symbol	Symbolic Meaning	Subscript and Superscript Meaning
<i>1 Urban traffic network of multi-mode and multi-user</i>		
$G_{ik}$	the network formed by the $i$ -th group choosing the $k$ -th traffic mode	the $i$ -th group; traffic mode $k$
$q^w$	the total travel amount of OD pair $w$	Any OD pair $w$
$q_i^w$	the travel demand of the $i$ -th group OD to $w$	the $i$ -th group
$q_{i,k}^w$	the travel demand of the $i$ -th group on OD pair $w$ choosing the traffic mode $k$	traffic mode $k$
$x_k^a$	the travel demand on the segment $a$ in the subnet $G_{ik}$	segment $a$ ; the $i$ -th group; traffic mode $k$
$J_{i,k}^{w,r}$	the travel demand of the OD pair $w$ on the route $r$ in the subnet $G_{ik}$	OD pair $w$ ; route $r$ ; the $i$ -th group; traffic mode $k$
$\delta_{i,k}^{w,r,a}$	Boolean variables associated with routes and segments	OD pair $w$ ; route $r$ ; segment $a$ ; the $i$ -th group; traffic mode $k$
$v_{i,k}^a$	the road segment traffic on the segment $a$ in the subnet $G_{ik}$	segment $a$ ; the $i$ -th group; traffic mode $k$
$E_k$	the equivalent car conversion coefficient of the traffic mode $k$	traffic mode $k$
$N_k$	the average number of passengers in the traffic mode $k$	traffic mode $k$
$h_{i,k}^{w,r}$	the road traffic on the route $r$	OD pair $w$ ; route $r$ ; the $i$ -th group; traffic mode $k$
$t_{i,k}^a$	the travel time on the segment $a$ in the subnet $G_{ik}$	segment $a$ ; the $i$ -th group; traffic mode $k$
$t_{i,k}^{a(0)}$	the free flow time on the segment $a$ in the subnet $G_{ik}$	segment $a$ ; the $i$ -th group; traffic mode $k$
$C_k^a$	the travel capacity of the $k$ -th travel mode for all groups on the segment $a$	segment $a$ ; traffic mode $k$
$d_{i,k}^{w,r}$	the travel time on the route $r$ in the subnet $G_{ik}$	OD pair $w$ ; route $r$ ; the $i$ -th group; traffic mode $k$
$p_{i,k}^{w,r}$	the probability that the $i$ -th travel group chooses the route $r$ in the subnet $G_{ik}$	OD pair $w$ ; route $r$ ; the $i$ -th group; traffic mode $k$
$\theta^i$	the familiarity of different travel groups with the road network	the $i$ -th group
<i>2 Model Construction and Solution</i>		
$\gamma_1$	the relative deprivation of travel costs incurred by all travelers in the multi-mode transportation network	Serial number
$H_{i,k}^w$	the travel cost of selecting the $k$ -th way for the $i$ -th resident group on the OD pair $w$ in the subnet $G_{ik}$	OD pair $w$ ; the $i$ -th group; traffic mode $k$
$H_{j,k}^w$	the travel cost of selecting the $k$ -th way in the subnet $G_{ik}$ of the $i$ -th resident group on the OD pair $w$	OD pair $w$ ; the $j$ -th group; traffic mode $k$
$T_{i,k}^w$	the travel time of the $k$ -th mode in the sub- $G_{ik}$ of the $i$ -th resident group on the OD pair $w$	OD pair $w$ ; the $i$ -th group; traffic mode $k$
$\tau_i$	the time value of the $i$ -th resident group	the $i$ -th group
$L_k$	the cost of vehicle use for selecting the $k$ -th traffic mode	traffic mode $k$
$Q_{i,k}^w$	the probability that the $i$ -th group selects the traffic mode on OD pair $w$	OD pair $w$ ; the $i$ -th group; traffic mode $k$
$\lambda$	the correction parameter	
$U_1$	the ratio of low income group	Serial number
$U_2$	the ratio of medium income group	Serial number
$U_3$	the ratio of high income group	Serial number
$Z_1$	the ratio of road resources occupied by the low income group	Serial number
$Z_2$	the ratio of road resources occupied by the medium income group	Serial number
$Z_3$	the ratio of road resources occupied by the high income group	Serial number
$S_k$	the road area occupied by the single vehicle of the $k$ -th travel mode	traffic mode $k$
$N_k$	the standard passenger number of the single vehicle of the $k$ -th travel mode	traffic mode $k$
$M_2$	the area $M_2$ enclosed by the line of points O, $D_1$ , $D_2$ , A, B	Serial number
$M_1$	the area enclosed by the Lorenz curve and the absolute equality line	Serial number
$\gamma_2$	the road area Gini coefficient	Serial number
$\gamma_3$	The sum of travel time	Serial number

## Data Availability

The example data used to support the finding 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

The research is funded by the National Natural Science Foundation of China (no. 71462005; no. 71563023) and the Basic Competence Improvement Project for Young and Middle-Aged Teachers in Higher Education of Guangxi (no. 2018KY0341).

## References

- [1] T. Litman, "Evaluating transportation equity: guidance for incorporation distributional impact," *World Transport Policy & Practice*, vol. 8, pp. 50–65, 2002.
- [2] E. A. Vasconcellos, "Transport metabolism, social diversity and equity: the case of São Paulo, Brazil," *Journal of Transport Geography*, vol. 13, no. 4, pp. 329–339, 2005.
- [3] J. Preston and F. Rajé, "Accessibility, mobility and transport-related social exclusion," *Journal of Transport Geography*, vol. 18, no. 3, pp. 151–160, 2007.
- [4] Q. I. Ahmed, H. Lu, and S. Ye, "Urban transportation and equity: A case study of Beijing and Karachi," *Transportation Research Part A: Policy and Practice*, vol. 42, no. 1, pp. 125–139, 2008.
- [5] C. Nuworsoo, A. Golub, and E. Deakin, "Analyzing equity impacts of transit fare changes: Case study of Alameda-Contra Costa Transit, California," *Evaluation and Program Planning*, vol. 32, no. 4, pp. 360–368, 2009.
- [6] J. Bröcker, A. Korzhenevych, and C. Schürmann, "Assessing spatial equity and efficiency impacts of transport infrastructure projects," *Transportation Research Part B: Methodological*, vol. 44, no. 7, pp. 795–811, 2010.
- [7] X. Ying and J. Shi, "Road resources distribution and evolution analysis using a species competition model for improving road equity," *Tsinghua Science and Technology*, vol. 13, no. 5, pp. 651–659, 2008.
- [8] J. Farrington and C. Farrington, "Rural accessibility, social inclusion and social justice: Towards conceptualisation," *Journal of Transport Geography*, vol. 13, no. 1, pp. 1–12, 2005.
- [9] L. D. Olvera, D. Plat, and P. Pochet, "Household transport expenditure in Sub-Saharan African cities: measurement and analysis," *Journal of Transport Geography*, vol. 16, no. 1, pp. 1–13, 2008.
- [10] Y. Duvarci and S. Mizokami, "A suppressed demand analysis method of the transportation disadvantaged in policy making," *Transportation Planning and Technology*, vol. 32, no. 2, pp. 187–214, 2009.
- [11] K. Lucas, "Making the connections between transport disadvantage and the social exclusion of low income populations in the Tshwane Region of South Africa," *Journal of Transport Geography*, vol. 19, no. 6, pp. 1320–1334, 2011.
- [12] E. M. Ferguson, J. Duthie, A. Unnikrishnan, and S. T. Waller, "Incorporating equity into the transit frequency-setting problem," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 1, pp. 190–199, 2012.
- [13] M. Nikolić and D. Teodorović, "Transit network design by bee colony optimization," *Expert Systems with Applications*, vol. 40, no. 15, pp. 5945–5955, 2013.
- [14] P. N. Kechagiopoulos and G. N. Beligiannis, "Solving the urban transit routing problem using a particle swarm optimization based algorithm," *Applied Soft Computing*, vol. 21, pp. 654–676, 2014.
- [15] R. O. Arbex and C. B. da Cunha, "Efficient transit network design and frequencies setting multi-objective optimization by alternating objective genetic algorithm," *Transportation Research Part B: Methodological*, vol. 81, pp. 355–376, 2015.
- [16] K. An and H. K. Lo, "Two-phase stochastic program for transit network design under demand uncertainty," *Transportation Research Part B: Methodological*, vol. 84, pp. 157–181, 2016.
- [17] O. Dib, L. Moalic, M.-A. Manier, and A. Caminada, "An advanced GA-VNS combination for multicriteria route planning in public transit networks," *Expert Systems with Applications*, vol. 72, pp. 67–82, 2017.
- [18] 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.
- [19] L. Zhen, "Tactical berth allocation under uncertainty," *European Journal of Operational Research*, vol. 247, no. 3, pp. 928–944, 2015.
- [20] L. Zhen, E. P. Chew, and L. H. Lee, "An integrated model for berth template and yard template planning in transshipment hubs," *Transportation Science*, vol. 45, no. 4, pp. 483–504, 2011.
- [21] K. Deb, A. Pratap, S. Agarwal et al., "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

## Research Article

# Incorporating a Bayesian Network into Two-Stage Stochastic Programming for Blood Bank Location-Inventory Problem in Case of Disasters

Shutong Chen  and Changjun Wang 

*Glorious Sun School of Business and Management, Donghua University, No. 1882 West Yan'an Road, Shanghai 200051, China*

Correspondence should be addressed to Changjun Wang; [cjwang@dhu.edu.cn](mailto:cjwang@dhu.edu.cn)

Received 7 September 2018; Revised 27 January 2019; Accepted 11 February 2019; Published 7 March 2019

Academic Editor: Lu Zhen

Copyright © 2019 Shutong Chen and Changjun Wang. 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 is concerned with the optimal decisions of blood banks in a blood logistics network (BLN) with the consideration of natural disasters. One of the biggest challenges is how to deal with unexpected disasters. Our idea is to consider the disasters as the natural consequences of interaction among multiple interdependent uncertain factors, such as the locations and the levels of disasters, the number of casualties, and the availabilities of rescue facilities, which work together to influence the rescue effects of the BLN. Thus, taking earthquakes as the example, a Bayesian Network is proposed to describe such uncertainties and interdependences and, then, we incorporate it into a dedicated two-stage multi-period stochastic programming model for the BLN. The planning stage in the model focuses on blood bank location and inventory decisions. The subsequent operational stage is composed of multiple periods, some of which may suffer disasters and initiate corresponding rescue operations. Numerical tests show that the proposed approach can be efficiently applied in blood management under the complicated disaster scenarios.

## 1. Introduction

The supply of human blood relies on Blood Logistics Networks (BLNs), which are usually composed of donor points, blood banks, and relief facilities, etc. As the facilities to collect, test, inspect, store, and distribute human blood, blood banks play a fundamental role in BLNs to ensure the effective and efficient supply. Cooperating with other facilities in BLNs, blood banks should satisfy not only the daily demand of human blood but also the emergency demand caused by natural and anthropogenic disasters. Especially in recent years, frequent occurrences of disasters, such as earthquakes, hurricanes, fires, and terrorist attacks, have caused significant personnel and property losses due to its uncertain and destructive nature [3]. A well-designed blood bank in the BLN, which includes the prepositioning of emergency aid before a disaster strike and the supply of the human blood during a response process, can effectively overcome the suddenness of disasters.

However, the blood banks related decisions are always fraught with different challenges, in which the first and most important one is the availability of the disaster information that is indispensable to make the decisions. Taking the demand on human blood as example, the daily demand may be stable. However, the emergency demand in a disaster, especially the suddenly-occurring disaster such as the earthquake in natural disasters or terrorist attack, is hard to be estimated using historical data directly [4]. Besides the demand, other stochastic events in disasters, such as the place where the disaster occurs and facility failure, which also influence the relief effects significantly, would bring extra difficulties to decision-making. Moreover, the decisions on the blood banks include not only the strategic issues by considering the locations, but also the operational activities, such as inventory control and distribution. The interaction between the strategic level and the operational level requires an integrated model at the whole BLN level, taking the costs of facility location, transportation, and inventory into account

[5]. Finally, human blood has its distinctive characteristics, such as multiple categories (e.g., red blood cells, platelets, white blood cells, etc.) with different blood types (type A, B, AB, and O) and the mutual substitution between various types [6]. For example, type O could serve as the substitution of blood types A, B, and AB [7].

Some researchers have adopted the integrated framework to study relief logistics network taking disasters into account. However, few works have addressed blood banks with planning stage and multiple operational stages simultaneously as well as the features of blood products. More importantly, the uncertainties of disasters are dealt with in a relatively simple way in previous BLN-related studies. Specifically, besides daily demand, blood banks should also offer emergency aid to the hospitals in case of disasters. Demand uncertainty is a crucial challenge that must be handled in disaster management [8]. To handle this challenge, the uncertain demand for relief goods caused by the disasters is usually assumed available in many studies [9]. Such assumption may be too strict, especially for suddenly-occurring disasters like earthquakes. Actually, besides the uncertain demand, each disaster usually contains multiple sources of uncertainty, such as the magnitude of the disaster, the number of casualties, and the availability of rescue facilities, which should be considered when we design an efficient emergency plan [10]. It is worth noting that these factors are mutually interdependent. For example, the higher magnitude of earthquake in some places is, the more casualties there are, the higher failure probability a local hospital has, and finally the more human blood is required. Thus, it inspires us to consider more stochastic factors involved in disasters and focus on how these factors interact with each other and contribute together to the final disaster consequences. With the consideration of these random factors, the decisions we study would be more comprehensive and reasonable. However, it also means a dedicated mathematical programming model is required. For example, if the failure of rescue facilities, which often occurs in disasters, is considered, we should design the corresponding model which can handle this case and make appropriate decisions on blood allocation.

Therefore, our study mainly deals with the location-inventory decisions on blood banks and related decisions in the BLN with the consideration of possible occurrences of disasters. We propose a two-stage multi-period Stochastic Programming (SP) model which takes the characteristics of human blood (such as multiple blood products, blood lifespan, and blood substitution) and multiple stochastic factors involved in disasters into account. In the planning stage of the model, we determine the blood bank location and the inventory levels in the selected blood bank and each hospital in the BLN. The following operational stage involves multiple periods, in which the distribution and replenishment decisions on human blood would be made periodically. When a disaster occurs in some periods, the BLN should have the ability to continuously satisfy the emergency blood transfusion requirements at the first-time treatment by its inventory. The corresponding transfusion quantity, as one of the consequences of disasters, is decided by multiple disaster-related factors, such as the magnitude of the disaster, the

number of the injured, the level of injury, and the distribution of blood type among the injured. Besides, some region-related stochastic factors, such as the potential disaster points and the unavailable probability of each hospital, would decide which hospital provides continuous blood supply during a transfusion time, and, then, influence the prepositioning inventory-level decisions in this rescue hospital because there is a replenishment time due to its distance from the blood bank. Thus, we apply the Bayesian Network (BN), which is a probabilistic modelling approach in the context of data mining, to synthesize all these random factors to generate discrete scenarios which can be used to describe the possible disaster. The integrated objective is given to optimize the total cost of the strategic and operational decisions of the whole BLN during a planning horizon. A case study based on earthquakes in Sichuan Province, China, is studied and the corresponding model, which is a mixed-integer linear program, is optimally solved by IBM ILOG CPLEX to demonstrate the effectiveness of the proposed methods.

The contribution of this study lies in two aspects. Firstly, we propose a BN model to utilize the interdependences of multiple uncertain factors to generate plausible scenarios and incorporate it into a dedicated optimization model. It allows us to synthesize multiple stochastic elements in disasters and study their aggregated impacts on rescue management. To the best of our knowledge, a modelling effort for location and inventory in BLNs considering BN-based scenarios generation does not appear in existing studies. Our work is a new attempt to combine optimization method with the technique from data science to address a complicated decision-making issue in real-world. Secondly, we formulate an integrated SP model to handle with the location decision on blood banks and related inventory as well as distribution decisions in the BLN with possible disasters. The aforementioned features (e.g., coordination of location and operational activities in the whole BLN, the typical characteristics of human blood, and multiple uncertain factors considered in the BN) are considered in our model. In previous works on the BLNs, the consideration of some features of human blood (e.g., multiple blood products, blood substitution) is limited. Also, few BLN-related papers consider multiple interdependent uncertain factors, some of which are not studied as we know, such as the blood type among the injured and the level of injury. Thus, we take an early step to incorporate both multiple random factors and unique characteristics of blood into blood management considering disasters.

The remainder of this study is presented as follows. Section 2 reviews some of the more recent and related articles on correlative areas. In Section 3, a two-stage multiperiod SP model is developed, and the corresponding BN is introduced to quantify the effect of multiple sources of uncertainty on scenarios. An application in earthquake crisis faced by Sichuan Province, China, is given in Section 4 and the sensitivity analysis is implemented to validate the proposed model. The Expected Value (EV) model is given in Section 5 to prove the value of the stochastic solution. The last section presents conclusions and remarks on some directions for future research.

## 2. Literature Review

In this section, the literature regarding relief logistics networks under disasters and the applications of Bayesian Networks (BNs) are unfolded as follows.

Recently, frequent-occurring disasters drive researchers to study relief logistics networks with the aim of mitigating the impacts of disasters. In a disaster, since two stages exist naturally, i.e., pre/postdisaster, most of the existing studies formulate two-stage models to study the related issues of relief logistics networks. Barbarosoğlu and Arda's work [11] is the pioneer study which models predisaster and postdisaster stages by a two-stage SP and plans the transportation of relief goods in case of disaster. Sequentially, some researchers focus on the joint location-inventory problems (e.g., [12–14]), location-distribution problems (e.g., [8, 15]), location-routing problems [12, 16], etc. under disasters based on the two-stage paradigm, in which multiple sources of disaster uncertainties like demand, transportation time, and facility disruption are considered. However, most existing studies handle these uncertainties independently. For example, Mete and Zabinsky [12] determine the location of storage points from possible warehouses and the required inventory levels for medical supplies in the first stage with the consideration of uncertain transportation time and demand in the second stage. A multiobjective SP model is proposed by Haghi et al. [13] for determining the locations of distribution centers (DCs) and health centers, as well as the inventory of relief goods stored by suppliers before a disaster. Random demand, the number of casualties, and uncertain transportation time are considered independently in scenarios. Mohamadi and Yaghoubi [15] address the location of relief facilities in which several random factors, such as the number of injured people, the failure probabilities of roads, and the travel time, are also separately considered.

Besides, only a few authors have incorporated the interdependences between random disaster-related factors within humanitarian operations. For example, Verma and Gaukler [17] study the location decision by two-stage SP. Their contribution is a distance-damage function, which is used to model the damage caused by the disaster, and thus corresponding scenarios could be constructed. To determine the locations and capacities of DCs, Paul and MacDonald [10] treat the magnitude as an original stochastic factor, making the facility damage and casualty losses as a function of the magnitude of the earthquake. Similar to work studied by Mete and Zabinsky [12], the prepositioning of relief goods in logistics networks is addressed in Alem et al. [8]. The set of scenarios in the second stage of the model is defined as the combination of four factors, i.e., demand, supply, donation, and route damage, totaling 40 scenarios. The generation procedure of scenarios is based on a classification of the magnitudes of disasters, historical data, and the related assumptions. For example, route damage is partially proportional to the number of victims. However, studies mentioned above deal with interdependences between random factors, tend to base on single original factor, such as magnitude of earthquake, and describe other factors, e.g., number of victims, facility damage, etc. as the simple proportional relationship with

the original factor. Until now, considering multiple uncertain sources of a disaster in location-inventory planning is still limited and lacks systematic method.

Moreover, it is worth noting that the studies above only focus on general relief goods. Due to the importance of human blood in disaster rescue, more and more studies focus on emergency BLN design, taking the special structure of blood supply network as well as distinctive features of blood products into account. For example, Kochan et al. [18] and Gunpinar and Centeno [19] focus on inventory decisions before disaster for a given BLN to mitigate the impacts of disasters, in which the uncertain demand and perishability of human blood are considered. Besides, facility location related decisions before disasters, mainly blood donor points and permanent blood facilities, as well as inventory decision in the BLNs in case of disasters, are also studied. For example, Fahimnia et al. [20] present a bi-objective SP model to determine the number of blood facilities in the preparedness stage and the decisions like transportation routing and blood collection in the response stage. Samani et al. [21] develop a multi-objective two-stage stochastic programming to optimize blood center location before disaster happens and assignment of donors, blood facilities, and demand zones after disaster. Shelf life of blood products and demand uncertainty are involved.

SP is extensively applied to the design of emergency logistics networks, including BLNs. However, in most previous studies, the stochastic scenario which is indispensable in SP is assumed to be known. Obviously, such assumption is not always realistic, especially in the case of disasters, due to their unpredictability. This motivates more researchers to utilize Robust Optimization (RO), which only requires the worst-case information instead of entire distribution, to study related problems. For example, Jabbarzadeh et al. [22] develop a robust model to investigate the design of the BLN, in which the location decisions of the permanent blood collection facilities must be made before a disaster, and adjustments in location and capacity of temporary facilities can be implemented with the consideration of the worst scenario. Khalilpourazari and Khamseh [23] extend the mathematical model presented by Jabbarzadeh et al. [22] to the bi-objective setting which aims to minimize total blood supply chain costs as well as total transportation time. Salehi et al. [24] present a robust two-stage stochastic model for the emergency BSN design, in which multiple blood products and the compatibility of the blood group are considered. Rahmani [25] proposes a robust model for the BSN design by using p-criterion technique to protect the solution against the risk of disruptions. Similar works can refer to Shishebori and Babadi [26], Zahiri et al. [27], Fereiduni and Shahanaghi [28], Ramezani and Behboodi [29], Samani and Hosseini-Motlagh [30], and Kamyabniya et al. [31]. These works indicate that RO can provide robust solutions to unexpected disasters. However, RO fails to capture some valuable information contained in available data and, then, sometimes achieves too conservative solutions [32, 33].

To clearly demonstrate the existing BLN research under the setting of disasters, Tables 1(a) and 1(b) show the features of model, decisions, and the formulation methods of them.

TABLE 1  
 (a) Characteristics of related literature in disaster management

Reference	Model features											
	Blood	Perishability	Blood substitution	Single	Multiple	Demand	Disaster point	Rank of disaster	Stochastic factors Severity of injury	Distribution of blood type	Facility failure	Interdependence
Barbarosöglu and Arda 2004				✓	✓	✓						
Mete and Zabinsky 2010				✓	✓	✓						
Jabbarzadeh et al. 2014	✓			✓		✓					✓	
Gunpinar and Centeno 2015	✓	✓	✓		✓	✓						
Tofighi et al. 2015				✓	✓	✓						
An et al. 2015				✓							✓	
Verma and Gaukler 2015				✓			✓		✓		✓	✓
Zahiri et al. 2015	✓	✓		✓		✓						
Shishebori and Babadi 2015	✓			✓							✓	
Alem et al. 2016				✓		✓						✓
Kochan et al. 2016	✓	✓		✓		✓						
Paul and MacDonald 2016				✓		✓			✓		✓	✓
Rezaci-Malek et al. 2016 [1]		✓		✓		✓						
Fereiduni and Shahanaghi 2016	✓	✓		✓		✓					✓	
Mohamadi and Yaghoubi 2017				✓		✓					✓	
Fahimnia et al. 2017	✓			✓		✓						

(a) Continued.

Reference	Model features											
	Blood	Perishability	Blood substitution	Product		Demand	Disaster point	Rank of disaster	Stochastic factors Severity of injury	Distribution of blood type	Facility failure	Interdependence
				Single	Multiple							
Ramezani and Behboodi 2017	✓			✓		✓						
Haghi et al. 2017	✓			✓	✓	✓					✓	
Salehi et al. 2017	✓		✓	✓	✓	✓					✓	
Khalilpourazari and Khamseh 2017	✓			✓							✓	
Samani et al., 2018	✓	✓		✓		✓						
Samani and Hosseini-Motlagh 2018	✓	✓		✓		✓						
Rahmani 2018	✓			✓		✓						
Kamyabniya et al. 2018	✓	✓		✓		✓						
Eskandari-Khanghahi et al. 2018 [2]	✓			✓		✓					✓	
This paper	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

(b) Decisions and methods of related literature in disaster management

Reference	Predisaster	Postdisaster	Single period	Multiple periods	Location	Decisions			Methods			
						Prepositioning	Inventory level	Distribution flow	SP	RO		
Barbarosöğlü and Arda 2004	✓	✓		✓			✓	✓		✓		
Mete and Zabinsky 2010	✓	✓		✓	✓		✓	✓		✓		
Jabbarzadeh et al. 2014	✓	✓		✓	✓		✓	✓				✓
Gunpinar and Centeno 2015	✓	✓		✓		✓				✓		
Tofighi et al. 2015	✓	✓			✓		✓	✓		✓		
An et al. 2015	✓	✓	✓		✓		✓	✓		✓		

(b) Continued.

Reference	Predisaster	Postdisaster	Single period	Multiple periods	Location	Decisions			Methods		
						Prepositioning	Inventory level	Distribution flow	SP	RO	
Verma and Gaukler 2015	✓	✓	✓		✓	✓				✓	
Zahiri et al. 2015		✓		✓	✓						✓
Shishebori and Babadi 2015	✓	✓	✓		✓						✓
Alem et al. 2016	✓	✓		✓		✓	✓			✓	
Kochan et al. 2016	✓	✓	✓			✓	✓			✓	
Paul and MacDonald 2016	✓	✓	✓		✓	✓				✓	
Rezaei-Malek et al. 2016	✓	✓		✓	✓	✓					✓
Fereiduni and Shahanaghi 2016	✓	✓		✓	✓						✓
Mohamadi and Yaghoubi 2017	✓	✓	✓		✓					✓	
Fahimnia et al. 2017	✓	✓		✓	✓					✓	
Ramezani and Behboodi 2017		✓		✓	✓						✓
Haghi et al. 2017	✓	✓	✓		✓					✓	
Salehi et al. 2017	✓	✓		✓	✓		✓			✓	✓
Khalilpourazari and Khamseh 2017	✓	✓		✓	✓					✓	✓
Samani et al. 2018	✓	✓		✓	✓					✓	
Samani and Hosseini-Motlagh 2018	✓	✓		✓	✓						✓
Rahmani 2018	✓	✓		✓	✓						✓
Kamyabniya et al. 2018		✓		✓							✓
Eskandari-Khanghahi et al. 2018	✓	✓		✓	✓						✓
This paper	✓	✓		✓	✓		✓				✓

In summary, the rescue effects of BLNs are significantly influenced by the consequences of disasters, which embody in multiple aspects, including the emergency demand, facility failure, and the place where the disaster occurs. In previous studies, such consequences are treated in a relatively simple way. Specifically, most works only consider the random demand. Also, SP-based studies usually assume there is a known discrete distribution about the uncertain demand, etc. RO, which does not require a probabilistic distribution, only considers extreme scenarios and thus may lead to conservative decisions. Recently, only few BLN-related works, such as Fereiduni and Shahanaghi [28], Fahimnia et al. [20], and Salehi et al. [24], notice the importance of considering multiple disasters-related factors. However, the uncertainties considered in these works are still limited and their relationships are also ignored. Actually, if we can explore the internal relationships between multiple stochastic factors involved in disasters, it would be possible to give a precise description of the consequences of disasters and then, facilitate the application of sophisticated mathematical programming methodologies, such as SP. Thus, what we need is a quantitative method which can systematically formulate the relationship between multiple random factors in disasters.

In recent years, BNs have become a popular method for extracting knowledge from data in complex and uncertain systems, which encode the joint probability distribution of a set of random variables by making conditional dependence assumptions [34]. As suggested by Galindo and Batta [9], the Bayesian approach could be suitable for incorporating additional observations into stochastic operations management. Till now, BN has been applied to evaluate the occurrence probability and consequence of the accident in different areas, such as the reliability of semiconductor manufacturing [35], mine water inrush [36], navigation safety [37], gas explosion [38], and tropical cyclones [39]. But, its applications in operation research are still rare, not to mention BLNs.

These works indicate that BNs can be used to describe complicated disaster scenarios and improve the quality of decisions. However, it is worth noting that different types of disasters involve distinctive stochastic factors and their relationships are also diverse. Fortunately, for some common natural disasters, such as earthquakes and hurricanes, there are lots of studies in the context of disasters research which can provide solid foundation for our BN modelling. Moreover, with the consideration of more stochastic factors, we need to consider more decisions of BLNs, which means a dedicated model is required. Thus, in the next section, we describe the problem we studied, give the corresponding mathematical formulation and then, propose our BN for the case of earthquake.

### 3. Problem Definition and Model Formulation

**3.1. Problem Statement.** Consider a region where a government plans to launch a central blood bank to satisfy both the daily demand and the uncertain demand caused by possible future disasters. Hence, the location-inventory problem is proposed in a three-layer BLN composed by donor points, a

central blood bank, and hospitals, taking uncertain disasters and the characteristics of human blood into account. After the blood bank is located, the donor points collect blood from donors and provide blood to the prespecified blood bank, which is responsible for distributing blood to hospitals. The operational stage is divided into multiple periods (time interval of each period is  $T$ ), some of which suffer suddenly-occurring disasters. The activities mentioned above repeatedly happen in each daily period. When a disaster occurs in a period, the casualties are transferred from the disaster point to the rescue hospital which is the closest available one. Then, the rescue hospital and the prespecified blood bank should satisfy the emergency demand at the first-time treatment by their prepositioning inventory without interruption. Obviously, this is critical to the medical relief for the victims. The objective is to minimize the total cost of the BLN while ensuring that the blood demands are met under both daily and disaster situations in a given planning horizon.

The problem consists of four different decisions: (1) the central blood bank location; (2) inventory level in the prespecified blood bank and each hospital at the beginning of each operational period; (3) the distribution flow from donor points to the prespecified blood bank at each operational period; and (4) the distribution flow from the prespecified blood bank to each hospital at each operational period. The decisions are made in two stages. Taking the BLN composed by three donor points, two central blood bank candidates and two hospitals as the example, the planning stage includes the strategic decisions on the selection of a central blood bank from two candidates and the inventory level of both the selected blood bank and each hospital. Given these strategic decisions, the activities in one period of the operational stage under the daily or disaster setting are given respectively in Figure 1.

The daily operations of the whole BLN are displayed in Figure 1(1). In each period, the donor points collect the blood as the request of the central blood bank, and then the central blood bank implements blood test, inspection, etc., and replenishes each hospital by using its inventory at the same time.

More specifically, the inventory status of each facility in Figure 1(1) can be further illustrated by Figure 2 which shows the change of the inventory level of each corresponding facility.

As shown in the right part in Figure 2, the shipments from donor points to the blood bank and from the blood bank to each hospital should begin at the beginning of each period. Actual arrival time is dependent on the corresponding transportation distance. The in-transit inventory is a part of the inventory level of the corresponding receiver which also accounts for a portion of its holding cost. The right-upper part of Figure 2 shows that three donor points collect the different quantity of blood in each period which composes the total blood supply. Moreover, as shown in the right-middle part of Figure 2, the inventory of blood in the prespecified blood bank consists of the daily inventory and the emergency stock. The former is used to meet the daily deterministic demand from hospitals at each daily period,

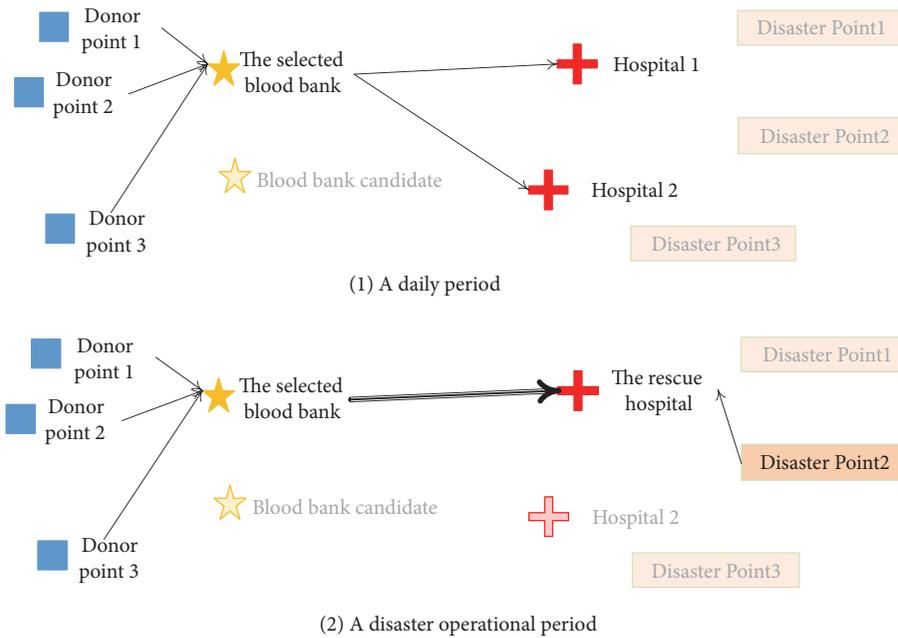


FIGURE 1: The illustration of activities in one period of the operational stage.

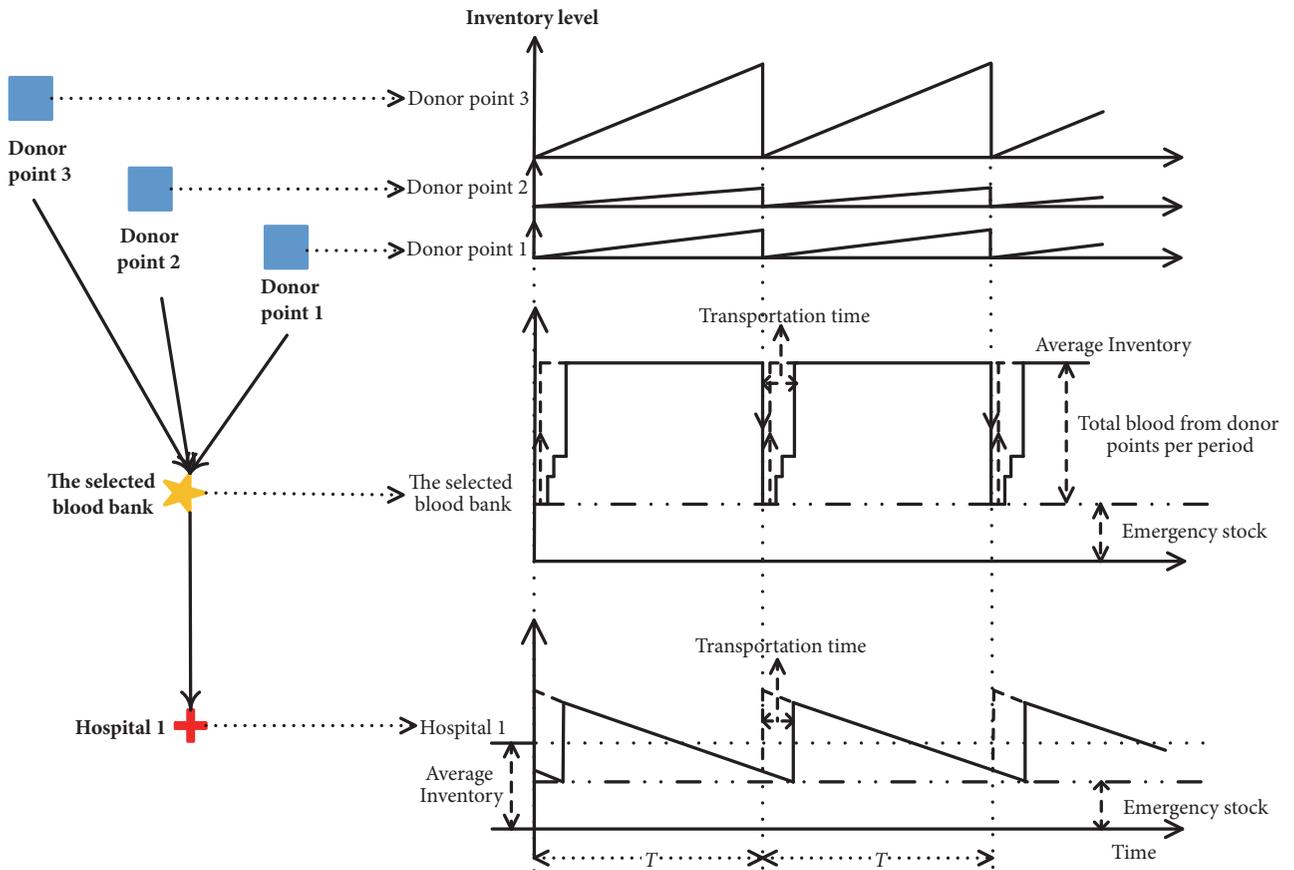


FIGURE 2: The status of inventory level in three-layers BLN in the operational stage.

and the latter is prepositioned to respond to the disaster demand. Both of them are decided at the planning stage. The total inventory level should be the sum of its physical inventory (denoted as the solid staircase lines) and the in-transit inventory (denoted as the dashed arrows). Finally, the inventory in each hospital (see the right-lower part of Figure 2) also includes the daily inventory and the emergency stock, in which the latter is prepositioned for the case that the casualties arrive at the hospital earlier than the delivery of blood from the prespecified blood bank. Similarly, the holding cost of each hospital is calculated based on the sum of its in-transit inventory (denoted as the dashed lines) from the prespecified blood bank and its physical inventory (denoted as solid lines).

Figure 1(2) illustrates the case that the disaster happens in some potential disaster points (e.g., disaster point 2), and then the casualties would be sent to the closest available hospital, namely, the rescue hospital (e.g., hospital 1) in some periods. Thus, the rescue hospital needs to use its inventory to satisfy the emergency demand caused by the disaster at first. To ensure the continuous transfusion, the prespecified blood bank also should have the ability to use its prepositioning inventory to replenish the blood to the rescue hospital before its inventory runs out. The continuous transfusion time in the first-time treatment is  $\Delta T$ , and the blood transfusion quantity is directly influenced by the consequences of the disaster, such as the number of the injured and the severity of the injury. These uncertain factors and the stochastic disaster points, as well as the availability of hospitals, would be directly or indirectly considered in the proposed model.

In summary, the BLN itself should have the ability to satisfy the first-time blood demand. The consumed inventory in the BLN can be recovered before the next period. Regarding the blood demand of subsequent treatments on the injured, it also can be satisfied by other neighbor regional BLNs or extra blood donations, which is not discussed in our study.

The main assumptions of the above problem are listed as follows:

- (i) A fixed construction cost for each blood bank candidate is available.
- (ii) The holding cost of in-transit inventory from each donor point to the prespecified blood bank and the prespecified blood bank to each hospital will be included in the inventory cost of the blood bank and corresponding hospital, respectively.
- (iii) The inventory holding cost is proportional to the inventory duration and quantity.
- (iv) The transportation cost is proportional to the distance and distribution flow.
- (v) In each period, at most one disaster occurs.
- (vi) At least one hospital is available during the disaster period, and the unavailable hospital can recover in the next period.

The last two assumptions are reasonable in our study. First, the disasters that require emergency blood rescue usually would not occur frequently. Moreover, a hospital may

be unavailable because of surrounding road damage or the damage to the hospital itself in a disaster. However, even all hospitals are unavailable, in order to implement the relief, the government would restore the most slightly damaged one or build a temporary rescue facility nearby some hospital as soon as possible. Hence, it is reasonable to assume there is at least one available hospital. Furthermore, some hospitals that are seriously damaged may not be restored soon; however, the daily demand in the area covered by the damaged hospital still exists and may be guided to a nearby substituting facility in a short time. Thus, the similar daily operational activities still exist. Therefore, we give the last assumption.

### 3.2. Model Formulation

(a) *Model Formulation and Explanation.* As it was previously pointed out, the decisions are made in two stages. The planning stage includes the decisions on the location of a central blood bank and the inventory levels in both the prespecified blood bank and hospitals. A binary location decision variable  $x_i$  equals to 1 if the central blood bank candidate  $i$  is selected, or 0 otherwise. Moreover,  $C_{iab}$  represents the maximum inventory level of blood products  $a$  with type  $b$  in the candidate  $i$ .  $ss_{iab}$  and  $s_{hab}$  are emergency stock levels in the candidate  $i$  and the hospital  $h$ , respectively. The operational stage includes the distribution flow in both each daily period and disaster period. In each daily period, the distribution flow of blood products  $a$  with type  $b$  from the donor point  $k$  to the candidate  $i$  is represented as  $g_{kiab}$ ,  $w_{ihab}$  and  $w'_{ihabb'}$  are the distribution flows from the candidate  $i$  to the hospital  $h$ . The latter is the quantity of blood product  $a$  with type  $b'$  used to substitute type  $b$ . When a disaster occurs,  $g_{kiab}^s$  is used to represent the distribution flow of blood products  $a$  with type  $b$  from the donor point  $k$  to the candidate  $i$  under scenario  $s$ . Let  $y_{ihab}^s$  and  $y'_{ihabb'}$  be the distribution flow from the candidate  $i$  to the rescue hospital  $h$  with the consideration of blood substitution under scenario  $s$ .

Sets, parameters, and variables used are given in Tables 2–4. The mutual substitution relationship, i.e., variable  $q_{bb'}$ , of different blood types in set BLN, is given in Table 5 [7].

The objective of the two-stage multiperiod SP model aims at minimizing the total cost of the BLN in both the planning stage and operational stage, which can be divided into three terms: (1) the construction cost in the planning stage, (2) the operational cost of the daily period, and (3) the operational cost of the disaster period in the operational stage.

(1) *The Construction Cost in the Planning Stage.* It can be formulated as formula (1):

$$\sum_{i \in I} c_i \cdot x_i \quad (1)$$

(2) *The Operational Cost of the Daily Period in the Operational Stage.* The daily operational cost includes the inventory holding cost of the prespecified blood bank, the inventory holding cost of hospitals, the transportation cost of distribution flow from donor points to the prespecified blood bank, and the

TABLE 2: Sets and indices.

Symbol	Description
$K$	Set of donor points
$I$	Set of blood bank candidates
$H$	Set of hospitals
$A$	Set of blood products (plasma, platelets, et al.)
$BL$	Set of blood types (A, B, AB, O, et al.)
$S_1$	Set of disaster scenarios
$S$	Set of whole scenarios (union of $S_1$ and the scenario without a disaster)
$k$	Indices to donor points, and $k \in K$
$i$	Indices to blood bank candidates, and $i \in I$
$h$	Indices to hospitals, and $h \in H$
$a$	Indices to blood products, and $a \in A$
$b$	Indices to blood types, and $b \in BL$
$s$	Indices to scenarios, and $s \in S$

transportation cost of distribution flow from the prespecified blood bank to hospitals. Thus, the operational cost in one period can be calculated as

$$\begin{aligned}
& T \cdot \sum_{i \in I} \left[ r_i \cdot \left( \sum_{a \in A} \sum_{b \in BL} ss_{iab} + \sum_{k \in K} \sum_{a \in A} \sum_{b \in BL} g_{kiab} \right) \right] + T \\
& \cdot \sum_{h \in H} \left[ u_h \right. \\
& \cdot \left. \sum_{a \in A} \sum_{b \in BL} \left( \sum_{i \in I} \frac{x_i \cdot (T + t_{ih}) \cdot d_{hab}}{2} + s_{hab} \right) \right] + e \cdot v \quad (2) \\
& \cdot \sum_{i \in I} \sum_{k \in K} \left( t_{ki} \cdot \sum_{a \in A} \sum_{b \in BL} g_{kiab} \right) + e \cdot v \cdot \sum_{i \in I} \sum_{h \in H} \left( t_{ih} \right. \\
& \cdot \left. \sum_{a \in A} \sum_{b \in BL} \left( w_{ihab} + \sum_{b' \in BL, b' \neq b} w'_{ihabb'} \right) \right)
\end{aligned}$$

in which,  $\sum_{a \in A} \sum_{b \in BL} ss_{iab} + \sum_{k \in K} \sum_{a \in A} \sum_{b \in BL} g_{kiab}$  is the inventory level of blood bank candidate  $i$  which involves the inventory on-hand and the corresponding in-transit inventory. When candidate  $i$  is not selected, this level would be set to zero by the constraints later. Similarly,  $\sum_{a \in A} \sum_{b \in BL} (\sum_{i \in I} (x_i \cdot (T + t_{ih}) \cdot d_{hab}/2) + s_{hab})$  is the inventory level of hospital  $h$  in which  $\sum_{i \in I} (x_i \cdot (T + t_{ih}) \cdot d_{hab}/2)$  is the average cycle inventory per  $T$ .

(3) *The Operational Cost of the Disaster Period in the Operational Stage.* In the period which suffers a disaster, the operational cost includes four items, in which the first two items are the same as these in formula (2). The third one is the

transportation cost from the donor points to the prespecified blood bank in the disaster period. Then the last one is the transportation cost caused by the daily and emergency demand of the rescue hospital as well as the daily demand of other available hospitals. Hence, the costs in a disaster period under one scenario  $s$  can be represented as follows:

$$\begin{aligned}
& T \cdot \sum_{i \in I} \left[ r_i \cdot \left( \sum_{a \in A} \sum_{b \in BL} ss_{iab} + \sum_{k \in K} \sum_{a \in A} \sum_{b \in BL} g_{kiab} \right) \right] + T \\
& \cdot \sum_{h \in H} \left[ u_h \cdot \sum_{a \in A} \sum_{b \in BL} \left( \sum_{i \in I} \frac{x_i \cdot (T + t_{ih}) \cdot d_{hab}}{2} + s_{hab} \right) \right] \\
& + e \cdot v \cdot \sum_{i \in I} \sum_{k \in K} \left( t_{ki} \cdot \sum_{a \in A} \sum_{b \in BL} g'_{kiab} \right) + e \cdot v \quad (3) \\
& \cdot \sum_{i \in I} \sum_{h \in H} \left( t_{ih} \cdot L_h^s \right. \\
& \cdot \left. \sum_{a \in A} \sum_{b \in BL} \left( y_{ihab}^s + \sum_{b' \in BL, b' \neq b} y'_{ihabb'} \right) \right)
\end{aligned}$$

In our study, the scenario set in each operational period is denoted as  $S$ , which contains the disaster scenario  $s$  ( $\in S_1$ ) with the probability  $p^s$  and the normal scenario  $s_0$  ( $\in S - S_1$ ) with the probability  $p^{s_0}$ . Thus, we have  $p^{s_0} + \sum_{s \in S_1} p^s = 1$ . The disaster scenarios vary with the different disaster points, the ranks of disasters, the unavailability of each hospital, etc.

Notice that the first two cost items in formula (2) (or formula (3)) always exist in the operational stage, while the last two in formula (2) and formula (3) are associated with the corresponding probability. Then, the final objective function of the two-stage multiperiod SP model is formulated as

$$\begin{aligned}
\min Z = & \sum_{i \in I} c_i \cdot x_i + \Phi \cdot \left\{ T \cdot \sum_{i \in I} \left[ r_i \cdot \left( \sum_{a \in A} \sum_{b \in BL} ss_{iab} \right. \right. \right. \\
& \left. \left. \left. + \sum_{k \in K} \sum_{a \in A} \sum_{b \in BL} g_{kiab} \right) \right] + T \cdot \sum_{h \in H} \left[ u_h \right. \right. \\
& \left. \left. \cdot \sum_{a \in A} \sum_{b \in BL} \left( \sum_{i \in I} \frac{x_i \cdot (T + t_{ih}) \cdot d_{hab}}{2} + s_{hab} \right) \right] \right\} + \Phi \\
& \cdot \left\{ p^{s_0} \cdot e \cdot v \cdot \left[ \sum_{i \in I} \sum_{h \in H} \left( t_{ih} \right. \right. \right. \\
& \left. \left. \left. \sum_{a \in A} \sum_{b \in BL} \left( w_{ihab} + \sum_{b' \in BL, b' \neq b} w'_{ihabb'} \right) \right) \right] \right. \\
& \left. + \sum_{i \in I} \sum_{k \in K} \left( t_{ki} \cdot \sum_{a \in A} \sum_{b \in BL} g_{kiab} \right) \right] + \sum_{s \in S_1} \left[ p^s \cdot e \cdot v \right. \\
& \left. \left. \left. \cdot \left[ \sum_{i \in I} \sum_{h \in H} \left( t_{ih} \cdot L_h^s \cdot \sum_{a \in A} \sum_{b \in BL} \left( y_{ihab}^s + \sum_{b' \in BL, b' \neq b} y'_{ihabb'} \right) \right) \right] + \sum_{i \in I} \sum_{k \in K} \left( t_{ki} \cdot \sum_{a \in A} \sum_{b \in BL} g'_{kiab} \right) \right] \right] \right\}
\end{aligned} \quad (4)$$

Moreover, the constraints include the following three groups.

(1) *The Constraints in the Planning Stage.* In the planning stage, we would select one central blood bank among several candidates and make decisions on the maximum inventory of

TABLE 3: Model parameters.

	Symbol	Description
Deterministic parameters	$\Phi$	The number of periods in planning horizon
	$T$	Time interval of each period
	$e$	Unit transportation fee (RMB/ km•U, U is the abbreviation of blood Units. 1 Unit equals 200 ml)
	$v$	The velocity of transportation (km/h)
	$\alpha$	Demand service level (confidence level)
	$r_i$	Inventory holding cost per unit of blood bank $i$ (RMB/U•h)
	$c_i$	Fixed construction cost of blood bank $i$ (RMB)
	$u_h$	Inventory holding cost per unit of hospital $h$ (RMB/U•h)
	$t_{ki}$	Transportation time from donor point $k$ to blood bank $i$ (h)
	$t_{ih}$	Transportation time from blood bank $i$ to hospital $h$ (h)
	$dis_{nh}$	Distance from potential disaster point $n$ to hospital $h$ . (km)
	$l_a$	Lifespan of blood product $a$ (h)
	$f_{kab}$	Supply capacity of blood product $a$ with type $b$ in donor point $k$ (U)
	$d_{hab}$	Daily demand for blood product $a$ with type $b$ per hour in hospital $h$ (U/h)
	$q_{bb'}$	1 if blood type $b'$ is able to substitute type $b$ , 0 otherwise.
	$\Delta T$	The continuous blood transfusion time in the first-time treatment (h)
Stochastic parameters	$p^s$	Probability of scenario $s$ ( $\in S$ )
	$t_h^s$	Transportation time from the actual disaster point to hospital $h$ under scenario $s$ (h)
	$L_h^s$	1 if hospital $h$ is available to the injured under scenario $s$ , 0 otherwise.
	$z_h^s$	1 if the injured are sent to hospital $h$ under scenario $s$ , 0 otherwise.
	$d_{ab}^s$	Emergency demand for blood product $a$ with type $b$ under scenario $s$ (U/h)

TABLE 4: Decision variables.

	Symbol	Description
The planning stage	$x_i$	1 if blood bank candidate $i$ is chosen, 0 otherwise.
	$C_{iab}$	Maximum inventory of blood product $a$ with type $b$ in blood bank candidate $i$ (U)
	$ss_{iab}$	Emergency stock of blood product $a$ with type $b$ in blood bank candidate $i$ (U)
	$s_{hab}$	Emergency stock of blood product $a$ with type $b$ in hospital $h$ (U)
The operational stage	$g_{kiab}$	Supply quantity of blood product $a$ with type $b$ from donor point $k$ to blood bank candidate $i$ per period (U)
	$w_{ihab}$	Distribution flow of blood product $a$ with type $b$ from blood bank $i$ to hospital $h$ per period (U)
	$w'_{ihabb'}$	Distribution flow of blood product $a$ with type $b'$ used to substitute type $b$ from blood bank $i$ to hospital $h$ per period (U)
	$g^s_{kiab}$	Supply quantity of blood product $a$ with type $b$ from donor point $k$ to blood bank candidate $i$ under scenario $s$ (U)
	$y^s_{ihab}$	Distribution flow of blood product $a$ with type $b$ from blood bank $i$ to hospital $h$ under scenario $s$ (U)
	$y^s_{ihabb'}$	Distribution flow of blood product $a$ with type $b'$ used to substitute type $b$ from blood bank $i$ to hospital $h$ under scenario $s$ (U)

each blood product with different blood type. The constraints are represented as follows:

$$\sum_{i \in I} x_i = 1 \quad (5) \quad i \in I; a \in A; b \in BL \quad (8)$$

$$\max_{k \in K; i \in I; h \in H} (t_{ki} + t_{ih}) \cdot x_i \leq \min l_a, \quad a \in A \quad (6)$$

$$\sum_{a \in A} \sum_{b \in BL} C_{iab} \leq MM \cdot x_i, \quad i \in I \quad (7) \quad x_i \in \{0, 1\}, \quad i \in I$$

$$C_{iab}, ss_{iab}, s_{hab}, \geq 0, \quad i \in I; h \in H; a \in A; b \in BL \quad (9)$$

TABLE 5: The mutual substitution among common blood types (set  $BL$ ).

$q_{bb'}$ $b, b' (\in BL)$	$b$							
	O+	O-	A+	A-	B+	B-	AB+	AB-
O+	1	0	1	0	1	0	1	0
O-	1	1	1	1	1	1	1	1
A+	0	0	1	0	0	0	1	0
A-	0	0	1	1	0	0	1	1
B+	0	0	0	0	1	0	1	0
B-	0	0	0	0	1	1	1	1
AB+	0	0	0	0	0	0	1	0
AB-	0	0	0	0	0	0	1	1

Constraint (5) ensures only one central blood bank would be chosen from all blood bank candidates in a region. Constraints (6) require that the time of transportation from any donor point to any hospital should be within the shortest lifespan of corresponding blood product.  $MM$  in (7) is a number big enough. Thus, inequalities (7) ensure that the corresponding maximum inventory of unselected blood bank candidate is zero. Constraints (8) enforce that the total quantity of blood provided by donor points within each  $T$  and the emergency stock is less than the maximum inventory in corresponding blood bank. Constraints (9) specify the domains of the decision variables.

(2) *The Constraints of Daily Periods in the Operational Stage.* In the daily operational period, the distribution flow should not be more than the quantity supplied by donor points and should meet the demand of all hospitals. The constraints are shown as follows:

$$0 \leq \sum_{i \in I} g_{kiab} \leq f_{kab}, \quad k \in K; a \in A; b \in BL \quad (10)$$

$$\sum_{h \in H} \left( w_{ihab} + \sum_{b' \in BL, b' \neq b} q_{bb'} \cdot w'_{ihab'b} \right) \leq \sum_{k \in K} g_{kiab}, \quad (11)$$

$$i \in I; a \in A; b \in BL$$

$$x_i \cdot (T + t_{ih}) \cdot d_{hab} \leq w_{ihab} + \sum_{b' \in BL, b' \neq b} q_{bb'} \cdot w'_{ihabb'}, \quad (12)$$

$$i \in I; h \in H; a \in A; b \in BL$$

$$x_i \in \{0, 1\}, \quad i \in I$$

$$g_{kiab}, w_{ihab}, w'_{ihabb'} \geq 0, \quad (13)$$

$$i \in I; h \in H; a \in A; b, b' \in BL$$

Constraints (10) express the maximum supply quantity in each donor point. Constraints (11) specify that the quantity of blood supplied to hospitals should be less than the quantity provided by all donor points within each  $T$ . Constraints (12) ensure that the blood delivery to hospitals should meet the demand during the sum of  $T$  and the transportation time with

the consideration of mutual substitution among blood types. Constraints (13) specify the domains of the decision variables.

(3) *The Constraints of Disaster Periods in the Operational Stage.* When a disaster happens, the injured are sent to the rescue hospital which should meet both the daily and emergency blood demand by its inventory at the first time. Besides, the subsequent blood should be supplied by the central blood bank to ensure continuous transfusion. Other available hospitals still need to satisfy their own daily demand. Thus, the constraints are formulated as

$$\sum_{h \in H} L_h^s \cdot \left( y_{ihab}^s + \sum_{b' \in BL, b' \neq b} q_{bb'} \cdot y'_{ihab'b} \right) = \sum_{k \in K} g_{kiab}^s, \quad (14)$$

$$s \in S_1; i \in I; a \in A; b \in BL$$

$$\sum_{h \in H} L_h^s \cdot \left( y_{ihab}^s + \sum_{b' \in BL, b' \neq b} q_{bb'} \cdot y'_{ihab'b} \right) \leq \sum_{k \in K} g_{kiab} + s s_{iab}, \quad (15)$$

$$s \in S_1; i \in I; a \in A; b \in BL$$

$$\frac{s_{hab}}{d_{ab}^s} + t_h^s \geq x_i \cdot z_h^s \cdot t_{ih}, \quad (16)$$

$$s \in S_1; h \in H; i \in I; a \in A; b \in BL$$

$$y_{ihab}^s + \sum_{b' \in BL, b' \neq b} q_{bb'} \cdot y'_{ihabb'} \geq x_i \cdot (T + t_{ih}) \cdot d_{hab} \cdot L_h^s, \quad (17)$$

$$s \in S_1; i \in I; h \in H; a \in A; b \in BL$$

$$P \left\{ \sum_{i \in I} \left[ \left( y_{ihab}^s + \sum_{b' \in BL, b' \neq b} q_{bb'} \cdot y'_{ihabb'} - x_i \cdot L_h^s \cdot (T + t_{ih}) \cdot d_{hab} \right) + s_{hab} \geq \Delta T \cdot d_{ab}^s \right] \geq z_h^s \cdot \alpha, \quad (18)$$

$$s \in S_1; h \in H; a \in A; b \in BL$$

TABLE 6: Variables (nodes) and states in the BN.

Variables (Nodes)		States	Description
Set	Name		
C	Potential epicenters	Point 1	Consider several potential epicenters and each with corresponding earthquake occurrence probabilities.
		...	
		Point N	
F	Level of earthquake	Level 1	Earthquake can be divided into different levels, and each of them has the corresponding probabilities.
		...	
		Level M	
G	Level of injury	Seriously	The level of injury of each injured person is subject to the two-point distribution. Each level of injury has corresponding probability and demand of blood.
		Slight	
D	Distribution of blood type	Distribution 1	The distribution of different blood types in the injured may be uncertain.
		...	
		Distribution J	
L	States of hospitals	Unavailable	Use unavailable probability, which is affected by the random events in C and F, to represent the state of hospitals.
		Available	

$$\sum_{s \in S_1} \left[ \sum_{k \in K} \sum_{a \in A} \sum_{b \in BL} g_{kiab}^{I^s} + \sum_{h \in H} \sum_{a \in A} \sum_{b \in BL} \left( y_{ihab}^s + \sum_{b' \in BL, b' \neq b} y_{ihabb'}^{I^s} \right) \right] \leq MM \cdot x_i, \quad i \in I \quad (19)$$

$$\sum_{i \in I} \sum_{a \in A} \sum_{b \in BL} \left( y_{ihab}^s + \sum_{b' \in BL, b' \neq b} y_{ihabb'}^{I^s} \right) \leq MM \cdot L_h^s, \quad (20)$$

$s \in S_1; h \in H$

$$x_i \in \{0, 1\}, \quad i \in I$$

$$g_{kiab}, s_{siab}, s_{hab}, g_{kiab}^{I^s}, y_{ihab}^s, y_{ihabb'}^{I^s} \geq 0, \quad (21)$$

$s \in S_1; i \in I; h \in H; a \in A; b, b' \in BL$

Constraints (14) represent the quantity of blood delivered from donor points to the prespecified blood bank in disaster period. Constraints (15) show the capacity of the prespecified blood bank should be able to deal with blood demand under the disaster scenario  $s$ . Constraints (16) guarantee that the blood replenished by the prespecified blood bank should be sent to the rescue hospital before its inventory runs out. The rescue hospital can be determined by formulas (22):

$$z_h^s = \begin{cases} \left\lfloor \frac{\max_{h \in H} t_h^s - t_h^s}{\max_{h \in H} t_h^s - \min_{h \in H} t_h^s} \right\rfloor, & \text{if } L_h^s = 1 \\ 0, & \text{if } L_h^s = 0, \end{cases} \quad (22)$$

$s \in S_1$

Moreover, constraints (17) and (18) show that the daily demand of available hospitals should be satisfied and the emergency demand should be probabilistically satisfied in

case of disasters, respectively.  $\alpha$  in the chance constraints (18) is the ratio of demand satisfaction, and constraints (18) can be transformed into corresponding deterministic linear constraints (23) and (24) [40]:

$$\sum_{i \in I} \left[ \left( y_{ihab}^s + \sum_{b' \in BL, b' \neq b} q_{bb'} \cdot y_{ihabb'}^{I^s} - x_i \cdot L_h^s \cdot (T + t_{ih}) \cdot d_{hab} \right) \right] + s_{hab} \geq z_h^s \cdot k_{ab}^s, \quad (23)$$

$h \in H; a \in A; b \in BL; s \in S_1$

where

$$k_{ab}^s = \sup \{ k_{ab}^s \mid k_{ab}^s = (F_{ab}^s)^{-1}(1 - \alpha) \}, \quad (24)$$

$s \in S_1; a \in A; b \in BL$

in which  $F_{ab}^s$  is the distribution function of  $\Delta T \cdot d_{ab}^s$ , and  $f_{ab}^s$  is the corresponding probability density function. A primary challenge in above modelling is to estimate  $F_{ab}^s$  and  $f_{ab}^s$ , which will be discussed in next section in detail.

Finally, constraints (19) guarantee that the distribution flows through unselected blood bank candidate are zero. Constraints (20) show that the distribution flows to unavailable hospitals in disaster periods are zero. Constraints (21) specify the domains of the decision variables.

**3.3. Bayesian Network (BN) and BN-Based Scenario Set.** In this section, based on earthquakes, we propose the corresponding BN model to generate the scenario set  $S$  used in our two-stage SP model. The variables used in the BN model can be referred to Table 6. Obviously, variables  $C$ ,  $F$ , and  $L$  vary with different earthquakes. Variables  $G$  and  $D$  would influence the final demand directly. Hence, it is natural to consider these factors in our BN model.

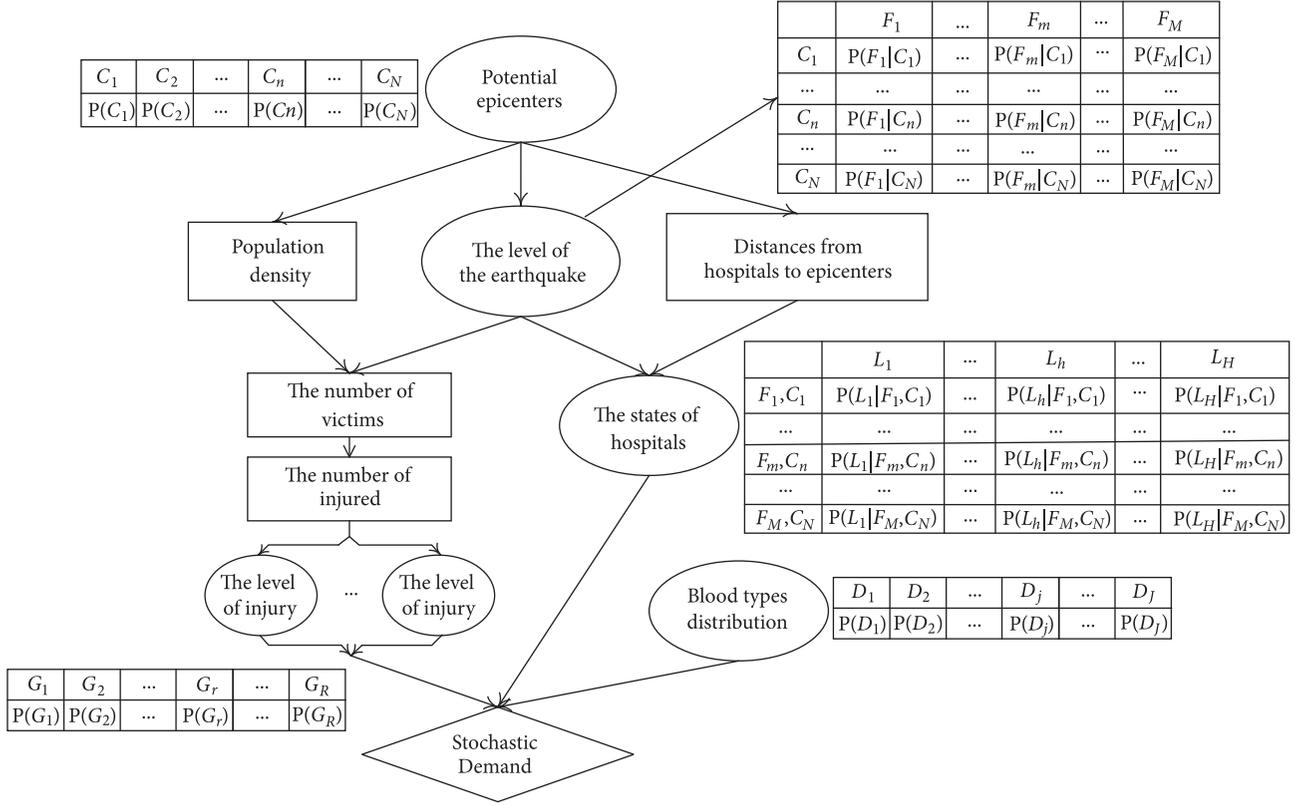


FIGURE 3: Proposed BN for the earthquake scenarios.

Thus, the proposed BN, as shown in Figure 3, consists of above five random variables (see the elliptical frame) contributing to stochastic blood demand (see the rhombic frame) by different intermediate variables (see the rectangular frame) for each blood product. The tables associated with these random variables in Figure 3 also show their possible states and corresponding possibilities.

Next, their inter-relationship can be quantified as follows.

(1) *The States of Hospitals.* That the hospitals may be unavailable in a disaster would obviously influence the effects of disaster relief. However, it is rarely considered in the BLNs related studies. Here, we assume the probability of the unavailability of a hospital is inversely proportional to its distance to the epicenter, being influenced by the level of the earthquake as well. This assumption is reasonable since the closer a hospital is to the epicenter, the higher the earthquake's level is, and the higher its unavailable probability is. Thus, the range normalization method is used to transfer the distance into the closed interval  $[0, 1]$ , taking the level of the earthquake into account. Hence, the probability of the unavailability of a hospital is the conditional probability of  $C$  and  $F$ :

$$p(L_h = 0 | F_m, C_n) = \max \left\{ 0, \frac{dis_{nh} - \max_{0 \leq \varepsilon \leq \vartheta_{F_m}} (\vartheta_{F_m} - \varepsilon)}{\min_{0 \leq \varepsilon \leq \vartheta_{F_m}} (\vartheta_{F_m} - \varepsilon) - \max_{0 \leq \varepsilon \leq \vartheta_{F_m}} (\vartheta_{F_m} - \varepsilon)} \right\},$$

$$= \max \left\{ 0, \frac{dis_{nh} - \vartheta_{F_m}}{0 - \vartheta_{F_m}} \right\},$$

$h \in H; F_m \in F; C_n \in C; n = 1, \dots, N$

(25)

in which  $C_n$  and  $F_m$  are used to represent the events of set  $C$  and  $F$ , respectively.  $\vartheta_{F_m}$  is the maximum affected epicentral distance calculated according to the earthquake level  $F_m$  and the given intensity, which results in facility or road damage [41]. Hence, we define the affected area as the circle with center point  $C_n$  and a radius of  $\vartheta_{F_m}$ .  $\varepsilon$  is the distance between epicenter and any point in the affected area. It is obvious that when the distance between epicenter and hospital is larger than  $\vartheta_{F_m}$ , the unavailable probability is zero. We use formula (26) to calculate the maximum epicentral distance ( $\vartheta_{F_m}$ ) when the level of the earthquake is  $F_m$  with the given intensity ( $\eta_0$ ) [42].

$$\eta_0 = 0.514 + 1.5F_m - 0.00659\vartheta_{F_m} - 2.014\lg(\vartheta_{F_m} + 10), \quad F_m \in F$$

(26)

(2) *The Stochastic Blood Demand in the Earthquake.* Notice that different potential epicenters may have distinctive possible levels of the earthquake. Thus, use  $P(F_m | C_n)$  to indicate the occurrence probability of level  $F_m$  in epicenter  $C_n$ , which is known before disasters. For a given  $F_m$ , Lee [43] presents the following formula to calculate the corresponding intensity, denoted as *Intensity*<sub>*m*</sub>.

$$\text{Intensity}_m = \frac{F_m - 1.5}{0.58}, \quad F_m \in F \quad (27)$$

Moreover, based on historical data in China, Ma [44] gives an empirical formula to evaluate the total number of affected victims (represented as  $NV$ ), which includes the injured and the fatal, by corresponding the intensity as follows:

$$\begin{aligned} \ln NV_{nm} = & -11.346 + 0.855 \ln F_m \cdot \text{Density}_n \\ & + 6.078 \ln \text{Intensity}_m, \quad F_m \in F, C_n \in C \end{aligned} \quad (28)$$

in which  $\text{Density}_n$  means the population density of epicenter  $C_n$ .

Then, with  $NV_{nm}$ , we can build its relationship with final stochastic blood demand by the following two steps.

At first, Wyss and Trendafiloski [45] use  $R$  to express the ratio of the injured to the fatal and calculate corresponding  $R$ s in some main areas and countries from historical statistics. For example, the value of  $R$  in China is 12.8. Thus, the number of the injured  $\Omega_{nm}$  can be obtained as follows.

$$\Omega_{nm} = NV_{nm} \cdot \frac{R}{1 + R} \quad (29)$$

Second, for each injured, the demand for blood products  $a$  per hour also could be stochastic, denoted as  $x_{1a}$  with the probability  $P(G_1)$  or  $x_{2a}$  with the probability  $P(G_2)$  which can be obtained by historical data. By Central Limit Theorem, the total demand for blood product  $a$  is approximately normally distributed especially when  $\Omega_{nm}$  is big enough. Let  $X_a$  be the demand for blood product  $a$  of each injured person per hour, and the mean and variance of total blood demand are explained as formulas (30) and (31):

$$\begin{aligned} E(\Omega_{nm} X_a) &= \Omega_{nm} \cdot E(X_a) \\ &= \Omega_{nm} \cdot [P(G_1) x_{1a} + P(G_2) x_{2a}] \end{aligned} \quad (30)$$

$$\begin{aligned} D(\Omega_{nm} X_a) &= \Omega_{nm} \cdot D(X_a) \\ &= \Omega_{nm} \cdot P(G_1) \cdot P(G_2) \cdot (x_{1a} - x_{2a})^2 \end{aligned} \quad (31)$$

The event that an earthquake occurs at potential epicenter  $C_n$  at level  $F_m$  constitutes a part of scenario  $s$ . Therefore, the demand for blood product  $a$  per hour under scenario  $s$  (denoted as  $d_a^s$ (U/h)) can be represented as

$$d_a^s \sim N(\Omega_{nm} E(X_a), \Omega_{nm} D(X_a)) \quad (32)$$

According to Peng [46] and Chen [47], the blood type distribution may differ from region to region. Also, in China, rapid economic development results in obvious human migration, making the blood type distribution in a specific

region fluctuate. Hence, for the people in a region, the blood type distribution is also stochastic. Thus, for each blood type distribution  $D_j \in D$ , denote  $p_{bj}$  as the proportion of blood type  $b$  in all the injured, and  $\sum_{b \in BL} p_{bj} = 1$ . Assume  $D_j$  occurs in scenario  $s$ ; thus, the corresponding emergency demand per hour for blood product  $a$  with type  $b$ , i.e.,  $d_{ab}^s$ , can be expressed as

$$d_{ab}^s = p_{bj} \cdot d_a^s \quad (33)$$

As we mentioned before, the blood prepositioned in the BLN needs to satisfy the blood demand of the transfusion time  $\Delta T$ . Thus, the emergency demand for blood product  $a$  with type  $b$  is  $\Delta T d_{ab}^s$ , and its probability density function  $f_{ab}^s$  used in formula (24) can be expressed as

$$f_{ab}^s = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(d_{ab}^s - \Delta T \cdot p_{bj} \cdot \Omega_{nm} E(X_a))^2}{2\Delta T^2 \cdot p_{bj}^2 \cdot \Omega_{nm} D(X_a)}\right), \quad (34)$$

$$s \in S_1; a \in A; b \in BL$$

When the hospital  $h$  is unavailable (i.e.,  $L_h=0$ ), the blood bank cannot replenish it at the present period. Hence, the states of hospitals will not change the emergency demand but would influence the final total demand.

Finally, one disaster scenario  $s$  composed by  $\{C_n, F_m, G_r, D_j, L_h\}$  and its impacts can be generated from the above procedure. Its corresponding probability is determined by the product of the probability of each random variable as

$$\begin{aligned} p^s &= P(C_n) \times P(F_m | C_n) \times P(G_r) \times P(D_j) \\ &\times \prod_{h=1}^H P(L_h | F_m, C_n), \quad s \in S_1 \end{aligned} \quad (35)$$

All disaster scenarios constitute disaster scenario set  $S_1$ . Although there are multiple hospitals and potential epicenters, we assume that at least one hospital is available and at most one point causes earthquake. Therefore,  $|S_1| = |C| \times |F| \times |G| \times |D| \times (2^{|H|} - 1)$ , in which the symbol  $|\bullet|$  is used to indicate the number of elements in set " $\bullet$ ". Notice, there is another scenario where no earthquake happens. Hence, the dimensionality of the whole scenario set  $S$  is  $|S| = |S_1| + 1$ .

## 4. Simulation

In this section, we implement simulation study based on earthquakes in Sichuan Province, China. In order to evaluate the impacts of the BN model, we would test and compare simulation cases under different scenario sets and parameters, such as the probability of disasters and the unit inventory holding cost.

### 4.1. The Construction of Disaster Scenario Sets Based on BN

4.1.1. *The Simulation Background.* We would introduce our simulation background based on the Longmenshan earthquake zone in Sichuan, China.

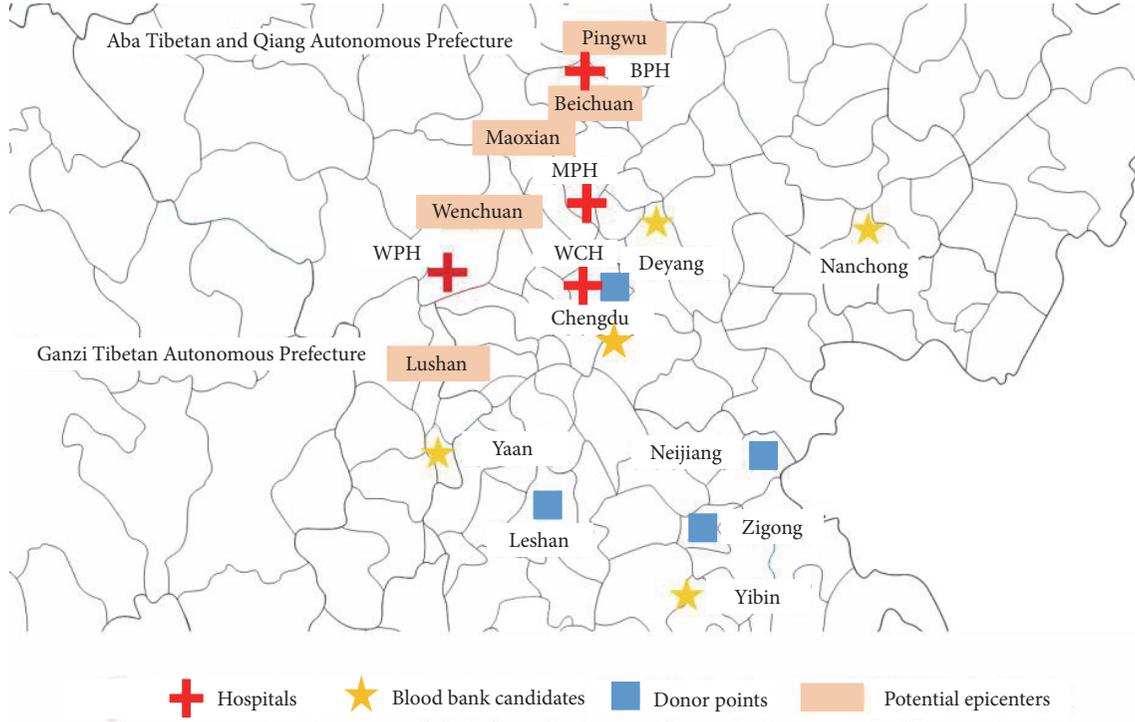


FIGURE 4: The geographic picture of blood bank supply chain.

TABLE 7: The parameters of the blood bank candidate  $i$ .

Blood bank candidates $i$	Fixed cost $c_i$ ( $10^7$ RMB)	Inventory holding cost $r_i$ (RMB/U·h)
Chengdu	1.2	0.012
Yibin	1.1	0.01
Nanchong	1.08	0.008
Yaan	1.05	0.006
Deyang	1.12	0.007

In the operational stage, each period  $T$  is set as a week. The planning horizon is 10 years (i.e.,  $\Phi \approx 520$  weeks). The transfusion time  $\Delta T$  is set to be 2(h), because the transfusion duration of each injured at one operation is about 2 hours according to clinical data. The ratio of demand satisfaction  $\alpha$  of blood supply is 95%.

The random epicenters, donor points, blood bank candidates, and hospitals are shown in the Figure 4.

Firstly, their parameters are elaborated below.

(1) Potential epicenters set  $C = \{Wenchuan, Beichuan, Maoxian, Lushan, Pingwu\}$  are all in the Longmenshan earthquake zone. The corresponding population densities are 0.0027, 0.0071, 0.0028, 0.0088, and 0.0031 (million/km<sup>2</sup>), respectively (see <http://data.stats.gov.cn/index.htm>).

(2) Let the blood bank candidates set  $I$  as  $\{Chengdu, Yibin, Nanchong, Yaan, Deyang\}$ , and each candidate's corresponding fixed cost and inventory holding cost are displayed in Table 7.

(3) Considering that the number of blood donors is positively correlated with the population density, top four cities in population density in Sichuan Province are designated as main blood donation points. Thus, by World Population Network [48], we have  $K = \{Chengdu, Neijiang, Zigong, Leshan\}$ .

(4) Set  $H$  is composed by hospitals which have superior medical conditions or are close to the random epicenters. Thus  $H = \{West\ China\ Hospital, Beichuan\ People's\ Hospital, Wenchuan\ People's\ Hospital, Mianzhu\ People's\ Hospital\}$  and is denoted as  $H = \{WCH, BPH, WPH, MPH\}$  for short. The unit inventory holding cost ( $u_h$ ) of the hospitals are (0.02, 0.015, 0.008, 0.01) RMB/U·h, respectively.

Secondly, the parameters of blood products are given as follows.

(1) The multiproduct set is denoted as  $A = \{Plasma, Red\ blood\ cells, Platelets\}$ , and the multitype set is denoted as  $BL = \{A, B, AB, O\}$ . Letters A, B, AB, and O are used to represent A+, B+, AB+, and O+ blood types and the negative cases are ignored because people with blood types A-, B-, AB-, and O- are rare. The mutual substitution relationship of different blood types can be referred to Table 5.

(2) By "The Minister of Health of the People's Republic of China [49]", the lifespans  $l_a$  ( $a \in A$ ) of plasma (thawed), red blood cells, platelets during transportation are set to be 1, 21, and 1 day, respectively.

(3) Table 8 gives the maximum quantity of blood that each donor point can supply per period  $T$  [50].

(4) By Wang et al. [51], the daily deterministic demand of each hospital is given in Table 9.

TABLE 8: The maximum supply quantity  $(f_{kab})(U)$  of donor point  $k$  per period  $T$ .

Donor point $k$	Plasma				Red blood cells			Platelets				
	A	B	AB	O	A	B	AB	O	A	B	AB	O
Chengdu	435	377	103	569	274	238	65	359	122	106	29	160
Neijiang	279	242	66	366	176	153	42	231	79	68	19	103
Zigong	305	264	72	399	192	167	46	252	86	74	21	112
Leshan	112	97	27	146	71	61	17	92	32	28	8	41

TABLE 9: The daily blood demand  $(d_{hab})(U/h)$  in hospital  $h$ .

Blood product $a$	Blood type $b$	Hospitals $h$			
		WCH	BPH	WPH	MPH
Plasma	A	3.050	0.213	0.081	1.191
	B	2.644	0.184	0.070	1.032
	AB	0.718	0.050	0.019	0.280
	O	3.007	0.281	0.107	1.562
Red blood cells	A	1.932	0.137	0.052	0.755
	B	1.675	0.118	0.045	0.655
	AB	0.455	0.032	0.012	0.178
	O	2.532	0.179	0.068	0.989
Platelets	A	0.829	0.058	0.022	0.324
	B	0.719	0.050	0.019	0.281
	AB	0.195	0.013	0.005	0.076
	O	1.087	0.076	0.029	0.425

TABLE 10: The transportation time  $(t_{ki})(h)$  from the donor point  $k$  to the blood bank candidates  $i$ .

Blood bank candidate $i$	Donor point $k$			
	Chengdu	Neijiang	Zigong	Leshan
Chengdu	0.15	2.5	3	2
Yibin	3.2	1.5	1	2
Nanchong	3	2.2	2.7	3.5
Yaan	1.8	2.8	2.8	1.5
Deyang	1.5	2.7	3	2.8

Furthermore, the parameters of transportation are given as follows:

(1) The transportation velocity  $v$  is assumed as an average value of 60 (km/h) in this paper, and the transportation fee  $e$  is set as 0.07 (RMB/km·U) according to Jabbarzadeh et al. [22].

(2) The transportation time from donor points to blood bank candidates and from blood bank candidates to hospitals and the distance from potential epicenters to hospitals, which can be obtained from Google Maps, are given in Tables 10–12, respectively.

4.1.2. *The Scenario Sets Based on BN.* Based on above parameters, we show how to construct scenario sets by the proposed BN model. Notice the final scenario set may vary due to the states and possibilities of the random factors. Hence, we only give one scenario set here. Other kinds of scenario sets can be given in similar way.

TABLE 11: The transportation time  $(t_{ih})(h)$  from the blood bank candidate  $i$  to the hospital  $h$ .

Hospitals $h$	Blood bank candidate $i$				
	Chengdu	Yibin	Nanchong	Yaan	Deyang
WCH	0.15	3.2	2.8	2	1.5
BPH	3	5.2	4.1	3.7	1.7
WPH	2.3	4.5	3.8	3	2.5
MPH	1.8	4	2.4	2.5	0.9

TABLE 12: The distance  $(dis_{nh})(km)$  from the epicenter  $n$  to the hospital  $h$ .

Epicenter $n$	Hospital $h$			
	WCH	BPH	WPH	MPH
Wenchuan	147	130	5	173
Beichuan	140	5	232	53.4
Maoxian	187	89.1	41.9	212
Lushan	170	305	261	241
Pingwu	273	124	368	211

(1) For  $C = (\text{Wenchuan, Beichuan, Maoxian, Pingwu, Lushan})$ , we give the probabilities of the earthquake as 0.03, 0.025, 0.02, 0.01, and 0.015, respectively. Thus, Table 13 could be obtained in which the corresponding probability of earthquakes in each single epicenter is shown. For example, the probability of a single-point earthquake in Wenchuan can be calculated as  $P(C_1) \times (1 - P(C_2)) \times (1 - P(C_3)) \times (1 - P(C_4)) \times (1 - P(C_5)) = 0.027953$ .

Next, since at most one epicenter at each period is assumed, the normalized results are displayed in Table 14.

(2) Let  $F = (8, 7, 6.5, 6)$  represent four levels, which are displayed in Table 15, and can be applied in calculating the intensity (refer to formulas (27)). Considering all potential epicenters are in the same earthquake zone,  $P(F_m | C_i)$  can be simplified to  $P(F_m)$ . The corresponding values for each  $F_m$  is 0.264, 0.189, 0.151, and 0.396, respectively [52].

(3) For each blood product,  $p_A, p_B, p_{AB}, p_O$  are used to represent the proportion of blood types A, B, AB, and O, respectively. Here, the blood type distribution, i.e.,  $p_A:p_B:p_{AB}:p_O$ , of injured people is considered to be random and has two kinds of blood type distribution, i.e.,  $D=(D_1, D_2)$ . Specifically, according to the mean blood type distribution in Sichuan Province [46], two equal-weighted distributions  $D_1=0.33:0.24:0.07:0.36$  and  $D_2=0.32:0.26:0.08:0.34$  are considered.

TABLE 13: The probability of single epicenter and no disaster.

	Epicenter					Probability
	Wenchuan	Beichuan	Maoxian	Lushan	Pingwu	
No earthquake						0.903803
	√					0.027953
		√				0.023174
Single-point earthquake			√			0.018445
				√		0.013763
					√	0.009129
Multipoint earthquake			2 <sup>5</sup> -5-1 cases			0.003733
<i>Sum of above probabilities</i>						1

TABLE 14: The normalization of probability.

	Epicenter					Probability
	Wenchuan	Beichuan	Maoxian	Lushan	Pingwu	
No earthquake						0.907
	√					0.028
		√				0.023
Single-point earthquake			√			0.019
				√		0.014
					√	0.009
<i>Sum of above probabilities</i>						1

TABLE 15: The levels of the earthquake and set  $F$ .

Rank	Magnitude	The level of the earthquake	Set $F$
I	≥7.0	Especially large earthquake	$F_1=8$
II	6.5~7.0	Large earthquake	$F_2=7$
III	6.0~6.5	Relatively large earthquake	$F_3=6.5$
IV	5.0~6.0	General earthquake	$F_4=6$

TABLE 16: The quantity of two-level demand for each blood product (per injured).

Blood product	Rank	Quantity(U/h)
Plasma	High	0.225
	Low	0.15
Red blood cells	High	1.136
	Low	0.522
Platelets	High	0.182
	Low	0.094

(4) Two demand levels, seriously and slightly, are assumed because the blood demand varies from person to person by injury level difference. Two random proportions,  $G=(G_1, G_2)$ , of two demand levels are considered, which are  $G_1=0.3:0.7$  and  $G_2=0.4:0.6$ . The probabilities of two random proportions (i.e.,  $P(G_r)$ ,  $r=1,2$ ) are both 0.5. Table 16 displays the corresponding quantities of blood demands [53, 54] for each level  $(x_{1a}, x_{2a}, a \in A)$ .

(5) Utilize formulas (26) to calculate the maximum epicentral distance with given  $\eta_0$  as VI [41]. Recall that at least one hospital is available in each scenario. Hence, the

conditional unavailable probability of each hospital under the combination of set  $C$  and  $F$ , which can be calculated by formulas (25), is shown in Table 17.

Till now, Figure 5 shows the scenario tree for each epicenter which is constructed according to parameters described above. The size of disaster scenarios set  $S_1$  is theoretically up to  $1200(= |S_1| = |C| \times |F| \times |G| \times |D| \times (2^{|H|}-1))=5 \times 4 \times 2 \times 2 \times (2^4-1)=5 \times 240$ , where the probability  $p_s$  ( $s \in S_1$ ) can be given by formulas (35). We denoted this dataset as 1200<sub>L</sub>.

Moreover, notice  $S_1$  varies with the random factors. In order to test the impacts of different  $S_1$  on the decisions, we consider more possible values of the proportions of two demand levels ( $G$ ) and the blood type distributions ( $D$ ). Specifically, the proportions  $G_3=0.5:0.5$ ,  $G_4=0.6:0.4$ ,  $G_5=0.7:0.3$  and the distribution  $D_3=0.318:0.25:0.082:0.35$  are considered. Thus, different scenario sets can be obtained as in Table 18. In the case of 1200 or 3600 scenarios, we only consider the extreme situations where emergency demand is extremely low ( $G_1, G_2$ ) or ( $G_1, G_2, G_3, G_4$ ) and high ( $G_3, G_4$ ) or ( $G_2, G_3, G_4, G_5$ ), respectively. Three combinations of blood type distribution are also incorporated into datasets with 1200 scenarios, resulting totally in 6 datasets. The case dataset  $|S_1| = 4500$  is the combination of all possible values of random variables we mentioned above; hence, there is only one corresponding dataset.

#### 4.2. The Computational Results and Sensitivity Analysis

4.2.1. Computational Results of All Scenario Sets. Assigning each dataset to the SP model which can be optimally solved by IBM ILOG CPLEX 12.6.3, the corresponding recourse

TABLE 17: The conditional unavailable probability of hospitals ( $P(L_h=0 | F_m, C_n)$ ).

Epicenter (Set C)	The level of the earthquake (Set F)															
	Rank IV (6)				Rank III (6.5)				Rank II (7)				Rank I (8)			
	WCH	BPH	WPH	MPH	WCH	BPH	WPH	MPH	WCH	BPH	WPH	MPH	WCH	BPH	WPH	MPH
Wenchuan	0	0	0.909	0	0	0	0.95	0	0.065	0.173	0.968	0	0.513	0.57	0.983	0.427
Beichuan	0	0.909	0	0.029	0	0.95	0	0.464	0.11	0.968	0	0.66	0.537	0.983	0.232	0.823
Maoxian	0	0	0.238	0	0	0.105	0.579	0	0	0.433	0.734	0	0.381	0.705	0.861	0.298
Lushan	0	0	0	0	0	0	0	0	0	0	0	0	0.437	0	0.136	0.202
Pingwu	0	0	0	0	0	0	0	0	0	0.211	0.978	0	0.096	0.589	0	0.301

TABLE 18: Characteristics of all datasets.

The dimensionality of the scenario sets	Dataset's name	Random factors	
		G	D
— $S_1$ —=1200	1200_1	$(G_1, G_2)$	$(D_1, D_2)$
	1200_2	$(G_1, G_2)$	$(D_1, D_3)$
	1200_3	$(G_1, G_2)$	$(D_2, D_3)$
	1200_4	$(G_4, G_5)$	$(D_1, D_2)$
	1200_5	$(G_4, G_5)$	$(D_1, D_3)$
	1200_6	$(G_4, G_5)$	$(D_2, D_3)$
— $S_1$ —=3600	3600_1	$(G_1, G_2, G_3, G_4)$	$(D_1, D_2, D_3)$
	3600_2	$(G_2, G_3, G_4, G_5)$	$(D_1, D_2, D_3)$
— $S_1$ —=4500	4500_1	$(G_1, G_2, G_3, G_4, G_5)$	$(D_1, D_2, D_3)$

TABLE 19: The decisions of maximum inventory and emergency stock of the blood bank.

The prespecified blood bank: Deyang							
Maximum inventory $C_{iab}(U)$	Plasma	A	4092.21	Emergency stock $ss_{iab}(U)$	Plasma	A	3379.45
		B	2460.53			B	1821.26
		AB	766.14			AB	598.29
		O	6154.89			O	5220.76
	Red blood cells	A	4914.84	Red blood cells	A	4428.12	
		B	3090.94		B	2701.67	
		AB	204.17		AB	98.43	
		O	9402.44		O	8721.99	
	Platelets	A	2842.62	Platelets	A	2633.79	
		B	1212.25		B	1029.25	
		AB	596.31		AB	549.32	
		O	4543.67		O	4269.8	

problem (RP) solution can be obtained. Taking the dataset 1200\_1 as the example, in its RP solution, Deyang is selected as the blood bank. The planning and daily operational cost (first four lines in formula (4)) is  $1.15 \times 10^8$ (RMB) and the disaster rescue cost (last line in formula (4)) is  $5.32 \times 10^6$  (RMB), totaling  $1.21 \times 10^8$ (RMB).

Specifically, Table 19 shows the maximum inventory and the emergency stock of the selected blood bank, i.e., Deyang, in this RP solution. Table 20 gives the distribution flow from Deyang to each hospital. The emergency stock in each hospital is shown in Table 21.

Notice there is no prepositioning emergency stock in WCH as shown in Table 21, since it is closer to the selected blood bank location Deyang than any potential epicenter. The blood bank can supply emergency blood timely; thus, no prepositioning inventory in WCH is required.

Similarly, other datasets can also be optimally calculated and Deyang is always selected as final blood bank, while the inventory results would be different. Specifically, the emergency stock in both blood bank and hospitals of each dataset is shown in Figure 6. It is obvious that more emergency stock will be incurred in both the blood bank and the hospitals

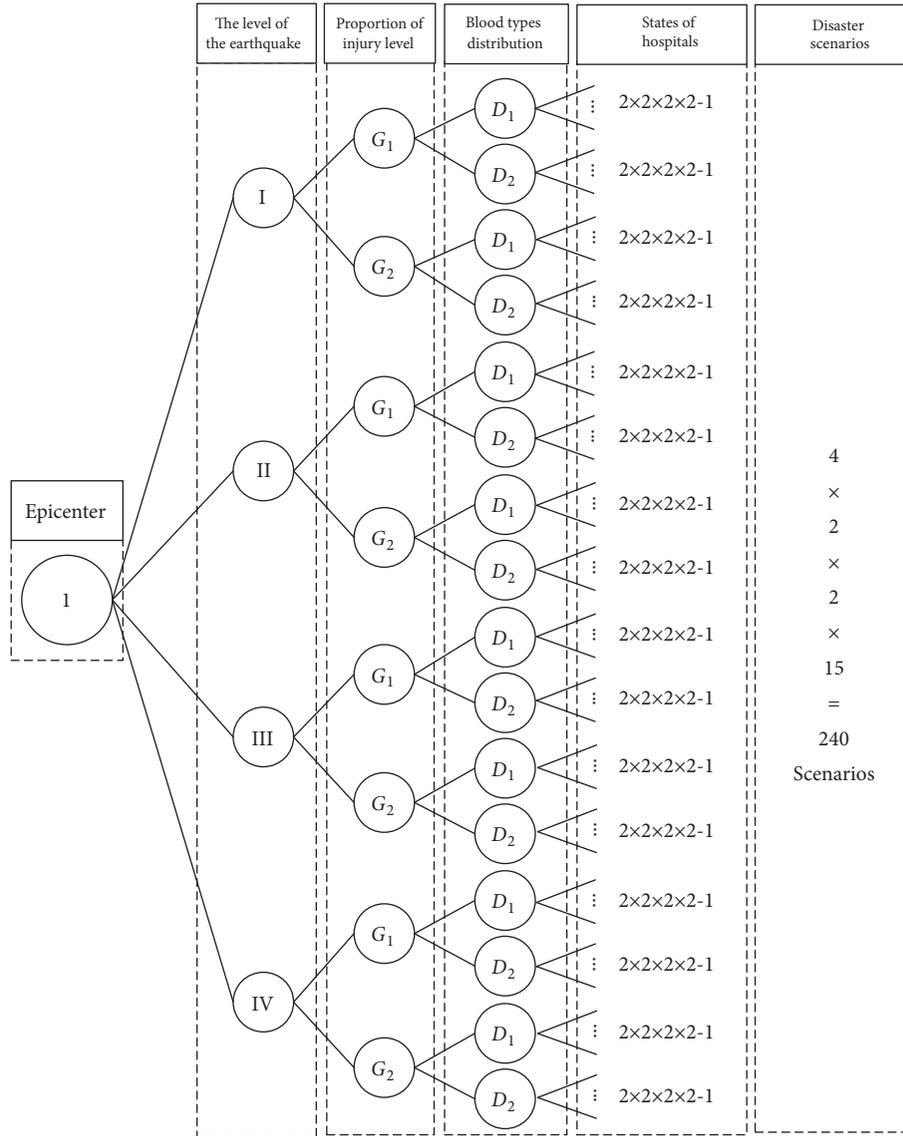


FIGURE 5: The scenario tree for one epicenter in the dataset 1200\_1.

TABLE 20: The distribution flow ( $w_{ihab}$ )(U) from the central blood bank to hospital  $h$  in each daily period.

Blood product $a$	Blood type $b$	Hospital $h$			
		WCH	BPH	WPH	MPH
Plasma	A	461.65	36.14	13.81	201.16
	B	448.15	31.22	12.11	147.79
	AB	108.83	8.49	3.24	47.29
	O	577.87	47.69	18.24	290.33
Red blood cells	A	327.47	23.24	8.87	127.52
	B	242.03	20.02	7.67	119.55
	AB	77.12	5.43	2.05	21.14
	O	471.43	30.38	11.59	167.04
Platelets	A	140.51	9.84	3.75	54.72
	B	121.87	9.57	4.09	47.46
	AB	33.05	1.11	0.00	12.84
	O	184.25	12.90	4.94	71.78

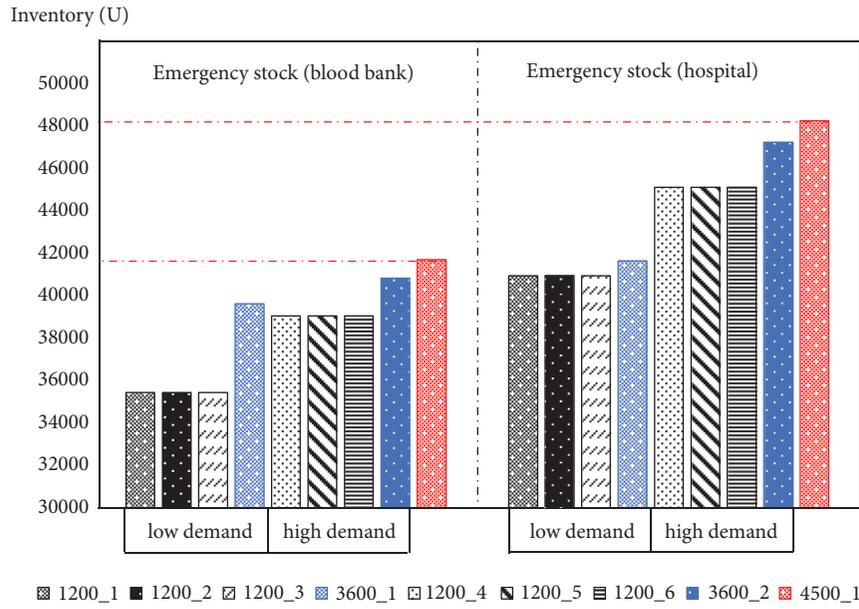


FIGURE 6: Total emergency stock in blood bank and all hospitals of different datasets.

TABLE 21: The emergency stock ( $s_{hab}$ )(U) in hospital  $h$ .

Blood product $a$	Blood type $b$	Hospital $h$			
		WCH	BPH	WPH	MPH
Plasma	A	0	2451.57	1604.92	15.13
	B	0	1932.33	1265.26	11.92
	AB	0	596.29	391.02	3.68
	O	0	2674.08	1750.46	16.5
Red blood cells	A	0	3547.9	2333.35	21.9
	B	0	2799.65	1842.68	17.28
	AB	0	870.96	576.4	5.37
	O	0	870.96	2543.47	23.87
Platelets	A	0	1886.98	1236.37	11.64
	B	0	1487.63	975.02	9.18
	AB	0	459.76	302.01	2.83
	O	0	2058.09	1348.35	12.70

under the higher level of the injury, e.g.,  $G = (G_4, G_5)$ . Besides, we can find that total emergency stock increases with the size of  $S_1$ . For example, the dataset with 4500 scenarios would result in the highest stock.

Moreover, for the datasets with the same  $G$ , different blood type distributions ( $D$ ) do not influence the extent of disasters, and, thus, the quantity of total inventory prepositioned in the BLN keeps unchanged. However, the inventory quantities with each blood type certainly would vary with different  $D$ . Due to space limitations, we do not list the results here.

In summary, the decision on blood bank location is consistent, but the different random variables would significantly influence the inventory decisions. Specifically, with the increase of the dimensionality of scenario set, more disaster scenarios, although their possibilities are getting smaller, would be taken into account. Thus, to handle these extreme

situations, larger inventory would be prepositioned which makes the solution more conservative.

**4.2.2. Sensitivity Analysis.** This section further implements sensitivity analysis by changing (1) the disaster probability, i.e.,  $P(C_n)$ , which only influences the probability of each BN-based scenario rather than the dimensionality, and (2) the unit inventory holding (UIH) cost, a cost-related parameter which is irrelevant to the BN.

*(1) Sensitivity Analysis of the Disaster Probability.* We focus on  $P(C_n)$  because it is usually not easy to estimate exactly. Considering earthquake is a small-probability event, we set the probabilities of the disaster happening in five potential epicenters to be 0.2, 0.25, 0.5, 1, 2, 3, 4, 5 times of the default value respectively.

TABLE 22: Daily transportation cost and rescue cost under varying disaster probabilities (million).

The disaster probability change ratio	0.2	0.25	0.5	1	2	3	4	5
Daily transportation cost								
All datasets	34.80	34.62	33.80	32.20	29.33	26.88	24.64	22.69
Rescue cost								
1200_1~1200_3	1.14	1.44	2.74	5.32	9.96	13.87	17.38	20.51
1200_4~1200_6	1.21	1.52	2.90	5.63	10.56	14.69	18.41	21.73
3600_1	1.18	1.46	2.77	5.39	10.10	14.07	17.63	20.80
3600_2	1.20	1.48	2.83	5.50	10.30	14.34	17.97	21.20
4500_1	1.20	1.49	2.82	5.43	10.18	14.18	17.76	20.96

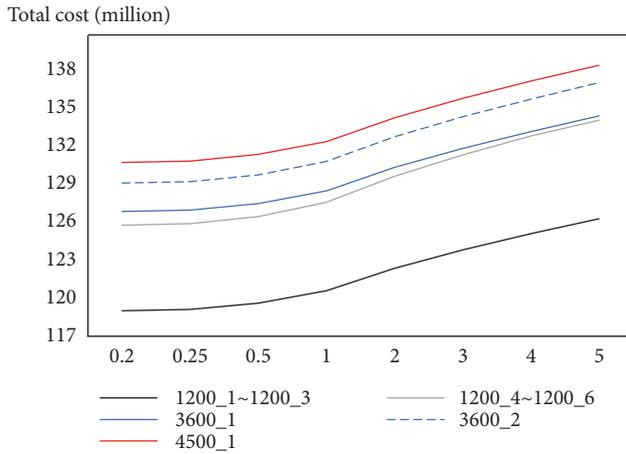


FIGURE 7: Total cost under varying disaster probabilities.

After calculation, it is found that, for each dataset, the decisions of blood bank location and emergency stock remain unchanged under the different disaster probabilities. And these two decisions are in the first stage, i.e., the planning stage, of two-stage SP which are what the decision-maker mainly care about. The reason is that  $P(C_n)$  only represents the occurrence probability but would not influence the extent of the disaster. Thus, it means, with the utilization of BN, we do not need a very exact disaster probability for the purpose of the BLN planning.

Since the decisions on location as well as emergency stock under different  $P(C_n)$  are the same, the inventory cost in each facility remains unchanged. However, the daily transportation cost, the rescue cost, and then the total cost would change along with the disaster probability as shown in Table 22 and Figure 7.

At first, for each disaster probability, the daily transportation cost of all datasets is identical. It is because their daily demand and the probability of normal scenario  $p^{s_0}$  are the same. Moreover, the growth of the disaster probability would increase the probability of total disaster scenarios  $\sum_{s \in S_1} p^s$  and thus decrease  $p^{s_0} (= 1 - \sum_{s \in S_1} p^s)$ . Based on formula (4), the daily transportation cost decreases and the rescue cost become larger when the disaster probability increases. Then, both of them work together and result in a higher total cost (see Figure 7).

(2) *Sensitivity Analysis of the Unit Inventory Holding (UIH) Cost.* To explore how the BN-unrelated parameters, i.e., the UIH cost, influence final decisions, sensitivity analysis is further implemented by varying  $r_i$  ( $i \in I$ ) and  $u_h$  ( $h \in H$ ) simultaneously. Their values can be 0.1, 0.2, ..., 1.8 times of their default values.

Firstly, Figure 8 shows the final location decisions in which the blood bank locations of all datasets change from Chengdu to Deyang with the increase of the UIH cost. The reason is that when the inventory cost is very low, the daily transportation cost would dominate the location decision. Thus, Chengdu is selected due to its geographical advantage. On the contrary, with the increase of this parameter, the decision-maker would choose a new blood bank, i.e., Deyang; thus, the inventory could be allocated in facilities with relatively lower cost. In addition, recall that, compared with 1200\_1~1200\_6, datasets with 3600 and 4500 scenarios would incur more inventory. Hence, they are more sensitive to the change of the UIH cost and thus would choose Deyang when this parameter is smaller than that of 1200 scenarios.

Moreover, it is found that the emergency stock prepositioned in the whole BLN, which is decided by the BN, remains unchanged under the different UIH costs. But the inventory in each facility may vary due the different location results. It also means the planning capacity of the blood bank is influenced by the UIH cost.

In summary, different from to the BN-related parameters which only affect the inventory decisions, the UIH cost would influence the location decision. Thus, it is necessary to measure the UIH cost exactly and analyze the location decision's sensitivity with respect to such cost-related parameters.

## 5. The Expected Value Model

To prove the effectiveness of the proposed two-stage SP model, in this section, we give the corresponding so-called Expected Value (EV) Model, in which the random variable is replaced by the expected value, and compare the results of two models.

*5.1. The Formulation of the Expected Value Model.* At first, we reformulate the two-stage SP in Section 3.2 to give the EV version as follows. The objective function (36) is the sum of cost in the planning stage and the daily periods of the

Unit inventory cost change ratio	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8
3600_1 and 3600_2	Chengdu <span style="margin-left: 150px;">Deyang</span>																	
4500_1																		
1200_1~1200_6																		

FIGURE 8: Blood bank location of each dataset under varying UIH costs.

operational stage, as well as the transportation cost of the expected value of blood demand when the disaster happens.

$$\begin{aligned}
 EV = \min & \sum_{i \in I} c_i \cdot x_i + \Phi \cdot \left\{ T \cdot \sum_{i \in I} \left[ r_i \right. \right. \\
 & \cdot \left. \sum_{k \in K} \sum_{a \in A} \sum_{b \in BL} \bar{g}_{kiab} \right] + T \cdot \sum_{h \in H} \left[ u_h \right. \\
 & \cdot \left. \sum_{a \in A} \sum_{b \in BL} \left( \sum_{i \in I} \frac{x_i \cdot (T + t_{ih}) \cdot d_{hab}}{2} + s_{hab} \right) \right] + e \cdot v \\
 & \cdot \sum_{i \in I} \sum_{k \in K} \left( t_{ki} \cdot \sum_{a \in A} \sum_{b \in BL} \bar{g}_{kiab} \right) \left. \right\} + \Phi \cdot e \cdot v \\
 & \cdot \sum_{i \in I} \sum_{h \in H} \left( t_{ih} \cdot \sum_{a \in A} \sum_{b \in BL} \left( w_{ihab} + z_{ihab} + \sum_{b' \in BL, b' \neq b} (w'_{ihab'b} + z'_{ihab'b}) \right) \right)
 \end{aligned} \quad (36)$$

In the EV model, in addition to the original constraints (5)~(7), (9), (12), and (13), new constraints (37)~(41) are added as follows:

$$\sum_{k \in K} \bar{g}_{kiab} \leq C_{iab}, \quad i \in I; a \in A; b \in BL \quad (37)$$

$$0 \leq \sum_{i \in I} \bar{g}_{kiab} \leq f_{kab}, \quad k \in K; a \in A; b \in BL \quad (38)$$

$$\begin{aligned}
 & \sum_{k \in K} \bar{g}_{kiab} \\
 & \geq ss_{iab} + \sum_{h \in H} \left( w_{ihab} + \sum_{b' \in BL, b' \neq b} q_{b'b} \cdot w'_{ihab'b} \right), \quad (39) \\
 & \quad \quad \quad i \in I; a \in A; b \in BL
 \end{aligned}$$

$$\begin{aligned}
 & ss_{iab} \geq \sum_{h \in H} \left( z_{ihab} + \sum_{b' \in BL, b' \neq b} q_{b'b} \cdot z'_{ihab'b} \right), \quad (40) \\
 & \quad \quad \quad i \in I; a \in A; b \in BL
 \end{aligned}$$

$$\begin{aligned}
 & \sum_{i \in I} \left( z_{ihab} + \sum_{b' \in BL, b' \neq b} q_{bb'} \cdot z'_{ihab'b} \right) \\
 & \geq \sum_{i \in I} (x_i \cdot \Delta T \cdot \bar{d}_{hab}), \quad h \in H; a \in A; b \in BL \quad (41)
 \end{aligned}$$

Let  $\bar{g}_{kiab}$  be the expected distribution flow for blood product  $a$  with type  $b$  from donor point  $k$  to blood bank candidate  $i$ . Constraints (37) enforce that the total expected quantity of blood provided by donor points within each  $T$ , which involves emergency stock, is not more than the maximum inventory in corresponding blood bank. Constraints

TABLE 23: The one-to-one match between potential epicenters and hospitals.

Potential epicenters	Hospitals
Wenchuan	WPH
Beichuan	BPH
Maoxian	WPH
Lushan	WCH
Pingwu	BPH

(38) express the capacity of blood supply by each donor point. Constraints (39) specify that the quantity of daily and emergency blood supplied to hospitals should be less than the quantity provided by donor points within each  $T$ . Constraints (40) ensure that enough blood is prepositioned in blood bank for an earthquake. Decision variables  $z_{ihab}$  and  $z'_{ihab'b}$  are the quantities of emergency blood supply for corresponding blood type  $b$  and type  $b'$  which can be substituted by type  $b$ , respectively. Constraints (41) show that the additional quantity of blood supply should satisfy the expected value of demand caused by the earthquakes, in which  $\bar{d}_{hab}$  is the expected value of demand caused by the earthquake for blood product  $a$  with type  $b$  in hospital  $h$ . In summary, the EV model is a mixed-integer linear programming which is tractable. The key is to give the expected emergency demand which would be shown at the next section.

**5.2. The Simulation Parameters of the EV Model.** Taking the dataset 1200\_1 in Section 4.1 as the example, we next show how the parameters, including the expected demand, in the EV model can be given.

At first, according to the distance from epicenters to hospitals in Table 12, Table 23 shows the one-to-one match between potential epicenters and hospitals. In the EV model, we assume that four hospitals are all available in the disaster.

According to the probabilities of four magnitudes, formulas (42) show the expected magnitude, and (43) give the mean intensity according to formulas (27), respectively.

$$\begin{aligned}
 \bar{F}_m &= 0.264 \times 8 + 0.189 \times 7 + 0.151 \times 6.5 \\
 &+ 0.396 \times 6 = 6.7925, \quad F_m \in F \quad (42)
 \end{aligned}$$

$$\overline{\text{Intensity}} = \frac{\bar{F}_m - 1.5}{0.58} = 9.125, \quad F_m \in F \quad (43)$$

Then the total number of the injured can be calculated by formulas (28) and (29) by using the following two variables.

The proportions of two-level demand and four blood types are both turned into corresponding expected value.

TABLE 24: The average blood demand (U) of each epicenter.

Blood product	Blood type	Epicenters				
		Wenchuan	Beichuan	Maoxian	Lushan	Pingwu
Plasma	A	282.28	645.10	291.22	775.02	317.68
	B	217.14	496.23	224.01	596.17	244.37
	AB	65.14	148.87	67.20	178.85	73.31
	O	304.00	694.72	313.62	834.63	342.12
Red blood cells	A	1180.22	2697.16	1217.58	3240.33	1328.23
	B	907.86	2074.74	936.60	2492.56	1021.71
	AB	272.36	622.42	280.98	747.77	306.51
	O	1271.01	2904.64	1311.24	3489.59	1430.40
Platelets	A	199.88	456.79	206.21	548.78	224.95
	B	153.75	351.37	158.62	422.14	173.04
	AB	46.13	105.41	47.59	126.64	51.91
	O	215.26	491.92	222.07	590.99	242.25

According to the proportion of two-level demand, blood type distribution and the corresponding probability stated in Section 4.1, let 0.35:0.65 be the mean proportion of two-level injury and  $p_A:p_B:p_{AB}:p_O=0.325:0.25:0.075:0.35$  be the distribution proportion of four types (A: B: AB: O). Same with the assumption in the two-stage SP, the blood bank is only required to meet the emergency blood demand in 2 hours. According to the expected proportions above, Tables 16 and 24 show the expected blood demand for each epicenter.

Multiplying the expected blood demand (see Table 24) by the corresponding probability of single-point earthquake (see Table 14), Table 25 is obtained in which the daily blood demand and expected emergency blood demand for each hospital are shown.

The parameters of the EV models under other datasets can be obtained in the same way.

5.3. *The EV Solution and Its Analysis.* Based on above parameters, IBM ILOG CPLEX 12.6.3 is used to optimally solve the corresponding EV model. In this EV solution, the central blood bank is located in Chengdu, and the total cost is  $4.86 \times 10^7$  (RMB). The maximum inventory and emergency stock of Expected Value Model are shown in Table 26.

By comparing the results of the RP and EV solution, the EV solution performs better in terms of the total cost. However, the EV model, which is under the deterministic setting, only takes the expected demand into account. It is necessary to validate its performance under the disaster scenarios [40].

Thus, for each dataset as shown in Table 18, we bring the EV solution into each scenario to check whether blood shortage occurs. The performance of the EV solution can be evaluated by the total probability of blood shortage occurrence if disasters really happen. We name it as failure probability (FP).

At first, the FP results under different disaster probabilities are displayed in Table 27. It is found that all the EV solutions would lead to poor FP results. When we consider higher disaster probability, the more emergency

stock would be prepositioned in the EV solution, and then the corresponding FP would be improved but also larger than 65%.

Furthermore, as to the UIH cost, its change would not influence the EV model's decisions on inventory because they are actually decided by the expected disaster demand. Hence, the FP results would not vary with the UIH cost and their values are same with the case of the disaster probability change ratio 1 as shown in Table 27.

In summary, compared with the EV solution, the RP solution can ensure the supply while reducing cost as possible by taking all possible disaster scenarios into account. Hence, the SP model which can generate the RP solution has the significant value for the blood management in the setting of disasters.

## 6. Conclusions

Human blood is a crucial and precious relief good. In this paper, we proposed a BN-based two-stage SP model to study the BLN planning problem, taking the suddenly-occurring disasters into account. The planning decisions in the BLN and the long-term operational decisions under the setting of daily and disaster scenarios are optimized globally by the two-stage SP model. The stochastic scenarios used in the SP model is generated by the proposed BN model in which multiple disaster-related factors that contribute the emergency blood demand are considered in a systematic way. Based on earthquakes in Sichuan Province, China, the simulations and related sensitivity analysis are implemented. Several observations and corresponding implications can be obtained from the results.

First, the BNs are able to describe the disaster scenarios and thus, work with the SP model to generate the effective solutions. Results indicate all these solutions significantly outperform the corresponding EV solutions in terms of the effects of rescue. Moreover, the scenario set with different dimensionalities can give the consistent location results. Hence, it is important and necessary to put efforts on the modelling of the BNs.

TABLE 25: The daily demands and expected emergency demands caused by the earthquake in hospital  $h$ .

	Blood product	Blood type	Hospitals			
			WCH	BPH	WPH	MPH
Daily demands $d_{hab}$ (U/h)	Plasma	A	3.050	0.213	0.081	1.191
		B	2.644	0.184	0.070	1.032
		AB	0.718	0.050	0.019	0.280
		O	3.007	0.281	0.107	1.562
	Red blood cells	A	1.932	0.137	0.052	0.755
		B	1.675	0.118	0.045	0.655
		AB	0.455	0.032	0.012	0.178
		O	2.532	0.179	0.068	0.989
	Platelets	A	0.829	0.058	0.022	0.324
		B	0.719	0.050	0.019	0.281
		AB	0.195	0.013	0.005	0.076
		O	1.087	0.076	0.029	0.425
Expected emergency demands $\Delta T \cdot \bar{d}_{hab}$ (U)	Plasma	A	10.850	17.696	13.437	0
		B	8.346	13.613	10.336	0
		AB	2.504	4.084	3.101	0
		O	11.685	19.058	14.471	0
	Red blood cells	A	45.365	73.989	56.180	0
		B	34.896	56.914	43.216	0
		AB	10.469	17.074	12.965	0
		O	48.854	79.680	60.502	0
	Platelets	A	7.683	12.531	9.515	0
		B	5.910	9.639	7.319	0
		AB	1.773	2.892	2.196	0
		O	8.274	13.494	10.246	0

TABLE 26: The decisions of maximum inventory and emergency stock of the blood bank in the EV solution.

The prespecified central blood bank: Chengdu							
Maximum inventory $C_{iab}$ (U)	Plasma	A	807.29	Emergency stock $ss_{iab}$ (U)	A	42	
		B	695.67		Plasma	B	33
		AB	189.74		AB	AB	10
		O	882.33		O	O	46
	Red blood cells	A	660.87	Emergency stock $ss_{iab}$ (U)	A	176	
		B	555.74		Red blood cells	B	136
		AB	154.74		AB	AB	41
		O	824.9		O	O	190
	Platelets	A	237.79	Emergency stock $ss_{iab}$ (U)	A	30	
		B	203.26		Platelets	B	23
		AB	55.63		AB	AB	7
		O	304.89		O	O	33

Furthermore, the larger scenario set can cover more detailed disaster scenarios, but it does not mean the bigger dimensionality is better. Because it not only requires more modelling efforts but also may cause conservative solutions, such as more prepositioning inventory. Considering the preciousness and perishability of human blood, over-stock also should be avoided.

At last, the parameters, including the random factors involving the BNs and the deterministic ones, would influence the final decisions in different ways. For example, the total quantity of the emergency stock is only influenced by the some of the BN-related factors, while the decision on blood bank location is decided by the cost-related parameters, such as the unit inventory holding cost. Hence, by the utilization

TABLE 27: The FP of the EV solutions of each dataset under varying disaster probabilities.

The disaster probability change ratio	1200						3600		4500_1
	1200_1	1200_2	1200_3	1200_4	1200_5	1200_6	3600_1	3600_2	
0.2	83.92%	84.30%	84.30%	83.92%	83.92%	83.92%	83.80%	83.80%	83.95%
0.25	84.08%	84.08%	84.08%	84.08%	84.08%	84.08%	84.06%	84.06%	84.22%
0.5	83.93%	83.93%	83.93%	83.93%	83.93%	83.93%	83.95%	83.95%	83.97%
1	83.98%	83.98%	83.98%	64.80%	64.80%	77.43%	78.61%	77.39%	77.45%
2	64.80%	64.80%	64.80%	64.80%	64.80%	64.80%	64.83%	64.83%	64.80%
3	64.96%	64.96%	64.96%	64.96%	64.96%	64.96%	64.96%	64.96%	64.97%
4	65.30%	65.30%	65.30%	65.30%	65.30%	65.30%	65.31%	65.31%	65.31%
5	65.30%	65.30%	65.30%	65.30%	65.30%	65.30%	65.29%	65.29%	65.30%

of the BNs, we can reveal the specific relationship between the random factors and the final decisions. For example, if the decision-maker cares about the planning of the total inventory, we should focus on the estimation of the severity of the injured, instead of the distribution of blood types, the disaster probability, and the unit inventory cost, etc. Thus, we can identify the specific efforts that should be made in terms of different planning purposes. This is the main benefit of the BNs.

There are some future directions as follows. At first, the BN model we proposed here only fits earthquakes. The applications of the BNs in the setting of other disasters still leave lots of opportunities. Moreover, we only consider the uncertainty of demand in this paper. But, when the earthquake occurs, blood shortage may be satisfied by urgent blood donations partially which are also random. And the more serious the disaster is, the more relief goods, including blood, the government and the people donate. Alfonso et al. [55], Jabbarzadeh et al. [22], and Chaiwuttisak et al. [56] notice the uncertainty of the donations but fail to utilize its underlying relationship with the random demand. The BN is still a possible way to model such relationships which can be addressed in future research. At last, another interesting research direction is to explore the application of distributionally robust approach which only requires moments information [57–59]. But the underlying correlations between different stochastic variables, which are involved in the second moment, usually are hard to obtain [32]. A natural idea is to address this issue by the BN.

## Data Availability

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

## Conflicts of Interest

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

## Acknowledgments

This work is supported by National Natural Science Foundation of P. R. China (Grant Nos. 71572033 and 71832001).

## References

- [1] M. Rezaei-Malek, R. Tavakkoli-Moghaddam, B. Zahiri, and A. Bozorgi-Amiri, "An interactive approach for designing a robust disaster relief logistics network with perishable commodities," *Computers & Industrial Engineering*, vol. 94, no. C, pp. 201–215, 2016.
- [2] M. Eskandari-Khanghahi, R. Tavakkoli-Moghaddam, A. A. Taleizadeh, and S. H. Amin, "Designing and optimizing a sustainable supply chain network for a blood platelet bank under uncertainty," *Engineering Applications of Artificial Intelligence*, vol. 71, pp. 236–250, 2018.
- [3] H. Mo, World Disasters Report, 2015, <http://ifrc-media.org/interactive/world-disasters-report-2015/>.
- [4] J. B. Sheu, "Dynamic relief-demand management for emergency logistics operations under large-scale disasters," *Transportation Research Part E: Logistics and Transportation Review*, vol. 46, no. 1, pp. 1–17, 2010.
- [5] B. Vahdani, M. Soltani, M. Yazdani, and S. Meysam Mousavi, "A three level joint location-inventory problem with correlated demand, shortages and periodic review system: robust meta-heuristics," *Computers & Industrial Engineering*, vol. 109, no. 7, pp. 113–129, 2017.
- [6] J. Beliën and H. Forcé, "Supply chain management of blood products: a literature review," *European Journal of Operational Research*, vol. 217, no. 1, pp. 1–16, 2012.
- [7] Q. Duan and T. W. Liao, "Optimization of blood supply chain with shortened shelf lives and ABO compatibility," *International Journal of Production Economics*, vol. 153, no. 7, pp. 113–129, 2014.
- [8] D. Alem, A. Clark, and A. Moreno, "Stochastic network models for logistics planning in disaster relief," *European Journal of Operational Research*, vol. 255, no. 1, pp. 187–206, 2016.
- [9] G. Galindo and R. Batta, "Review of recent developments in OR/MS research in disaster operations management," *European Journal of Operational Research*, vol. 230, no. 2, pp. 201–211, 2013.
- [10] J. A. Paul and L. MacDonald, "Location and capacity allocations decisions to mitigate the impacts of unexpected disasters," *European Journal of Operational Research*, vol. 251, no. 1, pp. 252–263, 2016.

- [11] G. Barbarosoğlu and Y. Arda, "A two-stage stochastic programming framework for transportation planning in disaster response," *Journal of the Operational Research Society*, vol. 55, no. 1, pp. 43–53, 2004.
- [12] H. O. Mete and Z. B. Zabinsky, "Stochastic optimization of medical supply location and distribution in disaster management," *International Journal of Production Economics*, vol. 126, no. 1, pp. 76–84, 2010.
- [13] M. Haghi, S. M. T. Fatemi Ghomi, and F. Jolai, "Developing a robust multi-objective model for pre/post disaster times under uncertainty in demand and resource," *Journal of Cleaner Production*, vol. 154, pp. 188–202, 2017.
- [14] S. Tofghi, S. A. Torabi, and S. A. Mansouri, "Humanitarian logistics network design under mixed uncertainty," *European Journal of Operational Research*, vol. 250, no. 1, pp. 239–250, 2016.
- [15] A. Mohamadi and S. Yaghoubi, "A bi-objective stochastic model for emergency medical services network design with backup services for disasters under disruptions: an earthquake case study," *International Journal of Disaster Risk Reduction*, vol. 23, pp. 204–217, 2017.
- [16] S. An, N. Cui, Y. Bai, W. Xie, M. Chen, and Y. Ouyang, "Reliable emergency service facility location under facility disruption, en-route congestion and in-facility queuing," *Transportation Research Part E: Logistics and Transportation Review*, vol. 82, pp. 199–216, 2015.
- [17] A. Verma and G. M. Gaukler, "Pre-positioning disaster response facilities at safe locations: an evaluation of deterministic and stochastic modeling approaches," *Computers & Operations Research*, vol. 62, pp. 197–209, 2015.
- [18] C. G. Kochan, S. S. Kulkarni, and D. R. Nowicki, "Efficient inventorying and distribution of blood product during disasters," in *Advances in Managing Humanitarian Operations*, International Series in Operations Research & Management Science, pp. 185–204, Springer International Publishing, Cham, Switzerland, 2016.
- [19] S. Gunpinar and G. Centeno, "Stochastic integer programming models for reducing wastages and shortages of blood products at hospitals," *Computers & Operations Research*, vol. 54, no. 54, pp. 129–141, 2015.
- [20] B. Fahimnia, A. Jabbarzadeh, A. Ghavamifar, and M. Bell, "Supply chain design for efficient and effective blood supply in disasters," *International Journal of Production Economics*, vol. 183, pp. 700–709, 2017.
- [21] M. R. Ghatreh Samani, S. A. Torabi, and S.-M. Hosseini-Motlagh, "Integrated blood supply chain planning for disaster relief," *International Journal of Disaster Risk Reduction*, vol. 27, pp. 168–188, 2018.
- [22] A. Jabbarzadeh, B. Fahimnia, and S. Seuring, "Dynamic supply chain network design for the supply of blood in disasters: a robust model with real world application," *Transportation Research Part E: Logistics and Transportation Review*, vol. 70, no. 1, pp. 225–244, 2014.
- [23] S. Khalilpourazari and A. Arshadi Khamseh, "Bi-objective emergency blood supply chain network design in earthquake considering earthquake magnitude: a comprehensive study with real world application," *Annals of Operations Research*, no. 1, pp. 1–39, 2017.
- [24] F. Salehi, M. Mahootchi, and S. M. Hussein, "Developing a robust stochastic model for designing a blood supply chain network in a crisis: a possible earthquake in Tehran," *Annals of Operations Research*, pp. 1–25, 2017.
- [25] D. Rahmani, "Designing a robust and dynamic network for the emergency blood supply chain with the risk of disruptions," *Annals of Operations Research*, no. 1, pp. 1–29, 2018.
- [26] D. Shishebori and A. Yousefi Babadi, "Robust and reliable medical services network design under uncertain environment and system disruptions," *Transportation Research Part E: Logistics and Transportation Review*, vol. 77, pp. 268–288, 2015.
- [27] B. Zahiri, S. A. Torabi, M. Mousazadeh, and S. A. Mansouri, "Blood collection management: methodology and application," *Applied Mathematical Modelling: Simulation and Computation for Engineering and Environmental Systems*, vol. 39, no. 23–24, pp. 7680–7696, 2015.
- [28] M. Fereiduni and K. Shahanaghi, "A robust optimization model for blood supply chain in emergency situations," *International Journal of Industrial Engineering Computations*, vol. 7, no. 4, pp. 535–554, 2016.
- [29] R. Ramezani and Z. Behboodi, "Blood supply chain network design under uncertainties in supply and demand considering social aspects," *Transportation Research Part E: Logistics and Transportation Review*, vol. 104, pp. 69–82, 2017.
- [30] M. R. G. Samani and S.-M. Hosseini-Motlagh, "An enhanced procedure for managing blood supply chain under disruptions and uncertainties," *Annals of Operations Research*, no. 1, pp. 1–50, 2018.
- [31] A. Kamyabniya, M. M. Lotfi, M. Naderpour, and Y. Yih, "Robust platelet logistics planning in disaster relief operations under uncertainty: a coordinated approach," *Information Systems Frontiers*, vol. 20, no. 4, pp. 759–782, 2018.
- [32] S. Agrawal, Y. Ding, A. Saberi, and Y. Ye, "Price of correlations in stochastic optimization," *Operations Research*, vol. 60, no. 1, pp. 150–162, 2012.
- [33] V. Gabrel, C. Murat, and A. Thiele, "Recent advances in robust optimization: an overview," *European Journal of Operational Research*, vol. 235, no. 3, pp. 471–483, 2014.
- [34] N. Bassamzadeh and R. Ghanem, "Multiscale stochastic prediction of electricity demand in smart grids using Bayesian networks," *Applied Energy*, vol. 193, pp. 369–380, 2017.
- [35] F. Archetti, A. Gaivoronski, and F. Stella, "Stochastic optimization on Bayesian nets," *European Journal of Operational Research*, vol. 101, no. 2, pp. 360–373, 1997.
- [36] J. Wu, S. Xu, R. Zhou, and Y. Qin, "Scenario analysis of mine water inrush hazard using Bayesian networks," *Safety Science*, vol. 89, pp. 231–239, 2016.
- [37] M. Afenyo, F. Khan, B. Veitch, and M. Yang, "Arctic shipping accident scenario analysis using Bayesian Network approach," *Ocean Engineering*, vol. 133, pp. 224–230, 2017.
- [38] Y. Huang, G. Ma, and J. Li, "Grid-based risk mapping for gas explosion accidents by using Bayesian network method," *Journal of Loss Prevention in the Process Industries*, vol. 48, pp. 223–232, 2017.
- [39] A. Sebastian, E. J. C. Dupuits, and O. Morales-Nápoles, "Applying a Bayesian network based on Gaussian copulas to model the hydraulic boundary conditions for hurricane flood risk analysis in a coastal watershed," *Coastal Engineering Journal*, vol. 125, pp. 42–50, 2017.
- [40] J. R. Birge and F. V. Louveaux, *Introduction to Stochastic Programming*, Springer, London, UK, 2011.
- [41] Y. X. Hu, *Earthquake Engineering*, Seismological Press, Beijing, China, 1988.
- [42] U. Chandra, "Attenuation of intensities in the united states," *Bulletin of the Seismological Society of America*, vol. 69, no. 6, pp. 2003–2024, 1979.

- [43] S. P. Lee, "Practical scale," *Chinese Journal of Geophysics*, vol. 7, no. 2, pp. 98–102, 1958.
- [44] X. Ma, *Study on the Location and Allocation of Social Emergency Materials Reservation*, Graduate University of Chinese Academy of Sciences, Beijing, China, 2011.
- [45] M. Wyss and G. Trendafiloski, "Trends in the casualty ratio of injured to fatalities in earthquakes," in *Advances in Natural and Technological Hazards Research*, R. Spence, E. So, and C. Scawthorn, Eds., vol. 29, Springer, Dordrecht, Netherlands, 2011.
- [46] D. R. Peng, "The distribution of ABO blood group of Han nationality in China," *Chinese Journal of Blood Transfusion*, vol. 4, no. 1, pp. 20–23, 1991.
- [47] Z. Y. Chen, "The distribution of ABO blood group in Chinese," *Hereditas*, vol. 4, no. 2, pp. 4–7, 1982.
- [48] World Population Network, "Ranking of Sichuan Population," 2014, <http://www.renkou.org.cn/china/sichuan/2014/2208.html>.
- [49] The Minister of Health of the People's Republic of China, *People's Republic of China Health Industry Standard (Standards for Blood Storage)*, People's Republic of China, 2012, <http://www.csbt.org.cn/plus/view.php?aid=94>.
- [50] X. L. Deng, X. H. Liang, L. Wang et al., "Survey on the situation of blood supply in 355 blood supply institutions," *Chinese Journal of Blood Transfusion*, vol. 24, no. 9, pp. 779–780, 2011.
- [51] J. Wang, L. Jiang, and Y. H. Wang, "Analysis of clinical blood and its significance in general hospital," *Psychologist*, vol. 22, no. 30, pp. 31–32, 2016.
- [52] Z.-T. Liu, D.-L. Wang, W.-J. Zhang, W. Feng, and T.-Y. Zheng, "Early judgment method of earthquake emergency response level based on Bayesian classification," *Earthquake*, vol. 31, no. 2, pp. 114–121, 2011.
- [53] J. Ling, Y. He, X. Q. Liu et al., "Analysis of characteristics about blood transfusion in 2065 cases of inpatients injured by earthquake," *Laboratory Medicine and Clinic*, vol. 7, no. 1, pp. 7–8, 2010.
- [54] C. H. Liu, C. L. Yuan, and Q. J. Huang, "Analysis of features of 320 blood transfusion earthquake victims in Deyang area," *Laboratory Medicine and Clinic*, vol. 7, no. 1, pp. 11–12, 2010.
- [55] E. Alfonso, X. Xie, and V. Augusto, "A simulation-optimization approach for capacity planning and appointment scheduling of blood donors based on mathematical programming representation of event dynamics," in *Proceedings of the 11th IEEE International Conference on Automation Science and Engineering, CASE 2015*, pp. 728–733, Sweden, August 2015.
- [56] P. Chaiwuttisak, H. Smith, Y. Wu, C. Potts, T. Sakuldamongpanich, and S. Pathomsiri, "Location of low-cost blood collection and distribution centres in Thailand," *Operations Research for Health Care*, vol. 9, pp. 7–15, 2016.
- [57] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Operations Research*, vol. 58, no. 3, pp. 595–612, 2010.
- [58] D. Bertsimas, K. Natarajan, and C.-P. Teo, "Probabilistic combinatorial optimization: moments, semidefinite programming and asymptotic bounds," *SIAM Journal on Optimization*, vol. 15, no. 1, pp. 185–209, 2004.
- [59] D. Bertsimas, X. V. Doan, K. Natarajan, and C.-P. Teo, "Models for minimax stochastic linear optimization problems with risk aversion," *Mathematics of Operations Research*, vol. 35, no. 3, pp. 580–602, 2010.

## Research Article

# Evaluating the Spatial Deprivation of Public Transportation Resources in Areas of Rapid Urbanization: Accessibility and Social Equity

Chunyang Han <sup>1</sup>, Xinquan Liu <sup>2</sup>, Xiaojing Shen <sup>2</sup>, Ling Zhang <sup>3,4,5</sup> and Nana Feng<sup>2</sup>

<sup>1</sup>School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China

<sup>2</sup>School of Logistics Management and Engineering, Nanning Normal University, Nanning 530001, Guangxi, China

<sup>3</sup>School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China

<sup>4</sup>Yunnan Integrated Transport Development and Regional Logistics Management Think Tank, Kunming University of Science and Technology, Kunming 650500, China

<sup>5</sup>Yunnan Engineering Research Center of Modern Logistics, Kunming University of Science and Technology, Kunming 650504, China

Correspondence should be addressed to Ling Zhang; [amilyzhangg@163.com](mailto:amilyzhangg@163.com)

Received 27 November 2018; Accepted 22 January 2019; Published 11 February 2019

Academic Editor: Lu Zhen

Copyright © 2019 Chunyang Han 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.

To better understand the transportation situation in rapid urbanization areas and to improve social equity, this study constructed an approach to assess the spatial differentiation of public transportation resources based on deprivation theory and an accessibility analysis. Chenggong New District in Kunming, a typical rapid urbanization area in China, was analyzed as a case study. We introduced 6 indexes to establish a public transportation spatial deprivation evaluation system and applied SPSS to screen out two main factors that reflected the spatial deprivation associated with public transportation resources and services. Then, we adopted the accessibility model and spatial cluster model to embody residents' opportunities to obtain access to public transportation and to judge whether public transportation resource allocation is appropriate. In addition, we used ArcGIS technology to better understand the spatial deprivation characteristics of public transportation. We found that the pattern of public transportation spatial deprivation in Chenggong could be summarized as "multicore and local radiation": the spatial accessibility characteristics of public transportation take the form of a circular layer along with the metro lines and decline progressively toward the peripheral areas, where public transportation resource allocation is lacking. These findings show that the public transportation situation in rapid urbanization areas is consistent with the local land-use context and the suitability of established methods for extracting spatial public transportation characteristics.

## 1. Introduction

With economic success and social progress, rapid urbanization has become a common phenomenon in developing countries in which the scale of cities has expanded rapidly due to urban expansion. As one of the fastest-growing countries in the developing world, China has undergone a phenomenal urbanization process. The average urbanization growth rate ranges from 1% to 1.5% per year, and China's urbanization rate reached 54.77% in 2014 [1]. The urbanization process has been transformed since the early stage, which mainly occurs in the city center as a medium-term accelerated stage

of attempts to expand a city [2]. Although urbanization has been associated with economic prosperity, it has also brought a series of problems, for example, local environmental changes caused by land-use transition [3–7]; increasing energy consumption [8]; and an increasing need for access to opportunities such as education, health care, and recreation. To meet the needs of urbanization areas, rural land around large and medium-sized cities has been expropriated, and the direct urbanization that has occurred in suburban and rural areas has led to problems when the residential spatial differentiation and commuting distance increased simultaneously. The formation of rapid urbanization can usually be

described as government-oriented and market-oriented due to the different forces driving urbanization. Government-oriented urbanization highlights the creation of a series of preferential policies or institutions through which the process of urban development is promoted under the influence of the government. In contrast, market-oriented urbanization is a spontaneous bottom-up process that appears to rely heavily on the market economy and fickle consumers. The rapid urbanization process involves three main aspects: (1) the rapid spread of an urban area to the periphery, which has strong social, economic, and cultural impacts on the entire area; (2) the inrush of a large number of rural residents from rural to nonrural areas; and (3) the transformation of rural areas into new towns or satellite towns through nonagricultural development. The government in China has increased its focus on development speed and demand, and the positive side of urbanization in terms of social development is continuously emerging. Therefore, a series of policies and institutions have been created to perform administrative interventions in urbanization and to promote and guide the urbanization process, making government-oriented urbanization the typical mode. However, even with high-level planning, the government cannot avoid differential development due to the different contexts of land use in rapid urbanization areas, which may lead to the nonuniform spatial distribution of the transportation infrastructure and services due to the broadening diffusion of the urban form [9]. Urbanization will continue to be a long-term general trend of urban development in developing countries.

Social equity is sometimes equated with the term “social stability,” which means that until all people are in a relatively socially stable condition, inadequacies in transportation provision could undoubtedly exclude a number of people from fully participating in daily activities, for instance, obtaining access to education, employment, and various social and leisure pursuits [10, 11], which may foster social exclusion and therefore inequity. With the widening economic gap in rapid urbanization areas, the unequal spatial diffusion of social resources makes people, especially the poor, feel deprived. In addition, those who believe they will benefit but instead suffer a disadvantage may experience a psychological change and even increased social dissatisfaction [12], thus causing a negative effect on social stability. Urban transportation resources are among the most frequently used public services, and the allocation of these resources can embody social fairness [13–15] and enable researchers to examine the spatial mismatches between the home locations of vulnerable population groups and the availability of urban services to assess equity. One of the most powerful instruments is the accessibility analysis [16, 17], in which the accessibility measures comprise two main parts: the cost of travel (e.g., money or time) and the availability of opportunities (e.g., the distance to public transportation services). Depending on the perspective of the researcher or the location of origin, transportation mode accessibility can be measured differently, but few scholars have paid attention to the linkage between spatial and temporal issues in social activities. An analysis of this relationship may help government officials identify the appropriate land-use arrangements, especially in rapid urbanization areas

where the high-level planning is imperfect. Currently, many countries enjoy automobile-dominated urban travel markets, which make it difficult to promote alternative modes of travel [18], but the detrimental effects of automobiles on both human physical health and the environment [19–21], such as traffic congestion and pollution, have exerted considerable pressure on transportation infrastructures. Therefore, access to public transportation is becoming increasingly important.

This paper aims to establish a relatively complete method to better understand transportation situations in rapid urbanization areas, thus improving the supply level of urban public transportation resources and effectively reducing the differences between different groups in the traffic environment. To build the evaluation index system of public transportation spatial deprivation, we attempted to determine the spatial deprivation characteristics by applying geographical information system (GIS) technique; then, using the minimum time as the travel cost, we constructed a public transportation accessibility evaluation model and obtained the time needed to access public transportation in each traffic analysis zone by calculating the average travel time from a public transportation station to all other public transportation stations. Finally, we introduced a spatial clustering model to evaluate the overall public transportation resource allocation.

## 2. Methodology

To explore the spatial differentiation characteristics of public transportation and evaluate whether the resource allocation is appropriate among different areas, the methodology was conducted in three steps as follows: first, we established an evaluation index system of the public transportation spatial deprivation; a factor analysis method was used in the evaluation model, and SPSS was used to analyze the factors in the evaluation index system. Then, we established a model to evaluate how public transportation resources are allocated, based on the accessibility evaluation method and spatial clustering analysis method, and to obtain the accessibility spatial pattern of public transportation.

*2.1. Establishment of Evaluation Index System.* Based on the actual accessibility of transportation resources in most rapid urbanization areas, this research constructed the evaluation index system of public transportation spatial deprivation with two aspects, considering both the transportation infrastructure and its services, shown in Table 1 as infrastructure indicators and service indicators. The former category embodies the distribution of public transportation infrastructure resources in different traffic analysis zones and in turn the convenience of residents in obtaining access to public transportation, which consists of public transportation network coverage, bus station occupancy, and the location quotient of public transportation network density. The latter category includes indexes related to transportation services, such as the availability of bus stations within a diameter of 300 and 500 meters, and the number of bus lines in traffic analysis zones, which can reflect the level of residents’ opportunities to obtain access to public transportation and

TABLE 1: Evaluation index system of spatial deprivation of public transportation.

Category	Index	Definition
infrastructure indicators	public transportation network coverage	proportion of the road length traveled by a bus in relation to the total road length in a traffic analysis zone
	bus station occupancy	proportion of the number of bus stops in each traffic analysis zone in relation to the total number of bus stops in the area
	location quotient of public transportation network density	ratio of network density in each traffic analysis zone to public transportation network density in the overall area
service indicators	number of bus lines per traffic analysis zone	the total number of bus lines in each traffic analysis zone
	location quotient of bus station availability within a 300 meter diameter	ratio of bus station availability in each traffic analysis zone to bus station availability in the overall area within a 300 meter diameter
	location quotient of bus station availability in a 500 meter diameter	ratio of bus station availability in each traffic analysis zone to bus station availability in the overall area within a 500 meter diameter

the public transportation services scope of different traffic analysis zones. The established index system and the index definitions are shown in Table 1.

**2.2. Main Factor Analysis Model.** Factor analysis is a statistical analysis method based on multiple factors. After the dimension processing of the collected primitive variables is reduced, those that are relatively fewer and play a supporting role for the primitive variables are found to account for most of the information. This paper adopted factor analysis to extract the main deprivation factors from a spatial deprivation evaluation index system, which is given as follows:

$$F_z = \beta_{z1}x_1 + \beta_{z2}x_2 + \cdots + \beta_{zp}x_p \quad (1)$$

where  $F_z$  is the score obtained by the main factor  $z$  in a traffic analysis zone;  $\beta_{zp}$  is the coefficient of  $p$ -th standardization index;  $x_p$  represents  $p$ -th standardization index of this traffic analysis zone, which is obtained automatically by SPSS processing; and  $p$  is a parameter denoting the total number of indexes. Based on the calculated main factor scores, the spatial distribution of public transportation resources and services could be determined. The average score of the main factors was 0, and a negative score indicated deprivation in the aspect of that main factor; the lower the score, the more severe the deprivation.

Based on the scores and contribution rates of the main factors obtained above, we could calculate the composite index of public transportation in each traffic analysis zone. The computation formula is given as follows:

$$C_i = \sum_{z=1}^m F_{iz} \times r_{iz} \quad (2)$$

where  $C_i$  is the composite index of public transportation in traffic analysis zone  $i$ ,  $F_{iz}$  represents the score obtained by main factor  $z$  in traffic analysis zone  $i$ ,  $r_{iz}$  is the contribution rate of main factor  $z$  in traffic analysis zone  $i$ , and  $m$  is a parameter denoting the total number of factors. The average

score of the composite index of each traffic analysis zone is 0, and a negative score indicates that it is below the average level. The lower the score, the more severe is the spatial deprivation of public transportation in that traffic analysis zone.

**2.3. Accessibility Evaluation Model.** Generally, accessibility refers to the degree of difficulty experienced by different residents in traveling through the urban spatial structure from their home to their destination by various transportation modes within the urban transportation network. As an important index that reflects the relationship between urban spatial structure and the transportation network, it can also measure the spatial distribution proportion of transportation resources in the urban spatial structure. To measure accessibility, there are three main methods: the spatial barrier model, the cumulative opportunity model and the spatial interactive model. To measure public transportation network accessibility in rapid urbanization areas, this paper considers the attraction of all public transportation stations to be equal; thus, a spatial barrier model was selected. With the minimum travel time as a spatial barrier, the time accessibility of each public transportation station can be characterized by calculating the shortest average travel time from a public transportation station to all other public transportation stations; the lower the accessibility value, the shorter the average travel time and the better the accessibility. The calculation formula is given as follows:

$$A_e = \frac{1}{n-1} \sum_{\substack{e=1 \\ e \neq f}}^n (d_{ef}) \quad (3)$$

where  $A_e$  is the time accessibility of public transportation station  $e$ ,  $d_{ef}$  is the shortest travel time from station  $e$  to station  $f$ , and  $n$  is the total number of public transportation stations in a traffic analysis zone. To characterize the convenience of travelers in obtaining access to public transportation in an integrative manner and to reflect the spatial structure of the time accessibility of public transportation stations, we constructed a comprehensive transportation accessibility

evaluation model based on formula (4), which is given as follows:

$$X_i = \frac{\sum_{e=1}^n A_{ei}}{n} \quad (4)$$

where  $X_i$  means the public transportation accessibility of traffic analysis zone  $i$ ,  $A_{ei}$  is the time accessibility of public transportation station  $e$  in traffic analysis zone  $i$ , and  $n$  is the total number of public transportation stations in traffic analysis zone  $i$ . When we calculated the average time accessibility of public transportation stations to reflect the public transportation accessibility of all traffic analysis zones, if there was no public transportation station in a traffic analysis zone, we used ArcGIS to search for the nearest stations. Therefore, the accessibility of such a traffic analysis zone was calculated by the time needed to walk to the nearest station.

**2.4. Spatial Clustering Model.** There are two main spatial clustering analysis methods: the global spatial autocorrelation model and the partial autocorrelation model. When a unit was verified, the particular attribute values in the study area were identified as related, alien, or independent of those in the neighboring regions; next, we determined whether the property value was spatial agglomeration or spatial proportionality. There are various indexes and calculation methods for the global spatial autocorrelation model, but the most commonly used Moran's I index, of which the range is  $[-1, 1]$ . Spatial agglomeration exists when the attribute value is positive, meaning each unit has a positive spatial correlation and good spatial proportionality. In contrast, a spatial differential exists when the attribute value is negative, indicating that each unit has a negative spatial correlation and the resource availability is spatially disproportionality. If the Moran's I index value is zero, it means that there is no spatial correlation, and the attribute values are independent. Usually, the standardized value  $Z(I)$  calculated in Moran's I is used to check the statistical significance of the spatial correlation based on the spatial autocorrelation analysis rules in ArcGIS. When  $Z(I)$  is greater than 1.65, it suggests a significant spatial relation and good spatial agglomeration; when it is in the range of  $(-1.65, 1.65)$ , there is a close spatial correlation; and if it is less than -1.65, it indicates that there is significant spatial discreteness. The Moran's I equations are expressed as follows:

$$I = \frac{\sum_{i=1}^q \sum_{j=1}^q W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{S^2 \sum_{i=1}^q \sum_{j=1}^q W_{ij}} \quad (5)$$

$$S^2 = \frac{1}{q} \sum_i (x_i - \bar{x})^2 \quad (6)$$

$$\bar{x} = \frac{1}{q} \sum_{i=1}^q x_i \quad (7)$$

where  $S^2$  is the variance of public transportation accessibility;  $x_i$  and  $x_j$  represent the public transportation accessibility of traffic analysis zones  $i$  and  $j$ , respectively; and  $W_{ij}$  is the

weight of the spatial distance, which is equal to 1 only if traffic analysis zones  $i$  and  $j$  are public boundaries and otherwise is equal to 0.  $q$  is the total number of traffic analysis zones in the study area.

However, it is difficult to use the global autocorrelation to identify the spatial position when there is agglomeration within a regional space mode; if no statistical significance is embodied, then partial agglomeration may exist. Therefore, in this paper, we combined the global spatial autocorrelation and partial autocorrelation models to reveal the spatial association and agglomeration of the attribute values between the spatial units and their neighborhoods as follows:

$$L(I) = Z_j \sum W_{ij} Z_i \quad (8)$$

$$Z_i = \frac{(x_i - \bar{x})}{\sqrt{(1/q) \sum_{i=1}^q (x_i - \bar{x})^2}} \quad (9)$$

$$Z_j = \frac{(x_j - \bar{x})}{\sqrt{(1/q) \sum_{j=1}^q (x_j - \bar{x})^2}} \quad (10)$$

The analysis depends mainly on whether the value  $L(I)$  is positive or negative. If the value is positive, it suggests that the public transportation accessibility of a traffic analysis zone and its surrounding zones are alike; that is, traffic analysis zones with high accessibility are surrounded by high-accessibility areas, and traffic analysis zones with low accessibility are surrounded by low-accessibility areas. In contrast, if the value is negative, traffic analysis zones with high accessibility are surrounded by low-accessibility areas, and traffic analysis zones with low accessibility are surrounded by high-accessibility areas. To consider and clarify the balance between service supply and demand diversity, we divided the traffic analysis zones into four categories: high-high superior, low-low trough, high-low outstanding, low-high backward. The first word in each category, high or low, represents the transportation resources in the traffic analysis zone, and the following words represent the resources in the surrounding areas.

### 3. Case Study

**3.1. Background of Study Area and Traffic Analysis Zone Division.** We chose Chenggong New District, which is located in the southeastern part of Kunming, the capital of Yunnan Province, as a case study because it is a microcosm representing many similar areas in China. The district is experiencing unbelievably rapid urbanization, and the infrastructure cannot support the increasing demand. Taking the lead in constructing a new modern version of Kunming, Chenggong is a typical rapid urbanization area. It was 160 km<sup>2</sup> in size with a population of 400 thousand in 2014 and consisted of several functional areas (e.g., the Dounan Flower Industry Area, Luoyang International Logistic Area, Dachong Industrial Area, Wulong Sports Area, Wujiaying Business District, Yuhayuxiu Education Area, and Dayuxiang Recreational Area) (Figure 1). With the rapid development of the economy

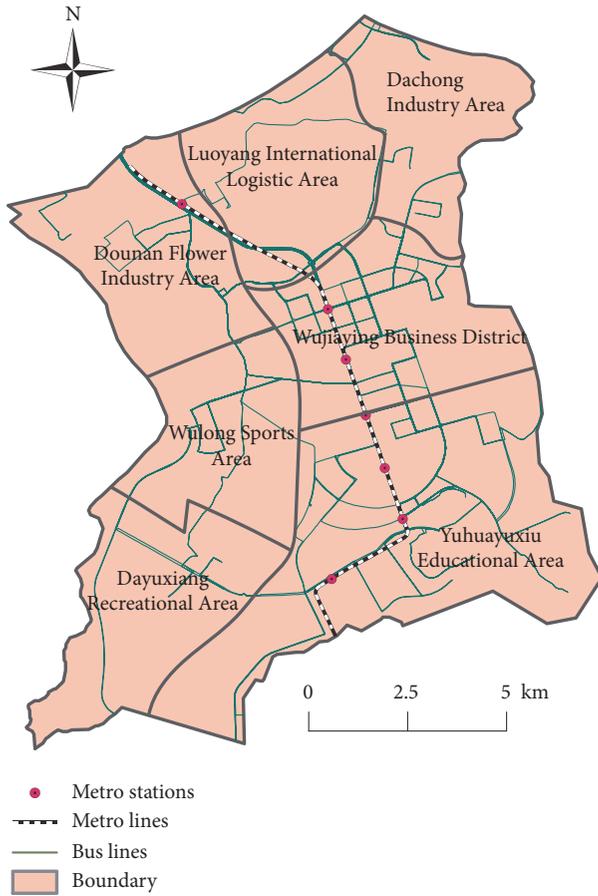


FIGURE 1: Distribution of the metro and bus line network in Chengggong New District.

and constant construction, the transportation infrastructure in Chengggong has been greatly improved. The total road length is approximately 285 kilometers, with a core backbone network of more than 100 kilometers. The coverage rate is 58.8%, and the rail transit network is nearly 14.5 km long and includes seven metro stations.

However, due to the phases and the specific nature of the rapid urbanization area under development, Chengggong will remain a diverse urban spatial structure of new central districts, relocated communities and rural areas for some time. Most of the residents of the new districts are farmers, college students, and migrant workers. Due to the limitations of their income, these people rely mainly on public transportation for their travel modes. For people who live in the more dispersed areas of Chengggong, where the public transportation options are bus or subway, different levels of public transportation services are significant differences in space that may produce public transportation spatial deprivation.

To investigate public transportation spatial deprivation and then identify spatial deprivation patterns, the study area was divided into 52 traffic analysis zones based on administrative regions and site characteristics. The traffic analysis zone division results of the Chengggong New District are presented in Figure 2.

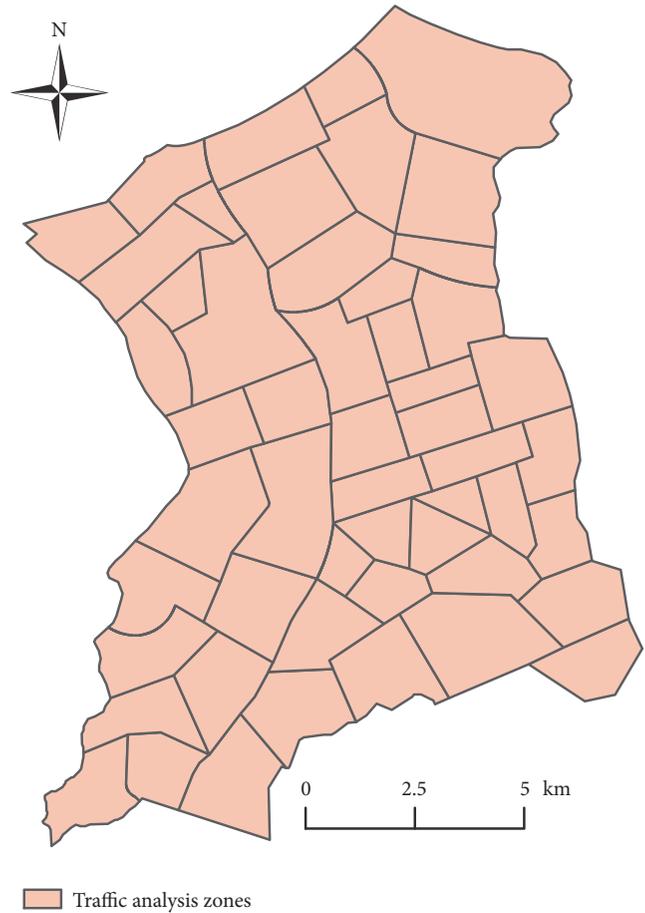


FIGURE 2: Division of Chengggong New District into traffic analysis zones.

**3.2. Data Sources.** The data adopted in this study were mainly extracted from ArcGIS, including the local public transportation and road transportation networks in Chengggong and roadway network data. To calculate the accessibility of the public transportation network, a network analysis method was used. First, a planning schematic diagram of the Chengggong New District was introduced into ArcGIS for geographic distribution and supplemented by Google Maps. The raster base maps needed in the study were obtained with the digital function of ArcGIS, in addition to the boundaries of the study area and the road transportation network area, and then vectorized in ArcGIS. Second, by querying the Baidu map to obtain bus line and bus station information, which was also vectorized in ArcGIS, a database of the transportation infrastructure in Chengggong was established. By analyzing the buffer zones of bus stations within a diameter of 300 meters and 500 meters in the study area, we clarified the availability of bus stations divided into different sections according to the boundaries of each traffic analysis zone. Finally, we searched the ArcGIS attribute table to obtain parameter values.

Because rail transit and bus transit are independent, they become synchronized only through metro stations. This

TABLE 2: Calculation results of main spatial deprivation factors.

Evaluation index	Main factor load	
	1	2
location quotient of public transportation network density	0.794	0.343
bus station occupancy	0.831	0.093
location quotient of bus station availability within a 300 meter diameter	0.905	0.172
location quotient of bus station availability within a 500 meter diameter	0.832	0.294
number of bus lines per traffic analysis zone	0.795	0.048
public transportation network coverage	0.162	0.972
<i>Eigenvalue</i>	3.435	1.191
<i>Contribution value (%)</i>	57.248	19.842
<i>Contribution rate (%)</i>	57.248	77.090

paper adopted methods based on a multimodal transportation network data set, meaning that we considered regular bus transit line and rail transit line networks to be connected only in metro stations. To establish a public transportation network and database, ArcCatalog was used to build a dataset of multiple patterns of public transportation networks. Stopping time was converted to time cost, and for the actual operations of public transportation in Kunming, we established two public transportation speeds: the bus was 20 km/h, and rail transit was 40 km/h. Finally, we calculated the shortest travel time from a bus/metro station to all other bus/metro stations and used the network analyst module in ArcMap to analyze the spatial structure of public transportation accessibility in Chenggong New District.

## 4. Results

*4.1. Spatial Deprivation Characteristics of Public Transportation.* By applying SPSS statistical analysis software for factor analysis and using orthogonal rotation to reduce the dimensions of the spatial deprivation evaluation index system of public transportation, we obtained results showing that the Kaiser-Meyer-Olkin (KMO) statistic was 0.778, meaning the partial correlation between the different variables was strong. The Bartlett spherical inspection probability was 0.00, less than 0.05 and therefore below the significance level. Therefore, we rejected the null hypothesis and found that the constructed evaluation index system of public transportation spatial deprivation was suitable for factor analysis. Furthermore, we screened out two main factors with eigenvalues exceeding one, and the cumulative explanatory variance reached 77.09%.

The calculation results of the factors are shown in Table 2. The eigenvalue of factor 1 is 3.435, and the contribution value is 57.248%, which is proportional to the density of the public transportation network. The occupation rate of the bus stations, the availability of bus stations within diameters of 300 and 500 meters, and the number of bus routes, as we mentioned before, are related to public transportation services. The eigenvalue of factor 2 is 1.191, with a contribution value of 19.842%, which is proportional to the public transportation line network coverage, indicating a relationship to the transportation infrastructure.

Figure 3 shows that two core regions have been formed based on the spatial distribution of the public transportation service factor scores: the Chenggong Downtown Area (including the Municipal Party Committee and the Longtan District) and the University Town Area. The main factor 1 scores are degressive from the core region to the outward areas, as are the public transportation services, while the spatial deprivation is the opposite (Figure 3(a)). The lower main factor 2 scores are mainly in the inner-city core areas, where the public transportation network has low coverage and where some road branches even have no public transportation lines. In contrast, most rural areas have higher scores with less road length than the core areas; the roads are used as connections to outer areas and thus have a higher coverage rate by the public transportation network (Figure 3(b)). Considering both service deprivation and infrastructure, it is apparent that the spatial deprivation pattern of public transportation in Chenggong is “multicore and local radiation” (Figure 3(c)), with the degree of public transportation spatial deprivation increasing gradually from the urban core to the periphery and the two core regions enjoying a lower level of deprivation. The spatial deprivation pattern of Chenggong is mainly due to inadequate funding to develop the overall area in a coordinated manner; thus, the government gives priority to developing areas with strong actual production demands and business opportunities. Therefore, spatial differentiation of public transportation resources exists.

*4.2. Spatial Accessibility Characteristics of Public Transportation.* Figure 4 shows that the public transportation accessibility pattern of Chenggong New District forms a circular layer along with the rail transit lines and declines progressively toward the peripheral areas (Figure 4(a)). The more peripheral an area is, the longer the time cost related to its accessibility. The time accessibility of each public transportation station was obtained by querying ArcGIS, and the public transportation accessibility level of each traffic analysis zone was then classified (Figure 4(b)), showing that the spatial diffusion outside the core region is similar to the distribution of public transportation accessibility. The areas with better transportation accessibility are mainly core regions; one of the best is Milan Garden, whose time accessibility is 15.72 minutes, while the lower-accessibility areas are focused on the

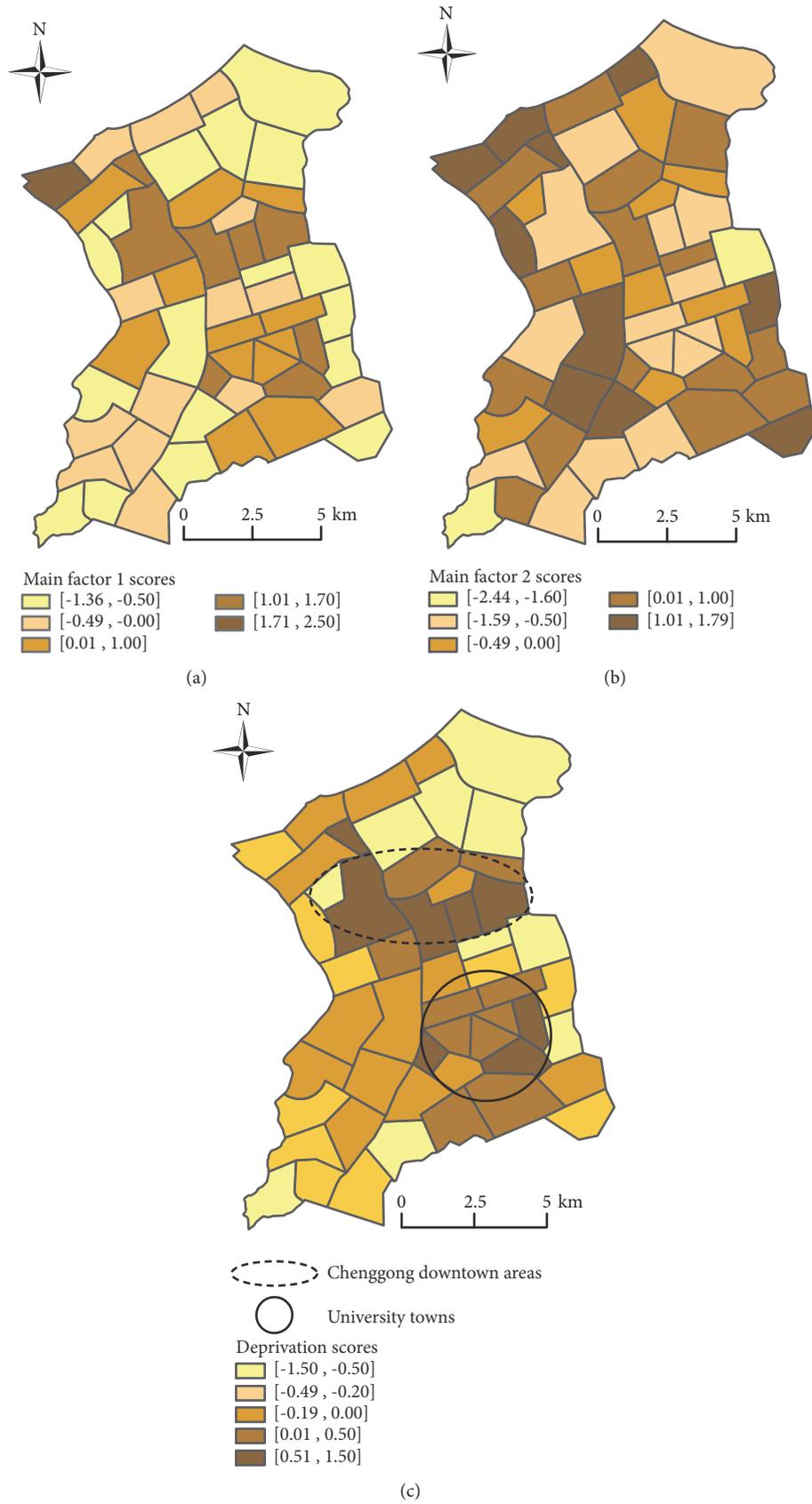


FIGURE 3: Spatial deprivation distribution of public transportation in Chenggong New District.

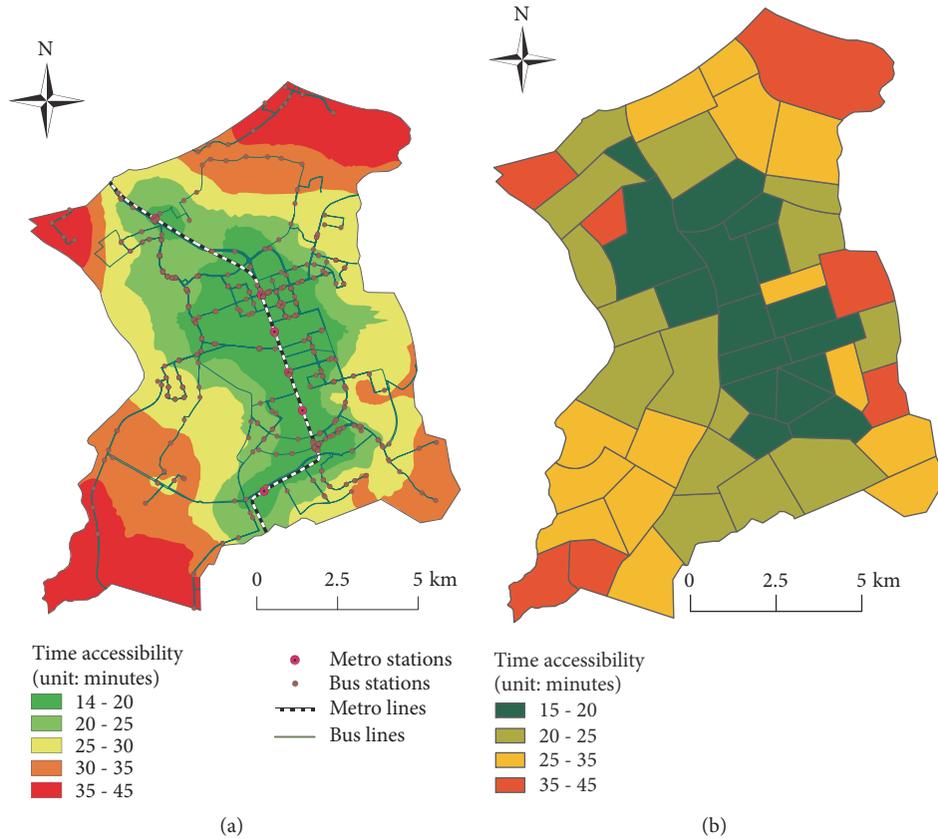


FIGURE 4: Spatial accessibility distribution of public transportation in Chenggong New District.

TABLE 3: Distribution frequency of time accessibility in traffic analysis zones.

Time (minutes)	Number of traffic analysis zones	Frequency (%)
15~20	17	32.69
20~25	15	28.85
25~33	13	25.00
35~45	7	13.46

peripheral northwestern, northeastern, southwestern, and southeastern areas. The area with the worst time accessibility is Dawan, at 43.93 minutes. There are certain special cases; for example, the time accessibility of the Plum community is 41.27 minutes, even though the area has a good accessibility atmosphere and is surrounded by areas with good time accessibility. After further investigation, we found that the poor accessibility of the Plum community is mainly due to its social organization, consisting of villages with no public transportation lines or with public transportation lines but no station.

The overall distribution frequency of time accessibility in all traffic analysis zones is presented in Table 3. Traffic analysis zones with good accessibility, meaning a time accessibility of less than 25 minutes, accounted for 61.54% of the zones, but since 38.46% of the traffic analysis zones have a time accessibility of more than 25 minutes, the overall accessibility of Chenggong New District remains low. Seven traffic analysis zones had a time accessibility that exceeded 35 minutes; the

residents of those areas experience serious spatial deprivation.

**4.3. Spatial Allocation Analysis of Public Transportation Resources.** In terms of the results calculated by the spatial clustering model, the values of Moran's  $I$  and  $Z(I)$  are 0.29 and 3.74, respectively;  $p$  is less than 0.01, which suggests that the test results are significant, and the public transportation accessibility of Chenggong is significantly positively correlated (Figure 5). According to the results, Chenggong New District can be divided into three categories as follows.

**High-High Superior.** In this type of area, traffic analysis zones and their neighborhoods have good transportation accessibility, and a positive spatial correlation is observed. Compared with other areas, the traffic analysis zones enjoy a complete infrastructure. These areas are mainly the core regions of the city.

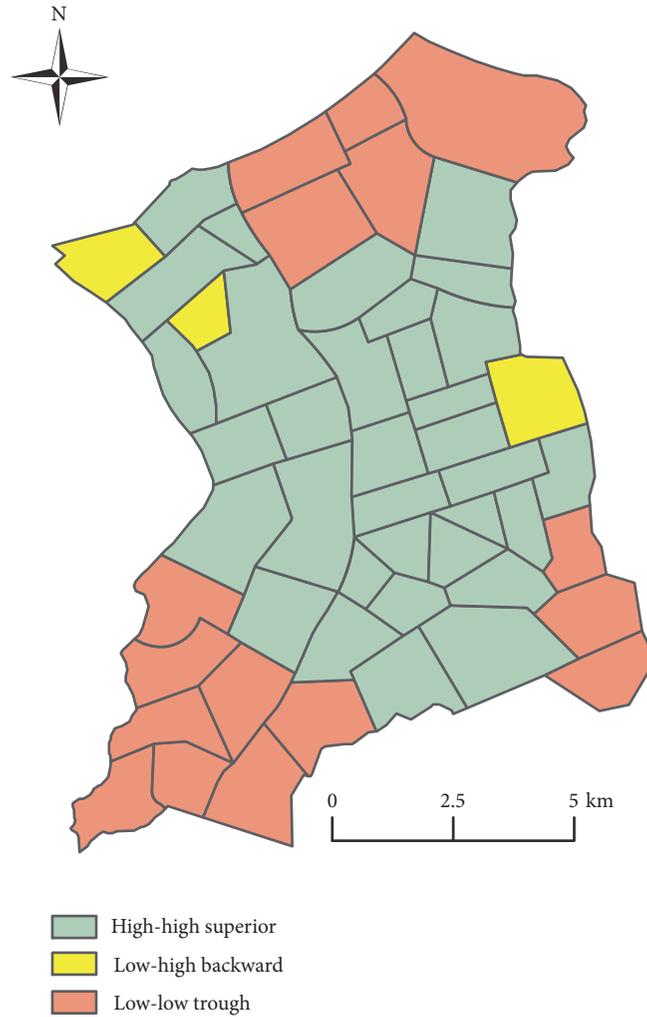


FIGURE 5: Spatial allocation pattern of public transportation resources in Chenggong New District.

*Low-High Backward.* In low-high backward areas, the traffic analysis zones have good transportation accessibility, while the neighborhoods have bad accessibility, meaning they are spatially negatively correlated. Three obvious trough areas in Chenggong New District, which also have poor transportation accessibility, are the Wangguan, Plum and White Longtan communities.

*Low-Low Trough.* Areas with poor transportation accessibility in both traffic analysis zones and neighborhoods and with a positive spatial correlation are low-low trough areas. In Chenggong, this type of area is mainly located in the northeastern, southwestern, and southeastern parts, accounting for 37%, which indicates that public transportation resource allocation in Chenggong remains obviously inequitable.

Compared with high-high superior regions, low-high backward and low-low trough areas are mainly located in the outer regions of Chenggong, which is consistent with the development pattern that usually expands outward in rapid urbanization areas.

## 5. Conclusion and Discussion

The assessment of transportation situations in rapid urbanization areas is critical to enable the creation of proactive policies for high-level planning, to reduce spatial deprivation and to improve social equity, which are urgent problems. However, few studies have focused on this issue. This study proposed a complete evaluation system based on the establishment of an index system, an assessment of transportation deprivation and an analysis of public transportation resource allocation to evaluate transportation situations in rapid urbanization areas on the basis of the actual data we could obtain. First, by considering the transportation infrastructure and transportation services that greatly influence residents' trips, an evaluation system was established that consisted of six indexes. Then, a main factor analysis was used to calculate the eigenvalues and contribution values of each evaluation index. Two main factors were chosen by SPSS to represent the public transportation network coverage and transport infrastructure services, and the spatial deprivation characteristics of public transportation in the case study

area were obtained. Second, by using the accessibility model and with the shortest time as the trip cost, ArcGIS was applied to determine the spatial structure of each traffic analysis zone. Third, global spatial autocorrelation and partial autocorrelation models were combined to analyze the spatial distribution of public transportation resources in the case study area. The results show that there are three types of resource allocation: high-high superior, low-high backward and low-low trough. These types appear most often in the outer regions of rapid urbanization areas, exhibiting obvious transportation inequity and spatial deprivation that should attract our attention.

This study proposed a useful method to evaluate the spatial deprivation of public transportation resources. This method will be beneficial for planners and managers seeking to better understand the spatial distribution of public transportation resources and develop appropriate strategies for equally allocating public transportation resource across regions to better meet residents' basic travel needs. However, the deprivation evaluation system constructed in this paper should be further discussed, as the number of indexes is relatively low. The limits of our capability to obtain more data indicate that a more comprehensive evaluation system can be created to take more deprivation factors into account, for example, the population of traffic analysis zones, average income level and/or public transportation service frequency. In addition, even though we take the transfer riding strategy between bus and rail transit into account for the transportation accessibility analysis, other factors, such as transfer fee affordability, that influence public transportation accessibility have been omitted. Thus, a certain gap with actual accessibility exists.

## Data Availability

The transportation network data used to support the findings of this study were supplied by Kunming Traffic Transport Bureau under license and so cannot be made freely available. Requests for access to these data should be made to Kunming Traffic Transport Bureau.

## Conflicts of Interest

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

## Acknowledgments

This work was jointly supported by (1) the Natural Science Foundation of China (no. 71462005); (2) the Basic Competence Improvement Project for Young and Middle-Aged Teachers in Higher Education of Guangxi (2018KY0341).

## References

- [1] China's National Bureau of Statistics. National Bureau of Statistics of China. 2014.
- [2] W. Y. Niu, *China's New-urbanization Report*, Science Press, Beijing, China, 2014.
- [3] H. Baral, R. J. Keenan, S. K. Sharma, N. E. Stork, and S. Kasel, "Economic evaluation of ecosystem goods and services under different landscape management scenarios," *Land Use Policy*, vol. 39, pp. 54–64, 2014.
- [4] K. Hubacek and J. Kronenberg, "Synthesizing different perspectives on the value of urban ecosystem services," *Landscape and Urban Planning*, vol. 109, no. 1, pp. 1–6, 2013.
- [5] C. F. Wu, Y. P. Lin, L. C. Chiang, and T. Huang, "Assessing highway's impacts on landscape patterns and ecosystem services: A case study in Puli Township, Taiwan," *Landscape and Urban Planning*, vol. 128, pp. 60–71, 2014.
- [6] 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.
- [7] L. Zhen, "Tactical berth allocation under uncertainty," *European Journal of Operational Research*, vol. 247, no. 3, pp. 928–944, 2015.
- [8] B. Lin and X. Ouyang, "Energy demand in China: Comparison of characteristics between the US and China in rapid urbanization stage," *Energy Conversion and Management*, vol. 79, pp. 128–139, 2014.
- [9] D. Scott and M. Horner, "Examining the role of urban form in shaping people's accessibility to opportunities: an exploratory spatial data analysis," *Journal of Transport and Land Use*, vol. 1, pp. 89–119, 2008.
- [10] K. Lucas, "Providing transportation for social inclusion within a framework for environmental justice in the UK," *Transportation Research Part A: Policy and Practice*, vol. 40, no. 10, pp. 801–809, 2006.
- [11] J. Preston and F. Raje, "Accessibility, mobility and transport-related social exclusion," *Journal of Transport Geography*, vol. 15, no. 3, pp. 151–160, 2007.
- [12] X. F. Ji, Y. P. Zhang, and F. Chen, "Deprivation-based transportation equity evaluation model for rapid urbanization areas," *Journal of Transportation Systems Engineering and Information Technology*, vol. 12, no. 5, pp. 7–13, 2010.
- [13] A. S. Pau, *Economics*, Post & Telecom Press, Beijing, China, 2008.
- [14] L. Wei and K. Yao, *Series of Economic Thinkers-Buchanan*, China Finance Press, Beijing, China, 2006.
- [15] J. Brocker, A. Korzhenevych, and C. Schürmann, "Assessing spatial equity and efficiency impacts of transportation infrastructure projects," *Transportation Research Part B*, vol. 44, no. 6, pp. 795–811, 2010.
- [16] M. Langford and G. Higgs, "Accessibility and public service provision: evaluating the impacts of the post office network change program in the UK," *Transactions of the Institute of British Geographers*, vol. 35, no. 4, pp. 585–601, 2010.
- [17] I. Omer, "Evaluating accessibility using house-level data: A spatial equity perspective," *Computers, Environment and Urban Systems*, vol. 30, no. 3, pp. 254–274, 2006.
- [18] E. J. Taaffe, H. L. Gauthier, and M. E. OKelly, *Geography of Transportation*, Prentice Hall, New Jersey, USA, 2nd edition, 1996.
- [19] L. Zhen, E. P. Chew, and L. H. Lee, "An integrated model for berth template and yard template planning in transshipment hubs," *Transportation Science*, vol. 45, no. 4, pp. 483–504, 2011.

- [20] T. Litman, "Integrating public health objectives in transportation decision-making," *American Journal of Health Promotion*, vol. 18, no. 1, pp. 103–108, 2003.
- [21] T. Shannon, B. Giles-Corti, T. Pikora, M. Bulsara, T. Shilton, and F. Bull, "Active commuting in a university setting: assessing commuting habits and potential for modal change," *Transport Policy*, vol. 13, no. 3, pp. 240–253, 2006.

## Research Article

# Discrete Optimization Model and Algorithm for Driver Planning in Periodic Driver Routing Problem

Lin Huang, Wenya Lv, Qian Sun, and Chengle Ma 

*School of Management, Shanghai University, Shanghai, China*

Correspondence should be addressed to Chengle Ma; [ma.chengle@foxmail.com](mailto:ma.chengle@foxmail.com)

Received 28 November 2018; Revised 31 December 2018; Accepted 10 January 2019; Published 7 February 2019

Guest Editor: Xinchang Wang

Copyright © 2019 Lin Huang 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.

Workforce planning is an operation management problem in the delivery industry to improve service quality and reliability, and the working attitude and passion of drivers, as the direct implementors of delivery service, affect the service level. Consequently, assigning equal workload for drivers so as to improve drivers' acceptance is a reasonable and efficient workforce plan for managers. This paper investigates a periodic driver routing problem to explore the relationship between workload differential among drivers and total workload; the objective of the optimization problem is to minimize the total workload. To tackle this problem, we first propose a mixed-integer linear programming model, which can be solved by an off-the-shelf mixed-integer linear programming solver, and use the local branching based method to solve larger instances of the problem. Numerical experiments are conducted to validate the effectiveness and efficiency of the proposed model and solution method, as well as the effect of small workload differential among drivers on the total workload.

## 1. Introduction

With the fierce competition of the service industry and the change in consuming ways of people, most competitors realize that decreasing service cost is not advisable for the long-term development of companies. And service reliability and customer satisfaction can be more decisive for long-term performance of service industries. Figure 1 shows the process of making profits in the service industry; we can see that drivers' satisfaction is the essential factor for increasing companies' profit, its improvement can make drivers provide high-quality service, and these two factors can be controlled by companies. Moreover, the direct influence of providing high-quality service is improving the customers' satisfaction. When customers' satisfaction rises to a high level, companies can obtain the customers' loyalty and gain long-term profit. Consequently, drivers' satisfaction is the essential internal factor, and customers' satisfaction is the most direct and important external factor for gaining profit to companies.

For improving customers' satisfaction, many companies have adopted service consistency strategies, like service time consistency. For example, many families have a long-term demand for fresh milk. And milk distributors will deliver

milk on a certain time every day; when the milk delivery time is regular, customers can predict the arrival time of milk every day and their satisfaction is improved. Moreover, the "vegetable basket" domestic project has developed to about 4000 wholesale markets for agricultural and sideline products in various provinces throughout the country. The whole country has initially formed a stable market system of "vegetable basket" which takes the central wholesale market as the core and links the production base with the retail market. In order to ensure the timely supply of vegetables and fresh products for customers, it is necessary to supply products regularly and distribute them to retail markets within a specified period of time. So we can see that the service time consistency is very important for meeting the customers daily demands.

In logistics companies, drivers are the critical executants of enhancing service quality and reliability. So, the management of drivers is critical for improving service quality. For example, the service pattern of UPS earns much concern of shipping industry in 2006; this company relies on its drivers, and the method of managing drivers is to work on the same routes to form a real bond with customers, which improves the service quality and maintains a high level of familiarity



FIGURE 1: The process of making profits in the service industry.

of customers. Many domestic companies also adopt the routes consistency for drivers. For example, school buses and scheduled company buses should keep routes consistency to ensure service quality because the routes consistency can make drivers more familiar with the driving routes so as to decrease the probability of accidents. And for the small-package shipping companies, it is also necessary to keep routes consistency. For example, S.F. Express needs to deliver more than 1,000 biological products and more than 580,000 vaccines in batches to disease prevention control centers in Western cities. In order to ensure the timely and safe arrival of drug transportation, the strategy of routes consistency is adopted to improve the drivers' familiarity with the driving routes in order to cope with traffic emergencies.

From the perspective of drivers, the realization of high service quality and reliability depends on their working aspiration, which can be enhanced by the equitable workload that companies provide, e.g., maintaining the balanced workload among drivers per day and letting each driver visit the same customers over the period. In reality, many companies focus on achieving workload balance for employees, especially in the transportation industry. For example, some companies have early, mid, and late buses to pick up and deliver employees to and from work, so drivers have three working time shifts, with each time shift at different working time. Companies will maintain the differential of the total workload among drivers in three shifts within a small range when distributing the operated routes for drivers. What is more, drivers' working content is relatively single, which is mainly measured by driving time and service time, so maintaining the workload balance among drivers is important and practical.

Aiming to improve the service quality and reliability, this paper investigates a driver routing problem (DRP) by considering the workload balance of drivers in a planning period for small-package shipping companies. In the classic vehicle routing problem, many scholars have studied the workload balance by limiting the maximum length or travel time of a route, or minimizing the differential among the longest and shortest routes (see, e.g., Mourgaya and Vanderbeck [1], Carlsson et al. [2], and Narasimha et al. [3]). But in period vehicle routing problem, as shown in Figure 2, the demands of 11 customers in the planning period  $T=3$  are denoted by  $d$ . We can see that, because of the different demands of the customers on each day, the operated routes by three drivers are different too from day 1 to day 3, which means there is a large differential in the workload among drivers over the

whole period. In order to achieve fair workload assignment, which further enhances the acceptance of drivers, this paper proposes a new strategy; i.e., in the planning period, keep the workload among drivers at the same level. In other words, if the workload difference among drivers is small, the assignment of drivers' workload is fair and more acceptable for drivers. Moreover, the proposed constraints are more flexible to maintain the relationship among drivers, because each driver can obtain the same level of payment with small workload differential.

The remainder of this paper is organized as follows. Related work is reviewed in Section 2. We describe the problem and present its mathematical model in Section 3. In Section 4, a local branching based method is used to solve larger scale instances for this problem. Numerical results are reported and discussed in Section 5 which is followed by a conclusion in Section 6.

## 2. Literature Review

The problem studied in this paper is a period driver routing problem (PDRP), which belongs to an NP-hard problem (see, e.g., Zhen [4]; Zhen, Chew, and Lee [5]). And this problem is the extension of the vehicle routing problem. Other extension problems include dynamic vehicle routing problem and stochastic time-dependent vehicle routing problem (see, e.g., Xu, Pu, and Duan [6]; Sun, Duan, and Yang [7]). In order to improve customers' satisfaction, many small-package shipping companies focus on service consistency, which attracts scholars' attention as well. Groër, Golden, and Wasil [8] first propose the consistent VRP (ConVRP), which contains driver consistency and arrival time consistency. And an algorithm based on the record-to-record travel algorithm is developed. Kovacs, Parragh, and Hartl [9] present an efficient solution method which calls template-based adaptive large neighborhood search to solve the ConVRP. And in their next published paper, cooperating with Golden [10], they propose a generalized ConVRP (GenConVRP), which relaxes the consistent constraints of ConVRP. An approach based on a flexible large neighborhood search to the entire solution is applied. In addition, they make a survey of vehicle routing problems by considering the consistent service [11]. In 2015, Kovacs, Parragh, and Hartl [12] extend the GenConVRP to a multiobjective GenConVRP by considering several independent objectives. Two exact solution approaches are proposed to solve small scale instances, and large instances are solved

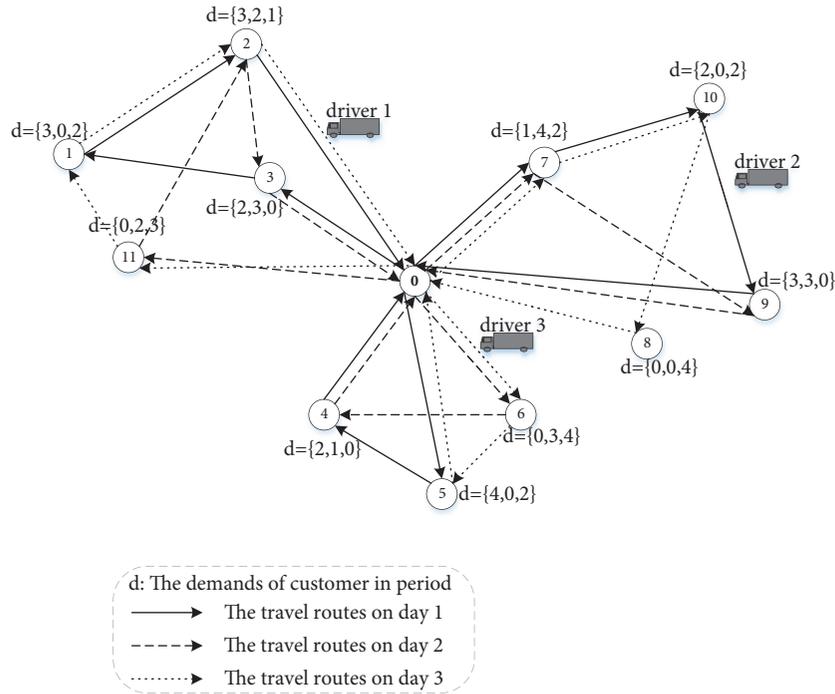


FIGURE 2: The operated routes by drivers in period.

by a heuristic which combines the multiple directional local search framework and large neighborhood search (LNS). In order to solve the same problem, Lian, Milburn and Rardin [13] develop an improved multiple directional local search (IMDLS) which makes use of the LNS to find improved solutions according to at least one objective to be added to the set of nondominated solutions at each iteration.

In other service industries, Macdonald, Dörner and Gandibleux [14] study the problem of consistent routing and scheduling for mobile nurses. The objective is minimizing the number of nurses assigned to one client. Nickel, Schröder, and Steeg [15] study short-term and mid-term home health nurse routing problems and use different heuristics combined with constraint programming strategy that aims to obtain flexible service consistency in treating different realistic constraints. Huang, Smilowitz, and Balcik [16] focus on the routing problem to assess damage and relief needs following a disaster. Campelo et. al [17] study the routes consistency of several warehouses to hundreds of orders. What should be mentioned is that the service consistency required the learning process and planning of drivers. Zhong, Hall, and Dessouky [18] study the construction of routes for local delivery of packages and create a two-stage vehicle routing model to balance the trade-offs between the vehicle dispatch consistency and the varying demand services. Coelho, Cordeau, and Laporte [19] introduce the concept of consistency in the inventory-routing problem (IRP) and discuss the situations with and without consistency requirements. And Coelho and Laporte [20] consider the consistency requirements in the multiproduct IRP and propose a branch-and-cut algorithm to solve it.

Other scholars, such as Francis, Smilowitz, and Tzur [21], Feillet et al. [22], Luo et al. [23], and Subramanyam and Gounaris [24], study PVRP with other different service consistency. Francis, Smilowitz, and Tzur [21] compare two delivery strategies for the PVRP: one is to visit each customer with his true demand, and the other is to visit each customer with his true demand or average demand, which is more flexible and consistent. The experiments indicate that cost can be decreased by adopting delivery consistency for some customers. Feillet et al. [22] define a multiday VRP called the time-consistent VRP considering the time consistency of the service. A large neighborhood search heuristic is used to solve the new problem. Luo et al. [23] investigate a PVRP with time windows and limited visiting quota. Subramanyam and Gounaris [24] develop three MIP formulations and use a class of valid inequalities to strengthen these formulations to solve the consistent traveling salesman problem.

In order to improve the service reliability, there are few studies which consider the workload balancing from a period view. Mourgaya and Vanderbeck [1] study the PVRP which aims to optimize the workload balance of each route and regional compactness of routes, and use a column generation procedure followed by a rounding heuristic to solve this problem. Gulczynski, Golden, and Wasil [25] develop a heuristic algorithm which combines the record-to-record travel algorithm with an integer program for the PVRP, and extend this heuristic algorithm to handle some real-world routing constraints such as balancing the workload of drivers and regenerating new routes. Smilowitz, Nowak, and Jiang [26] evaluate the effect

that balancing the visited customer and region for drivers has on routing costs and propose multiobjective PVRP models to balance the travel distance and driver planning.

In addition, a number of scholars have made research on the workload balancing in other VRP. Some of them aim to solve the capacitated VRP with route balancing, such as Borgulya [27], Jozefowicz, Semet, and Talbi [28], Oyola and Løkketangen [29], and Mandal et al. [30]. They study the capacitated VRP with two objectives; i.e., they aim to not only minimize the routing distance or cost, but also consider how to balance the workload of each route and propose different methods to solve this problem. Extension related to the capacitated VRP with workload balancing which considered the time window is also studied by few scholars (see, e.g., Baños et al. [31], de Freitas Aquino and Arroyo [32], and Melián-Batista et al. [33]); all of them aim to not only minimize the total travel cost but also balance the distances traveled among vehicles and propose several heuristic methods to solve this problem. In the min-max VRP, the biggest difference from other VRP is that it considered the workload balancing as the primary objective by minimizing the longest distance or traveling time among routes. And different methods are proposed by scholars (see, e.g., Carlsson et al. [2], Narasimha et al. [3], Wang, Golden, and Wasil [34], Carlsson, Carlsson, and Devulapalli [35], and Schwarze and Voß [36]).

The applications related to workload balancing are also reported in some literature. Apte and Mason [37] use the route balancing concept to develop models and heuristics for library delivery operations. Liu, Chang, and Huang [38] study the VRP with the balanced workload and delivery time of vehicles and present a multiobjective zero-one MIP to solve the problem. Huang, Smilowitz, and Balcik [39] focus on the balanced service that recipients received in humanitarian relief operations. And they also consider the goals of minimum cost as well as the quick and sufficient distribution. Zhen [40] proposes the workload balancing protocol for mitigating congestion in the optimization of container ports.

Different from the above-mentioned methods, to improve service reliability, we require that the workload differential among drivers over the period is within a small range to enhance drivers' acceptance and satisfaction, so as to enhance the service quality eventually. The contribution of this paper is summarized as follows. Firstly, this paper proposed a new strategy from the drivers' perspective to improve service quality and reliability. Secondly, a new mixed-integer linear programming (MILP) to minimize the workload of drivers, by considering the workload balance of drivers, is proposed. Thirdly, in order to solve this problem on a larger scale, we use the local branching based method to accelerate the computational time and obtain better solutions in a reasonable time. Lastly, a sensitivity analysis is conducted to explore the effect of workload differential among drivers on the total workload.

### 3. Problem Description and Model Formulation

The problem in this paper is a PDRP, so the basic background of PDRP is introduced in this section. In addition, the constraints concerning the workload balance are also given in this section. Lastly, a model formulation is presented. Different from the traditional PVRP, the problem we study is a PDRP with considering the workload balance of drivers, which aims to minimize the total workload over the period with small maximum workload differential among drivers. The workload of drivers is measured by travel time and service time.

In the standard PDRP, there is a single depot, with a homogeneous fleet departure from the single depot to serve set of customers in the planning period, and each customer has a specific demand request per day. Drivers should go back to the single depot after finishing service requirements. Each driver visits a series of customers with a limited capacity and limited travel time, and each customer cannot be visited by more than one driver each day. What should be mentioned is that the situation of a driver with no workload in one day is allowable. In order to achieve service quality and reliability, the consideration of workload balance for drivers is summarized as follows.

Keeping balanced workloads for drivers is propitious to increase the drivers' working passion and acceptance of the work schedule. Consequently, drivers' service quality and reliability are improved. The periodic workloads in this paper are measured by the total traveled time and service time during the whole planning period. The balanced strategy is setting a value of workloads' difference among drivers (denoted by  $\Theta$ ); i.e., the maximal workloads' difference among drivers is no more than  $\Theta$ . In this way, not only the service quality is improved, but also the relationship among drivers becomes more harmonious, which benefits the long-term development of small packages shipping industries.

The mathematical model is established as follow:

#### Indices and Sets

- $i, j$ : Index of a customer or the single depot;
- $N$ : Set of all customers with the single depot;
- $A$ : Set of all arcs;
- $k$ : Index of a driver;
- $K$ : Set of all drivers;
- $t$ : Index of a day;
- $T$ : Set of days in the planning horizon.

#### Parameters

- $t_{ij}$ : Travel time of the arc  $(i, j)$ ,  $i, j \in N$ .
- $s_i^t$ : Service time of customer  $i$  on day  $t$ ,  $i \in N, t \in T$
- $r_i^t$ : Demand quantity of customer  $i$  on day  $t$ ;  $i \in N \setminus \{0\}$ ,  $t \in T$ .

$e_i^t$ : Equaling one if customer  $i$  requires service on day  $t$  ( $t_i^t > 0$ ), otherwise zero.

$C$ : Capacity of the vehicle.

$L$ : Maximal workload of each driver per day

$\Theta$ : The allowed maximal workloads' difference among drivers.

### Decision Variables

$\alpha_{ijk}^t = \{0, 1\}$ : Equaling one if driver  $k$  traverses arc  $(i, j \in A)$  on day  $t \in T$ , and zero otherwise;

$\beta_{ki}^t = \{0, 1\}$ : Equaling one if driver  $k$  serves customer  $i$  on day  $t \in T$ , and zero otherwise;

$\delta_i^t \geq 0$ : A nondecreasing indicator, showing the sequence to visit each customer on day  $t$  (for subtour elimination purpose);

$\omega_k^t \geq 0$ : Workload of driver  $k$  on day  $t$ ;

$\lambda_k = \{0, 1\}$ : Equaling one if driver  $k$  is not idle over the period, and zero otherwise.

### MILP Formulation

$$(P1) \min \sum_{t \in T} \sum_{k \in K} \sum_{(i,j) \in A} t_{ij} \alpha_{ijk}^t + \sum_{k \in K} \sum_{i \in N \setminus \{0\}} \sum_{t \in T} s_i^t \beta_{ki}^t \quad (1)$$

$$\text{s.t.} \quad \sum_{j \in N} \alpha_{ijk}^t = \beta_{ki}^t \quad (2)$$

$$\forall k \in K, \forall i \in N \setminus \{0\}, \forall t \in T;$$

$$\sum_{j \in N \setminus \{0\}} \alpha_{0jk}^t \leq 1 \quad \forall k \in K, \forall t \in T; \quad (3)$$

$$\sum_{k \in K} \beta_{ki}^t = e_i^t \quad \forall i \in N \setminus \{0\}, \forall t \in T; \quad (4)$$

$$\sum_{i \in N} r_i^t \beta_{ki}^t \leq C \quad \forall k \in K, \forall t \in T; \quad (5)$$

$$\sum_{(i,j) \in A} t_{ij} \alpha_{ijk}^t + \sum_{i \in N \setminus \{0\}} s_i^t \beta_{ki}^t \leq L \quad (6)$$

$$\forall k \in K, \forall t \in T$$

$$\sum_{j \in N} \alpha_{ijk}^t - \sum_{j \in N} \alpha_{jik}^t = 0 \quad (7)$$

$$\forall k \in K, \forall i \in N, \forall t \in T;$$

$$\sum_{(i,j) \in A} t_{ij} \alpha_{ijk}^t + \sum_{i \in N \setminus \{0\}} s_i^t \beta_{ki}^t = \omega_k^t \quad (8)$$

$$\forall k \in K, \forall t \in T;$$

$$\beta_{ki}^t \leq \lambda_k \quad \forall k \in K, \forall i \in N \setminus \{0\}, \forall t \in T \quad (9)$$

$$\begin{aligned} & \sum_{t \in T} \omega_k^t - \sum_{t \in T} \omega_{k'}^t \\ & \leq \Theta + M(1 - \lambda_k) + M(1 - \lambda_{k'}) \end{aligned} \quad (10)$$

$$\forall k, k' \in K;$$

$$\delta_j^t \geq \delta_i^t + 1 - |N| \left( 1 - \sum_{r \in R} \alpha_{ijk}^t \right) \quad (11)$$

$$\forall (i, j) \in A : j \neq 0;$$

$$\alpha_{ijk}^t, \beta_{ki}^t = \{0, 1\} \quad (12)$$

$$\delta_i^t, \omega_k^t, \lambda_k \geq 0. \quad (13)$$

Objective (1) minimizes the total workload of all the drivers. The workload of drivers includes two parts: one is the traveled time, and the other is the service time. Constraints (2) ensure that customers are visited at most once on day  $t$ . Constraints (3) state that the drivers with the workload on day  $t$  should start from the depot '0' to visit customers, or this driver does not have workload on that day. Constraints (4) ensure that if a customer has demand on day  $t$ , then this customer must be visited by a driver on that day. Constraints (5) are the capacity constraints, which guarantee that the served total load requirement of a driver on day  $t$  should not exceed a vehicle's load capacity. Constraints (6) are the travel time constraints, which guarantee that the total workload of a driver on one day is no more than the maximal workload. Constraints (7) state that each customer has only one predecessor and one successor. Constraints (8) state the definition of the workload of drivers, which contains the traveled time and service time. Constraints (9) guarantee that if a driver has worked on one day, then this driver is not idle over the planning period. Constraints (10) guarantee that any two drivers' workload should not exceed maximal workload difference. Constraints (11) serve to eliminate subtours in the individual daily routes. Constraints (12)-(13) define the decision variables.

## 4. Solution Methodology

The local branching based method plays a very important role in improving the incumbent solution as early as possible during the computation when solving difficult MIP problems. It adopts a two-level branching strategy whose high-level strategic branching defines solution neighborhoods and low-level tactical branching explores them. The core idea is to find an initial solution and use the CPLEX solver as a 'lower-level' black-box tool for exploring solution subspaces, which are then controlled at a 'higher-level' by an external branching framework. The difference from other solution methods is that the neighborhoods are formulated by introducing some linear inequalities, which are called 'local branching cuts'. In our study, the local branching based method is to branch the binary variable  $\beta_{ki}^t$  within its solution space, because the variable  $\beta_{ki}^t$  is the critical decision variable in the model, as it determines the remaining variables such as  $\alpha_{ijk}^t$ . The basic

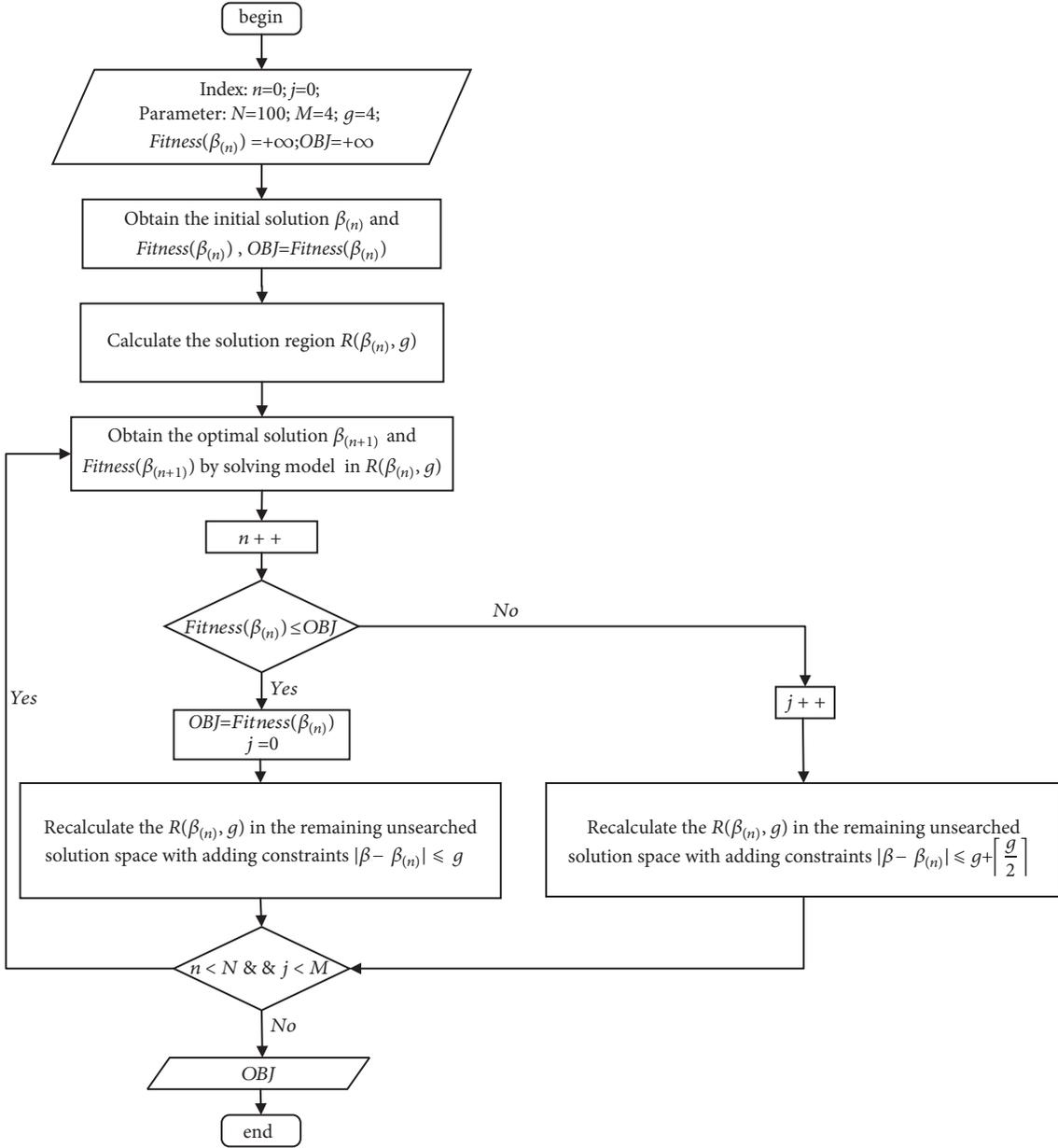


FIGURE 3: The basic procedure of local branching.

procedure of this local branching based method is shown in Figure 3.

In Figure 3, several important parameters should be determined firstly; therinto,  $g$  is the neighborhood-size parameter, and it should be chosen as small to reduce the computing time, but still large enough to ensure that better solutions can be found.  $N$  is the maximal number of iterations, and  $M$  is the allowable consecutive number of unimproved solutions found in the iterating process; both are the stopping conditions of the whole procedure; once one of them is exceeded, the procedure stops.  $OBJ$  is the objective value. In this procedure,  $\beta_{[n]}$  denotes the optimal solution that can be found in the  $n^{\text{th}}$  iteration, and the initial solution is denoted by  $\beta_{[n]}$  where  $n = 0$ . The fitness value in the  $n^{\text{th}}$

iteration is denoted by  $Fintness(\beta_{[n]})$ . The initial solution is obtained by initializing the binary variables  $\beta_{ki}^t$  ( $\forall k \in K, i \in N, t \in T$ ) when solving the model proposed in Section 3; then a solution region  $R(\beta_{[0]}, g)$  is calculated by a constraint  $|\beta - \beta_{[0]}| \leq g$ . Here  $|\beta - \beta_{[0]}|$  reflects the radius of  $\beta_{[0]}$ 's neighborhood in the solution space, and it is calculated as follows:  $|\beta - \beta_{[0]}| = \sum_{k \in K} \sum_{i \in N} \sum_{t \in T} |\beta_{ki}^t - \beta_{ki}^{t[0]}|$ . After that, the improved solution  $\beta_{[1]}$  will be obtained by solving model in the  $R(\beta_{[0]}, g)$ . In the next iteration, as the searched space (i.e.,  $|\beta - \beta_{[0]}| \leq g$ ) will no longer be searched, the solution region  $R(\beta_{[1]}, g)$  will be recalculated by constraints  $|\beta - \beta_{[0]}| \geq g + 1$  and  $|\beta - \beta_{[1]}| \leq g$ . And the optimal solution  $\beta_{[2]}$  obtained by solving the same model in a new solution

**The procedure of handling strategy to two cases:**  
**if**  $elapsed\_time(n) \geq time\_limit$  **then** //  $elapsed\_time(n)$  is the running time of solving model  
in the solution region  $R(\beta_{[n]}, g)$ .  
    **if**  $Fitness(\beta_{[n]}) < OBJ$  **then**  
         $OBJ \leftarrow Fitness(\beta_{[n]})$ ;  
        Recalculate the  $R(\beta_{[n]}, g)$  by the constraints  $|\beta - \beta_{[0]}| \geq g + 1, \dots, |\beta - \beta_{[n-2]}| \geq g + 1, |\beta - \beta_{[n]}| \leq g$ ;  
    **else**  
        Recalculate the  $R(\beta_{[n]}, g)$  by the constraints  $|\beta - \beta_{[0]}| \geq g + 1, \dots, |\beta - \beta_{[n-2]}| \geq g + 1, |\beta - \beta_{[n-1]}| \leq \lfloor g/\alpha \rfloor$  //  $\alpha$  is an integer.  
    **end if**  
**end if**

PSEUDOCODE 1

region  $R(\beta_{[1]}, g)$  will replace the  $\beta_{[1]}$  if  $Fitness(\beta_{[2]})$  is better than the  $OBJ$ . Repeating the updating procedure until the stop condition is reached, then we will obtain the final objective value.

The basic procedure of local branching method has two points to be noted: One is related to the solution region. As the searched space will no longer be searched again, the solution region will be recalculated according to constraints  $|\beta - \beta_{[0]}| \geq g + 1, \dots, |\beta - \beta_{[n-1]}| \geq g + 1, |\beta - \beta_{[n]}| \leq g$  in each iteration. The other is the situation that the solution found is worse than the incumbent solution in an iteration; once the situation occurs, a diversification consisting in enlarging the current solution region can be applied, e.g., changing the constraint  $|\beta - \beta_{[n]}| \leq g$  into  $|\beta - \beta_{[n]}| \leq g + \lceil g/2 \rceil$ .

Imposing a time limit on the subprocedure of solving model in the solution region  $R(\beta_{[n]}, g)$  will enhance heuristic performance of the local branching method. In case the time limit is exceeded, there will be two cases.

(1) The incumbent solution has been improved when the time limit is reached. In this situation, the handling strategy is solving the model in the solution region calculated by  $|\beta - \beta_{[0]}| \geq g + 1, \dots, |\beta - \beta_{[n-2]}| \geq g + 1, |\beta - \beta_{[n]}| \leq g$  in the  $n^{\text{th}}$  iteration, and the solution region in the next iteration will be within the remaining searchable space with a constraint  $|\beta - \beta_{[n+1]}| \leq g$ .

(2) The time limit is reached with unimproved solution. Once it happened, we can reduce the size of the solution region to speed up its exploration.

The handling procedure of the above two cases is illustrated by Pseudocode 1.

The basic procedure and enhancing strategies are the whole content of local branching method. Though not complex, the method proved quite effective when solving difficult MIPs.

## 5. Computational Experiments

This section presents a series of experiments to validate the effectiveness of the MILP model and the local branching based method. In Section 5.1, we solve a set of small problems and compare the results obtained by local branching based

method with the objective values obtained by CPLEX. And we solve a number of small and medium scale instances by local branching based method in Section 5.2 to conduct a sensitivity analysis for exploring the effect of workload differential on the total workload of drivers. To validate the efficiency of the proposed MILP model and solution method, we perform numerical experiments on a PC (Intel Xeon Gold 6154 Processor, 3.00G Hz; Memory, 512G). The algorithm is implemented by C# using the CPLEX of version 12.5.

*5.1. Performance Comparison of CPLEX and Local Branching Based Method.* The small scale data consists of five problems with 10 customer locations and five problems with 12 customer locations. The probability that a customer requires service on a given day was set to 70%, and the planning period is  $T=3$ . The data about the locations of customers and depot and the number of customers' demands is randomly generated. In addition, the allowed working time for each driver per day is  $L=35$ , and the vehicle's capacity is  $C=15$  for all problems. The maximal workload differential among drivers is set as  $\Theta=10$ , and the available number of divers is  $K=2$  in 10 customers and  $K=3$  in 12 customers to guarantee that all the instances can be solved. We solve these instances by CPLEX and local branching based method, and the results are shown in Table 1.

As shown in Table 1, column ' $GAP_L$ ' records the gaps between the solutions obtained by local branching based method and the solutions obtained by CPLEX's MILP solver. Note that, if the value of ' $GAP_L$ ' is negative, it implies the local branching based method is better than the CPLEX's solution. From Table 1, we can see that CPLEX can find feasible solutions of all instances within two hours. In comparison, the local branching based method can mostly obtain better solutions in a quite short time period. Moreover, with the instances' scale increasing, the average improvement ratio of local branching based method is from 7.33% to 27.63%, and the computational time is slightly increased. So, we can conclude that the performance of local branching based method is much better than the CPLEX, especially on the larger scale. In the next section, we conduct sensitive analysis by using local branching based method.

TABLE 1: The performance comparison of CPLEX and local branching based method.

Instances ID	CPLEX		Local Branching		$GAP_L$	$T_L/T_C$
	$Z_C$	$T_C$	$Z_L$	$T_L$		
10_1	138.0	7200	138.0	615	0.00%	0.09
10_2	124.0	7200	118.7	604	-4.27%	0.08
10_3	151.7	7200	148.3	605	-2.24%	0.08
10_4	158.5	7200	148.6	607	-6.25%	0.08
10_5	160.8	7200	122.3	603	-23.90%	0.08
Average					-7.33%	0.08
12.1	171.0	7200	161.3	915	-5.67%	0.13
12.2	165.9	7200	110.9	605	-33.20%	0.08
12.3	187.7	7200	129.5	608	-31.01%	0.08
12.4	225.9	7200	165.0	606	-26.96%	0.08
12.5	223.9	7200	131.4	606	-41.31%	0.08
Average					-27.63%	0.09

**Notes:** (1) The ID '10-1' means there are 10 customers in this instance, and '1' is the index of the instance. (2) Objective value and CPU time for CPLEX are denoted by  $Z_C$  and  $T_C$ . Objective value and CPU time for local ranching based method are denoted by  $Z_L$  and  $T_L$ , and  $GAP_L = (Z_L - Z_C) / Z_C$ .

TABLE 2: sensitivity analysis on workload differential.

ID	K	No consideration of $\theta$			Consideration of $\theta$			gap
		$Z_L$	$T_L$	$(1/5)\omega$	$\theta$	$Z_L'$	$T_L'$	
10-1	2	138.0	439	4.60	3	138.0	1930	0.00%
10-2	2	118.6	1960	3.95	3	118.8	1949	0.17%
10-3	2	148.2	1938	4.94	3	148.3	1938	0.07%
10-4	2	148.6	1929	4.95	3	151.3	1946	1.78%
Average			1567				1941	0.51%
15-1	2	809.4	2409	26.98	10	811.1	2714	0.21%
15-2	2	963.8	2470	32.13	12	966.4	2430	0.27%
15-3	3	1188.3	2471	26.41	10	1296.7	3405	8.36%
15-4	2	751.5	1809	25.05	10	751.6	2411	0.01%
Average			2290				2740	2.21%
20-1	2	861.9	1907	28.73	20	862.4	2510	0.06%
20-2	3	1121.4	2517	24.92	20	1238.6	2210	9.46%
20-3	3	1322.1	2235	29.38	20	1328.0	4645	0.44%
20-4	2	876.9	2444	30.20	20	879.6	3362	0.31%
Average			2276				3182	2.57%

**Notes:** (1) 'K' is the number of drivers; ' $\omega$ ' denotes the average workload of each driver on a day, which is obtained by the function  $\omega = Z_L/(T \times K)$ . (2)  $\theta$  denotes the value of maximal workload differential among drivers which is the input parameter to solve the problem. Here,  $\theta < (1/5)\omega$ . (3) 'gap' is calculated by the function  $gap = (Z_L' - Z_L)/Z_L$ ;  $Z_L$  and  $T_L$  are the objective value and CPU time of local branching which do not consider small workload differential, respectively; and  $Z_L'$  and  $T_L'$  are the objective value and CPU time of local branching which consider small workload differential, respectively.

**5.2. Sensitivity Analysis.** In order to explore the effect of small workload differential among drivers on the total workload, we conduct a sensitivity analysis on this section. Firstly, we generated three group instances which range from 10 customers to 20 customers. In each group, there are four instances. The data about locations of customers and depot and the number of customers' demands are randomly generated as well. The planning period in all instances is  $T=3$ . In addition, the allowed working time for each driver per day is  $L=35$ , the vehicle's capacity is  $C=15$  for small scale instances (10 customers), and  $C=160$ ,  $L=200$  in medium scale instances (15-20 customers). Secondly, we do not consider the workload differential among drivers, and we solve these instances

to obtain the minimal total workload. After obtaining the minimal total workload, in order to explore the effect of small workload differential among drivers on the total workload, we set the maximal workload differential among drivers to be no more than the one-fifth of the average workload of a driver on a day, solve these instances, and obtain the updated objectives. The results are shown in Table 2.

From Table 2 we can see that the objective value of considering small workload differential among drivers is larger than the objective value of no limit to workload differential among drivers, which means that achieving workload balance will increase the total workload (cost). With the instances' scale increasing, the gap between no consideration

and consideration workload differential becomes larger, the average gap is from 0.51% to 2.57%, and the computational time is longer as well. What should be mentioned is that the gap on the small and medium scale is in an accepted range, so it is advisable for managers to achieve a balanced workload for drivers with small increased cost.

## 6. Conclusion

This study examines an operations management problem of shipping industry which aims to balance the workload among drivers. For solving this problem, a mixed-integer linear programming model is proposed. Different from the general vehicle routing problem, this model adds a new constraint to limit the workload differential among drivers. Moreover, we use a local branching based method to solve this problem on different scales, which can obtain better and faster solutions for large instances of the problem. To validate the efficiency of the proposed model and the local branching based method, numerical experiments are conducted. The results indicate that local branching based method can accelerate the solving time and obtain better solutions compared with CPLEX. Moreover, from the result of sensitivity analysis, we can discover that considering a balanced workload among drivers will lead to an increase in cost, but the increased cost is in an accepted range. For managers of the delivery industry, it is advisable to balance workload among drivers with fewer increased cost, so as to improve the service quality and reliability.

The problem we studied was to minimize the total workload of all drivers with considering small workload differential among drivers, which provides a new perspective to improve service quality and reliability. At present, we used local branching based method to solve this problem on a small and medium scale. In the future, we aim to design a new heuristic algorithm to solve this problem on a larger scale in a reasonable time. Moreover, the essential goal is improving service quality and reliability, so we will consider more factors like consistent service with different implementation in our research.

## Data Availability

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

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] M. Mourgaya and F. Vanderbeck, "Column generation based heuristic for tactical planning in multi-period vehicle routing," *European Journal of Operational Research*, vol. 183, no. 3, pp. 1028–1041, 2007.
- [2] J. Carlsson, D. Ge, A. Subramaniam, A. Wu, and Y. Ye, "Solving min-max multi-depot vehicle routing problem," *Lectures on Global Optimization*, vol. 55, pp. 31–46, 2009.
- [3] K. V. Narasimha, E. Kivelevitch, B. Sharma, and M. Kumar, "An ant colony optimization technique for solving min-max multi-depot vehicle routing problem," *Swarm and Evolutionary Computation*, vol. 13, pp. 63–73, 2013.
- [4] L. Zhen, "Tactical berth allocation under uncertainty," *European Journal of Operational Research*, vol. 247, no. 3, pp. 928–944, 2015.
- [5] L. Zhen, E. P. Chew, and L. H. Lee, "An integrated model for berth template and yard template planning in transshipment hubs," *Transportation Science*, vol. 45, no. 4, pp. 483–504, 2011.
- [6] H. Xu, P. Pu, and F. Duan, "Dynamic vehicle routing problems with enhanced ant colony optimization," *Discrete Dynamics in Nature and Society*, vol. 2018, Article ID 1295485, 13 pages, 2018.
- [7] S. Sun, Z. Duan, and D. Yang, "Urban freight management with stochastic time-dependent travel times and application to large-scale transportation networks," *Discrete Dynamics in Nature and Society*, vol. 2015, Article ID 914309, 10 pages, 2015.
- [8] C. Groër, B. Golden, and E. Wasil, "The consistent vehicle routing problem," *Manufacturing and Service Operations Management*, vol. 11, no. 4, pp. 630–643, 2009.
- [9] A. A. Kovacs, S. N. Parragh, and R. F. Hartl, "A template-based adaptive large neighborhood search for the consistent vehicle routing problem," *Networks*, vol. 63, no. 1, pp. 60–81, 2014.
- [10] A. A. Kovacs, B. L. Golden, R. F. Hartl, and S. N. Parragh, "The generalized consistent vehicle routing problem," *Transportation Science*, vol. 49, no. 4, pp. 796–816, 2015.
- [11] A. A. Kovacs, B. L. Golden, R. F. Hartl, and S. N. Parragh, "Vehicle routing problems in which consistency considerations are important: A survey," *Networks*, vol. 64, no. 3, pp. 192–213, 2014.
- [12] A. A. Kovacs, S. N. Parragh, and R. F. Hartl, "The multi-objective generalized consistent vehicle routing problem," *European Journal of Operational Research*, vol. 247, no. 2, pp. 441–458, 2015.
- [13] K. Lian, A. B. Milburn, and R. L. Rardin, "An improved multi-directional local search algorithm for the multi-objective consistent vehicle routing problem," *Institute of Industrial Engineers (IIE) Transactions*, vol. 48, no. 10, pp. 975–992, 2016.
- [14] T. Macdonald, K. Dörner, and X. Gandibleux, *Metaheuristics for the consistent nurse scheduling and routing problem*, [Doctoral dissertation, MS Thesis], Université de Nantes, Nantes, France, 2009.
- [15] S. Nickel, M. Schröder, and J. Steeg, "Mid-term and short-term planning support for home health care services," *European Journal of Operational Research*, vol. 219, no. 3, pp. 574–587, 2012.
- [16] M. Huang, K. R. Smilowitz, and B. Balcik, "A continuous approximation approach for assessment routing in disaster relief," *Transportation Research Part B: Methodological*, vol. 50, pp. 20–41, 2013.
- [17] P. Campelo, F. Neves-Moreira, P. Amorim, and B. Almada-Lobo, "Consistent vehicle routing problem with service level agreements: a case study in the pharmaceutical distribution sector," *European Journal of Operational Research*, vol. 273, no. 1, pp. 131–145, 2019.
- [18] H. Zhong, R. W. Hall, and M. Dessouky, "Territory planning and vehicle dispatching with driver learning," *Transportation Science*, vol. 41, no. 1, pp. 74–89, 2007.
- [19] L. C. Coelho, J.-F. Cordeau, and G. Laporte, "Consistency in multi-vehicle inventory-routing," *Transportation Research Part C: Emerging Technologies*, vol. 24, pp. 270–287, 2012.
- [20] L. C. Coelho and G. Laporte, "A branch-and-cut algorithm for the multi-product multi-vehicle inventory-routing problem,"

- International Journal of Production Research*, vol. 51, no. 23-24, pp. 7156–7169, 2013.
- [21] P. Francis, K. Smilowitz, and M. Tzur, “Flexibility and complexity in periodic distribution problems,” *Naval Research Logistics (NRL)*, vol. 54, no. 2, pp. 136–150, 2007.
- [22] D. Feillet, T. Garaix, F. Lehuédé, O. Péton, and D. Quadri, “A new consistent vehicle routing problem for the transportation of people with disabilities,” *Networks*, vol. 63, no. 3, pp. 211–224, 2014.
- [23] Z. Luo, H. Qin, C. Che, and A. Lim, “On service consistency in multi-period vehicle routing,” *European Journal of Operational Research*, vol. 243, no. 3, pp. 731–744, 2015.
- [24] A. Subramanyam and C. E. Gounaris, “A branch-and-cut framework for the consistent traveling salesman problem,” *European Journal of Operational Research*, vol. 248, no. 2, pp. 384–395, 2016.
- [25] D. Gulczynski, B. Golden, and E. Wasil, “The period vehicle routing problem: new heuristics and real-world variants,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 47, no. 5, pp. 648–668, 2011.
- [26] K. Smilowitz, M. Nowak, and T. Jiang, “Workforce management in periodic delivery operations,” *Transportation Science*, vol. 47, no. 2, pp. 214–230, 2013.
- [27] I. Borgulya, “An algorithm for the capacitated vehicle routing problem with route balancing,” *Central European Journal of Operations Research*, vol. 16, no. 4, pp. 331–343, 2008.
- [28] N. Jozefowicz, F. Semet, and E.-G. Talbi, “An evolutionary algorithm for the vehicle routing problem with route balancing,” *European Journal of Operational Research*, vol. 195, no. 3, pp. 761–769, 2009.
- [29] J. Oyola and A. Løkketangen, “GRASP-ASP: An algorithm for the CVRP with route balancing,” *Journal of Heuristics*, vol. 20, no. 4, pp. 361–382, 2014.
- [30] S. K. Mandal, D. Pacciarelli, A. Løkketangen, and G. Hasle, “A memetic NSGA-II for the bi-objective mixed capacitated general routing problem,” *Journal of Heuristics*, vol. 21, no. 3, pp. 359–390, 2015.
- [31] R. Baños, J. Ortega, C. Gil, A. L. Márquez, and F. De Toro, “A hybrid meta-heuristic for multi-objective vehicle routing problems with time windows,” *Computers & Industrial Engineering*, vol. 65, no. 2, pp. 286–296, 2013.
- [32] R. De Freitas Aquino and J. E. C. Arroyo, “A hybrid multi-objective iterated local search heuristic for vehicle routing problem with time windows,” in *Proceedings of the 14th International Conference on Hybrid Intelligent Systems, HIS 2014*, pp. 117–122, Kuwait, December 2014.
- [33] B. Melián-Batista, A. De Santiago, F. Angelbello, and A. Alvarez, “A bi-objective vehicle routing problem with time windows: A real case in Tenerife,” *Applied Soft Computing*, vol. 17, pp. 140–152, 2014.
- [34] X. Wang, B. Golden, and E. Wasil, “The min-max multi-depot vehicle routing problem: Heuristics and computational results,” *Journal of the Operational Research Society*, vol. 66, no. 9, pp. 1430–1441, 2014.
- [35] J. G. Carlsson, E. Carlsson, and R. Devulapalli, “Balancing workloads of service vehicles over a geographic territory,” in *Proceedings of the 26th IEEE/RSJ International Conference on Intelligent Robots and Systems: New Horizon, IROS 2013*, pp. 209–216, IEEE, Japan, November 2013.
- [36] S. Schwarze and S. Voß, “Improved load balancing and resource utilization for the Skill Vehicle Routing Problem,” *Optimization Letters*, vol. 7, no. 8, pp. 1805–1823, 2013.
- [37] U. M. Apte and F. M. Mason, “Analysis and improvement of delivery operations at the San Francisco Public Library,” *Journal of Operations Management*, vol. 24, no. 4, pp. 325–346, 2006.
- [38] C. M. Liu, T. C. Chang, and L. F. Huang, “Multi-objective heuristics for the vehicle routing problem,” *International Journal of Operations Research*, vol. 3, no. 3, pp. 173–181, 2006.
- [39] M. Huang, K. Smilowitz, and B. Balcik, “Models for relief routing: equity, efficiency and efficacy,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 1, pp. 2–18, 2012.
- [40] 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

# Research on the Scheduling Problem of Movie Scenes

Yulian Liu,<sup>1</sup> Qiuji Sun <sup>1</sup>, Xiaotian Zhang,<sup>1,2</sup> and Yiwei Wu <sup>1</sup>

<sup>1</sup>School of Management, Shanghai University, Shanghai, China

<sup>2</sup>Oulu Business School, University of Oulu, Finland

Correspondence should be addressed to Yiwei Wu; [wyw0044@shu.edu.cn](mailto:wyw0044@shu.edu.cn)

Received 28 November 2018; Accepted 10 January 2019; Published 3 February 2019

Guest Editor: Xinchang Wang

Copyright © 2019 Yulian 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.

The scheduling problem of movie scenes refers to the reasonable ordering of the shooting sequence of the scene, thereby minimizing the total cost of the movie scenes. This paper considers the factors affecting the cost of movie scenes shooting in the real world and constructs an integer linear programming model. A tabu search based method (TSBM) and a particle swarm optimization based method (PSOBM) are designed to solve larger-scale problems. Numerical experiments show that both TSBM and PSOBM are suitable for solving small-scale movie scenes scheduling problem. Moreover, through the comparison experiments between TSBM and PSOBM, it is verified that the TSBM can effectively solve the large-scale movie scenes scheduling problem.

## 1. Introduction

In the real world, the sequence of movie scenes shoot is not the final sequence. Therefore, we do this research in order to find the optimal sequence. The movie scenes scheduling problem can be described as follows. Let  $S = \{s_1, s_2, \dots, s_n\}$  represent a set of  $n$  scenes and  $A = \{a_1, a_2, \dots, a_m\}$  denote a set of  $m$  actors. Let  $D = \{d_1, d_2, \dots, d_n\}$  show the duration of  $n$  scenes and  $P = \{p_1, p_2, \dots, p_m\}$  mean the wage of  $m$  actors in one day. The actor shot from the first scene belongs to his last scene. Meanwhile, the actors are in the crew. During this period, the actor will be paid every day. Even in some days, they may not be shot. The movie scenes scheduling problem is how to reduce the transfer cost in consideration of actors with different wages so as to minimize the total cost.

The movie scenes scheduling problem is a multifactor optimization problem. Most literatures study the scheduling problem at the same location, taking no account of the impact of different locations on the sequence of the scene. It is very likely that the transfer cost of the scene location will be generated. In this paper, the location of the movie scenes is considered. In addition, this paper also considers different wage of actors and different scenes shooting duration in order to construct an integer linear programming model. Then, this paper designs a tabu search based method (TSBM) and a particle swarm optimization based method (PSOBM) to solve

large-scale problems, comparing them through a series of numerical experiments.

The remainder of this paper is organized as follows. Section 2 is related works. In Section 3, we build an integer programming model for the movie scenes scheduling problem. Then, we establish a TSBM in Section 4. Section 5 develops a PSOBM to handle the movie scenes scheduling problem. The experiment results and settings are shown in Section 6. In Section 7 we conclude the whole paper.

## 2. Related Works

During the movie scenes shooting process, many factors affect the order in which movie scenes are shot, such as the cost of the actors, the duration of the scenes shot, and the duration of the actors. In addition, cost control is one of the issues that must be considered. Optimizing the sequence of scenes in the movie scenes shooting process, reducing the waiting time for the actors, and minimizing the total cost of the crew are the overall goals.

The movie scenes scheduling problem was proposed by Cheng et al. [1]. Considering the different daily wage of actors, the shooting duration of each scene was the same value. A joint model was proposed to minimize the waiting time of actors as well as the total minimized costs. And they used the branch and bound algorithm and heuristic algorithm

to solve large-scale problems. Nordström and Tufekci [2] applied hybrid genetic algorithm to solve the problem of movie scenes scheduling based on the research of Cheng et al. [1] and verified that the proposed algorithm is better than the heuristic algorithm of Cheng et al. [1]. Fink and Voß [3] solved the movie scenes scheduling problem as a special application of the general pattern ordering problem and used simulated annealing algorithm and TS to solve this problem. Bomsdorf and Derigs [4] put forward a conceptual model to solve the scheduling problem in the movie scenes shooting process and used a heuristic algorithm to generate a feasible scheme. Finally, a decision support system is proposed and the effectiveness of the decision support system is proved.

Some scholars assumed that each scene shooting duration is different, and different actors have different daily wages. de la Banda et al. [5] described the scheduling problem. In their study, actor's daily wage was different. And scenes may have different durations. They proposed the basic dynamic programming algorithm, through the preprocessing, search limits, to determine the upper and lower bound to improve the algorithm. Preprocessing mainly includes two simplified methods. On the one hand, referring to the same actors in the two scenes, you can combine the two scenes shooting duration and reduce the calculation. On the other hand, if any actor appears in only one scene, then the cost of the actor can be calculated as a fixed cost. Kochetov [6] studied the movie scenes scheduling problem in movie shooting of the column transformation problem that can be reconstructed into a binary matrix and solved it using the improved greedy algorithm and metaheuristic algorithm of iterative local search. Liang et al. [7] proposed a branch and bound algorithm to solve the movie scenes scheduling problem. They used a new lower bound function to delete redundant search nodes. Based on de la Banda et al. [5], Qin et al. [8] proposed an enhanced branch and bound algorithm to solve the movie scenes scheduling problem. They used an integer linear programming model and branch and bound algorithm to solve the problem. By preprocessing, double-ended search determines the lower bound, and cache search states and dominant rules enhance the branch and bound algorithm. The preprocessing is similar to that of de la Banda et al. [5]. Experiments results show that the improved branch and bound algorithm has better performance than the dynamic programming algorithm of de la Banda et al. [5]. Cheng et al. [9] presented two methods to improve the lower bound of the movie scenes scheduling problem. The result can give a partial timetable.

Most of studies consider the cost of the actor, nevertheless, a small number of studies, not only consider the cost of the actor, but also consider the operating costs of crew. Wang et al. [10] studied the minimization of movie costs, including the cost of the actors as well as the cost of operating the crew. They used the next fitting algorithm and the first fitting reduction algorithm to solve this problem.

Scheduling in movie shooting process is essentially a scheduling problem, which can draw lessons from the scheduling problem. There are also many studies that resemble the scheduling problem in movie. For example, the

problem of program rehearsal is optimized through reasonable scheduling, so as to reduce the waiting time and minimize the total cost. Therefore, we can learn from such research ideas and methods. Gregory et al. [11] solved the rehearsal scheduling problem by planning and model checking, setting the daily wage of actors differently, and the rehearsal duration of each program was determined. Sakulsom and Tharmmaphornphilas [12] proposed a two-stage method to solve the problem of rehearsal of music programs. They set the same actor daily wage, and the rehearsal duration of each program was determined.

From the above study, we can see that most scholars consider the cost of actors from the perspective of actors. They consider the cost of actors, including the daily wage of actor which is different, and each scene has the same shooting duration. As well as the daily wage of actor is different, each scene shooting duration for different values. Some scholars not only consider the cost of actors, but also consider the operational costs of the crew. Few scholars consider the transfer cost between scenes. Different movie scenes may be taken at different locations, so there is a high potential for transfer costs between the scene and the scene transition during movie scenes shooting process.

### 3. Model Formulation

Above all, the movie scenes scheduling problem is a scheduling problem. Zhen [13] studied the scheduling problem of tactical berth allocation problem. The first deterministic model established in the paper considers the certainty of the period. The research in this paper considers the certainty of the shooting duration. The model building process refers to the deterministic model in Zhen [13]. This paper establishes an integer programming model to solve the movie scenes scheduling problem considering the different wage of actors, different shooting duration of actors, and the transfer cost between scenes. The model mainly consists of two parts, one takes the cost of the actors into account, and the other considers the transfer cost between scenes.

Before formulating the model for the movie scenes scheduling problem, we list the notation used in the paper below.

#### 3.1. Notations

##### *Indices and Sets*

$i, j$ : Index of scenes.

$k$ : Index of an actor.

$m, n$ : Index of shooting locations.

$K$ : Set of all actors;  $k \in K$ .

$G$ : Set of all shooting locations;  $m, n \in G$ .

$S$ : Set of all scenes. Set two virtual scenes 0 and  $T$ , in which 0 indicates the scene where the crew begins to shoot and  $T$  indicates the scene the crew shot in the end.

$i, j \in S$ ;  $S^0 = S \cup \{0\}$ ;  $S^T = S \cup \{T\}$ ;  $\bar{S} = S \cup \{0, T\}$ .

## Parameters

$p_k$ : The daily wage of actor  $k$ .

$s_{ik}$ : Set to one if actor  $k$  appears in scene  $i$ , and zero otherwise.

$x_{ijk}$ : Set to one if actor  $k$  appears in scene  $i$  and scene  $j$ , and zero otherwise.

$c_{ij}$ : The cost of moving from scene  $i$  to scene  $j$ .

$u_i$ : The number of days required to shoot scene  $i$ .

## Decision Variables

$\delta_{ijk}$ : Binary variable; it equals one if the actor  $k$  first appears in scene  $i$  and then appears in scene  $j$ , and zero otherwise; note that it is not required that scene  $i$  and scene  $j$  are adjacent to each other.

$\alpha_{ijk}$ : Binary variable; it equals one if the actor  $k$  appears in two adjacent scenes  $i$  and  $j$  in which scene  $j$  is the next scene of scene  $i$ , and zero otherwise.

$\beta_{ijk}$ : Integer; the time interval of actor  $k$  plays between scene  $i$  and scene  $j$ . For the actor  $k$ , it first appeared in the scene  $i$ , followed by the scene  $j$ .

$\gamma_{ij}$ : Binary variable; it equals one if the scene  $j$  is the next scene of scenes  $i$ , and zero otherwise.

$\varepsilon_i$ : Integer; the time that scene  $i$  finishes.

## 3.2. Mathematical Model.

$$\text{Minimize } \sum_{i \in \bar{S}} \sum_{j \in \bar{S}} \sum_{k \in K} \beta_{ijk} p_k + \sum_{j \in S} \sum_{i \in S} \gamma_{ij} c_{ij} \quad (1)$$

$$\sum_{j \in S^T} \gamma_{0j} = 1 \quad (2)$$

$$\sum_{i \in S^0} \gamma_{iT} = 1 \quad (3)$$

$$\sum_{j \in S^T} \gamma_{ij} = 1 \quad \forall i \in S, i \neq j \quad (4)$$

$$\sum_{j \in S^0} \gamma_{ji} = 1 \quad \forall i \in S, i \neq j \quad (5)$$

$$\gamma_{ij} + \gamma_{ji} \leq 1 \quad \forall i \in \bar{S}, j \in \bar{S} \quad (6)$$

$$\varepsilon_j - \varepsilon_i - u_j \leq M(1 - \gamma_{ij}) \quad (7)$$

$$\forall i \in S^0, j \in S^T$$

$$\sum_{j \in S} \alpha_{0jk} = 1 \quad \forall k \in K \quad (8)$$

$$\sum_{j \in S} \alpha_{jTk} = 1 \quad \forall k \in K \quad (9)$$

$$\delta_{ijk} + \delta_{jik} = x_{ijk} \quad (10)$$

$$\forall i, j \in \bar{S}, i \neq j, \forall k \in K$$

$$\alpha_{ijk} \leq \delta_{ijk} \quad \forall k \in K, \forall i, j \in \bar{S} \quad (11)$$

$$s_{ik} \leq \sum_{j \in S^T} \alpha_{ijk} \quad \forall i \in S^0, \forall k \in K \quad (12)$$

$$s_{ik} \leq \sum_{j \in S^0} \alpha_{jik} \quad \forall i \in S^T, \forall k \in K \quad (13)$$

$$\varepsilon_j - \varepsilon_i + M(\alpha_{ijk} - 1) \leq \beta_{ijk} \quad (14)$$

$$\forall i, j \in S, i \neq j, \forall k \in K$$

$$u_j + M(\alpha_{0jk} - 1) \leq \beta_{0jk} \quad (15)$$

$$\forall j \in S, \forall k \in K$$

$$M(\alpha_{jTk} - 1) \leq \beta_{jTk} \quad \forall j \in S, \forall k \in K \quad (16)$$

$$s_{jk} - 0.5 \leq 0.5 \left( \sum_{i \in \bar{S}} \alpha_{ijk} + \sum_{i \in \bar{S}} \alpha_{jik} \right) \leq s_{jk} \quad (17)$$

$$\forall j \in \bar{S}, i \neq j, \forall k \in K$$

$$M(\delta_{ijk} - 1) \leq \varepsilon_j - \varepsilon_i \quad (18)$$

$$\forall i \in S^0, \forall j \in S^T, i \neq j, \forall k \in K$$

$$\varepsilon_j - u_j \leq M(2 - \alpha_{0jk} - \gamma_{ij}) + \varepsilon_i \quad (19)$$

$$\forall i \in S^0, \forall j \in S^T, \forall k \in K$$

$$\varepsilon_j - u_j \leq M(2 - \alpha_{jTk} - \gamma_{ij}) + \varepsilon_i \quad (20)$$

$$\forall i \in S^0, \forall j \in S^T, \forall k \in K$$

$$0 \leq \varepsilon_j - \varepsilon_i - u_j + M(1 - \alpha_{ijk}) \quad (21)$$

$$\forall i \in S^0, \forall j \in S^T, \forall k \in K$$

$$0 \leq \varepsilon_i \leq d \quad \forall i \in \bar{S} \quad (22)$$

$$\beta_{ijk} \geq 0 \quad \forall i, j \in \bar{S}, \forall k \in K \quad (23)$$

$$\delta_{ijk}, \alpha_{ijk}, \gamma_{ij} \in \{0, 1\} \quad \forall i, j \in \bar{S}, \forall k \in K \quad (24)$$

Objective function (1) minimizes the total costs of the movie scenes scheduling problem. The total costs include the wage of actors and the transfer cost of scenes. Constraints (2) and (3) guarantee that each scene has a start point 0 and an end point  $T$ . Constraints (4) and (5) ensure that there is only one scene before and after each scene. Constraint (6) states the priority relationship of scene  $i$  and scene  $j$ . Constraint (7) links the two decision variables  $\gamma_{ij}$  and  $\varepsilon_i$ . Constraints (8) and (9) ensure that each actor  $k$  starts with scene 0 and ends with scene  $T$ . Constraints (10) and (11) define the relation of  $\alpha_{ijk}$ ,  $x_{ijk}$ , and  $\delta_{ijk}$ . Constraints (12) and (13) guarantee that for actor  $k$ , each scene before and after there is only one scene. Constraints (14), (15), and (16) define  $\beta_{ijk}$ . Constraint (17) ensures that actor  $k$  completes the shooting scenes according

to the sequence decided by  $\alpha_{ijk}$ . Constraint (18) links the two decision variables  $\delta_{ijk}$  and  $\varepsilon_i$ . Constraints (19), (20), and (21) link the two decision variables  $\alpha_{ijk}$  and  $\varepsilon_i$ . Constraints (22), (23), and (24) define the decision variable.

The above model can be solved by ILOG CPLEX solver. However, ILOG CPLEX solver can only solve small-scale movie scenes scheduling problems. For some large-scale movie scenes scheduling problems, the ILOG CPLEX solver cannot solve very well. Therefore, this paper proposes a TSBM to handle the movie scenes scheduling problems under large-scale situation.

## 4. Tabu Search Based Method

**4.1. Introduction to the Tabu Search.** At present, many scholars have used tabu search (TS) algorithms to solve optimization problems. Xue et al. [15] studied the problem of local container shipping under the new operating mode, defining that the tractor can be removed from its associated trailer and assigned to a new mission. Aiming at this problem, the TS is proposed, which uses numerical experiments to evaluate the performance of the algorithm and the advantages of the new operation mode. Shen [16] designed a TS with complex neighborhood structure for job shop problems with sequence-related setup time. The method of inserting is used to define the neighborhood function without a simple exchange. The test results show that the proposed TS is superior to other algorithms. Qin et al. [17] studied a multicycle inspector scheduling problem, which is an extension of the multistroke vehicle routing problem with time windows, and then solved using the TS. In addition, a constrained backpack model is proposed, which can generate an upper limit of the problem, thereby improving the efficiency of the solution. Peng et al. [18] combined the TS process into the path relinking framework to study a new algorithm to solve the job shop scheduling problem. The experimental results show that the new algorithm has a good effect on the solution quality and computational efficiency. Li and Alidaee [19] use the TS to solve the black and white traveling salesman problem. The improved TS is built on a new efficient neighborhood structure, which utilizes the arrangement of black and white traveling salesman problems and the backpack function. Moreover, the TS is embedded as a heuristic process to improve the upper bound of the mixed integer linear programming equation. Li and Gao [20] use a hybrid algorithm of genetic algorithm and TS to solve the shop scheduling problem, in order to minimize the completion time. The genetic algorithm with powerful global search ability is used to explore the TS with good local search ability. The experimental results show that the hybrid algorithm has good search ability and can balance the intensification and diversification. Shahvari and Logendran [21] improved the TS for a batch scheduling problem on a set of uncorrelated parallel machines with the goal of minimizing the linear combination of total weighted completion time and total weighted delay. Therefore, we design a TSBM to solve the integer linear programming model. Wei et al. [22] studied the partner selection problem from the perspective of supplier network global optimization. The study established

an optimal combination model of supplier selection to choose the right suppliers for improving the efficiency of the supply chain. A dynamic programming-TS and an ant colony algorithm were developed to solve the model, respectively.

TS is a global progressive optimization algorithm. Factors affecting search performance in TS include initial solution, special rules, neighborhood structure, tabu list, candidate set, and termination rule. For the amnesty rules, if the target value of a solution is better than any of the best candidate solutions, it can be special. And if all objects are contraindicated, then a solution with the smallest evaluation value is selected. It is possible to specifically target objects that have a large impact on the target value. For the size of the candidate set, the excessive calculation increases the computational memory and computation time and too small will cause the premature fall into the local optimum. For termination criteria, it is to make the search terminate the search process in some cases. For the tabu list, it has a short-term memory function that prohibits searching for previously searched content. For the tabu length, the purpose is to prevent the search from appearing in an infinite loop. When other variables are unchanged, the shorter the tabu length, the less memory usage, and the larger the scope of the lifting, that is, the upper limit of the search range. But it is easy to fall into local optimum earlier. If the length of the tabu is too long, it will be time-consuming.

### 4.2. The Frame of Tabu Search

**4.2.1. The Basic Flow of the TSBM.** The basic flow of the TSBM is as follows.

*Step 1.* Determine an initial feasible solution and its objective function value. Set the tabu list to empty and set the length of the tabu list. Meanwhile, let the initial solution as the current solution  $x_{now}$ . In addition, define  $n$  as the number of iterations. Set  $N = 1$ ,  $nowL = 0$ .

*Step 2.* When  $n$  reaches the maximum number of iterations  $N$ , end the algorithm and output the optimal solution; otherwise, go to Step 3.

*Step 3.* If  $nowL < L$ , the neighborhood function that generates the current solution  $x_{now}$  will generate a number of neighborhood solutions, and go to Step 5; otherwise, go to Step 4;

*Step 4.* Use the diversification strategy to generate a new current solution  $x_{now}$ , and set  $nowL = 0$ , go to Step 3.

*Step 5.* Sort the candidate solutions according to the fitness value, and judge from the candidate solution with the smallest fitness value. If it is not in the tabu list, directly replace it with the current solution to become  $x_{now}$ ; if it is in the tabu list and meet the exemption criteria, let it as  $x_{now}$  and go to Step 7. Otherwise, go to Step 6.

*Step 6.* If the candidate solution value is in the tabu list and does not meet the exemption criteria, certain courage criterion is given. If the candidate solution value satisfies the

courage criteria, it is also selected as the current solution  $x_{now}$ . Go to Step 7.

*Step 7.* Update the global optimal  $x_{best}$ , and put the current solution into the tabu list. If the tabu list exceeds the set length, release the tabu element that firstly put into the tabu list. When  $x_{best}$  has not changed,  $nowL = nowL + 1$ ,  $n = n + 1$ , go to Step 2.

*4.2.2. Initial Solution.* The initial solution has great influence on the algorithm search performance because the TSBM is mainly based on the neighborhood search. Therefore, we firstly consider generating a number of feasible solutions for CPLEX solver randomly and select the optimal fitness value of the solution as the initial solution.

Each  $x$  represents one possible order of scene, i.e., the variable  $\gamma_{ij}$  is known. If  $x$  is given, then we select the optimal fitness value of the solution as the initial solution. For example, if there are 5 scenes, elements in  $x$  are 1-4-5-3-2, which means the order of scene is 1-4-5-3-2.

*4.2.3. Neighborhood Structure.* The neighborhood structure is critical in tabu search because it determines the form and number of neighborhood solutions of the current solution and the relationship between the solutions.

The neighborhood is generated by 2-swap and 2-opt methods. The 2-swap method is to generate neighborhood as explained below. On the basis of the current solution  $x_{now}$ , the order of the two scenes is randomly exchanged. For example,  $x_{now}$  is 1-4-5-3-2, and we exchange first and third scene to get a neighborhood, i.e., 5-4-1-3-2. However, the method of generating the neighborhood using the 2-opt method is random two scenes for 2-opt exchange. And 2-opt means that from the selected starting scene to the ending scene, all scenes are reversed to generate new neighborhoods. For example,  $x_{now}$  is 1-4-5-3-2, and the two random scenes are the second and fifth scenes, so a new neighborhood is 1-2-3-5-4.

*4.2.4. Tabu List.* The tabu list has a short-term memory function which prohibits the items that are searched before. TSBM uses tabu list to prevent the search process being caught in the local optimal.

The neighborhood exchange method and the exchanges of two points will be added to the tabu list. For example,  $x_{now}$  is 1-4-5-3-2, and we exchange the first and the third scene to get a neighborhood, i.e., 5-4-1-3-2. Use the 2-swap method, and exchange the first and the third scene, then add 1-1-3 three numbers to the tabu list. Besides, set  $x_{now}$  is 1-4-5-3-2, and two random scenes are the second and the fifth scenes, so the new neighborhood is 1-2-3-5-4. And we use the 2-opt method, and the random number is 2 and 5, then add 2-2-5 to the tabu list. This shows that when determining whether a candidate solution is in the tabu list, you need to know which method is used, and which two points should be random.

*4.2.5. Diversification Strategy.* In order to avoid the algorithm getting into local optimum prematurely, a diversified strategy is adopted in this paper. If the optimal value does not change after a certain number of iterations, we will adopt a diversified

strategy to regenerate the current solution and make the algorithm jump out of the local minimum. In addition, the method of generating the current solution is the same as that of generating the initial solution, i.e., it is generated randomly.

## 5. Particle Swarm Optimization Based Method

*5.1. Introduction to the PSO.* PSO is a new evolutionary algorithm proposed by Kennedy and Eberhart [23] to optimize continuous nonlinear functions. At present, PSO has been widely used to solve problems such as scheduling, combinatorial optimization, and supply chain network design. Garg and Sharma [24] use PSO to solve system assignment problems. Many of these goals are system reliability and design costs. Above all, a fuzzy nonlinear programming model is constructed, and then the particle swarm optimization algorithm is designed to solve the proposed problem. Hamta et al. [25] studied the assembly line balance problem of a single model. This considers the minimization of cycle time, total equipment cost, and smoothness index. A new method combining PSO and variable domain search is used to solve the proposed problem. Zhong et al. [26] use the PSO to solve the multiobjective hull assembly line balance problem. The goal is to minimize cycle time, multisite responsibility, and dynamic and static load balancing. In the particle swarm algorithm, the new method is used to encode the particles. Yao et al. [27] considered the carton heterogeneous vehicle routing problem at the collection station, constructed a mathematical model, and solved the model using the improved PSO. For improving the PSO, the local search is optimized in the algorithm. Experiments show that both the proposed model and the algorithm are feasible. Khatami et al. [28] constructed a PSO for flame detection, which enabled image detection techniques for flame detection. PSO can distinguish the different colors, and the color conversion matrix can detect different color flames. Hu et al. [14] established a mixed integer programming model for the joint optimization of multibatch production and vehicle routing problems involving a pickup. A two-stage hybrid heuristic algorithm is proposed to solve this model. The PSO was used in the second stage to allocate vehicles to each production batch. The results of numerical experiments validate the efficiency of the proposed algorithm.

PSO finds the best solution by simulating the process of bird flocking and sharing information between birds and birds. Each particle can change its position and speed, and the position of the particle is a viable solution in the search space. The speed of the particles determines the direction and distance of the flight of bird. Other particles will search for the direction of the current optimal particle. Each particle remembers, follows the current optimal particle, and searches in the solution space. But the process of each iteration is not completely random. If a better solution is found, it will be used as a basis to find the next solution to find the global optimal solution. The advantage of PSO is that the algorithm is easy to implement and there are not many parameters to adjust.

TABLE 1: The performances of the proposed solution methods.

Case id	TSBM		PSOBM		CPLEX		Gap	Gap1
	$OBJ_T$	$T_T$ (s)	$OBJ_P$	$T_P$ (s)	$OBJ_C$	$T_C$ (s)		
10-3-3-1	2960	14	3040	137	2960	8	0.00%	2.70%
10-3-3-2	3440	21	3980	154	3440	13	0.00%	15.70%
10-3-3-3	4440	7	4630	125	4440	22	0.00%	4.28%
11-3-3-1	4600	42	5010	178	4600	97	0.00%	8.91%
11-3-3-2	4810	7	5130	156	4810	104	0.00%	6.65%
11-3-3-3	4410	43	4630	147	4410	115	0.00%	4.99%
11-5-3-1	5700	14	5860	215	5690	48	0.20%	2.99%
11-5-3-2	6890	46	6930	246	6890	165	0.00%	0.58%
11-5-3-3	6710	21	7060	197	6710	170	0.40%	5.22%
13-5-4-1	22230	57	26430	236	22130	1167	0.00%	19.43%
13-5-4-2	32160	46	33790	318	31800	2850	1.00%	6.26%
13-5-4-3	20180	26	23570	115	20080	6613	0.50%	17.38%
15-5-4-1	36680	51	38800	53	N.A	>7200	N.A	N.A
15-5-4-2	22320	74	31300	39	N.A	>7200	N.A	N.A
15-5-4-3	32800	113	38430	137	N.A	>7200	N.A	N.A
					Avg.		0.18%	7.92%

5.2. *The Frame of Particle Swarm Optimization.* The flow of the PSOBM is as follows.

*Step 1.* Set the initial iteration number  $G = 1$  and the number of particles to  $H$ . The position of the particle represents a feasible solution, and the velocity of the particle represents the search direction. The position and velocity of the initial particle swarm are randomly generated.

*Step 2.* Calculate the fitness value of each particle.

*Step 3.* Find the individual optimal and global optimal values of each particle. If it is better than the current individual extremum, it is updated; if the best of all the individual extremum of the particles is better than the current global optimal, the global optimal is updated.

*Step 4.* Use the particle velocity and position formula to update the velocity and position of the particle.

*Step 5.* If one of the two termination conditions is reached, the iteration is stopped and the optimal solution is output; otherwise, go to Step 2.

Among the above steps, the two termination conditions in Step 5 are as follows. For the first, the number of iterations reaches the maximum number of iterations. And the global optimal value is no longer updated in the number of iterations.

## 6. Numerical Experiments

6.1. *Experiment Settings.* In this section, we apply the proposed model and methods to solve the movie scenes scheduling problem. The movie scenes scheduling problem is a scheduling problem. The numerical experiments in this paper

mainly refer to the experiments of Zhen [29] and Zhen et al. [30] to analyze the proposed method and model.

The experimental cases in this paper are set as follows. For the cost  $c_{ij}$  transferred from scene  $i$  to scene  $j$ , they are randomly generated within the range (0, 10000). Moreover, it is necessary to satisfy  $c_{ij} = c_{ji}$ . For the parameter  $q_k$  (i.e., the daily wage of actor  $k$ ), they are randomly generated within a range between 80 dollars/day and 1000 dollars/day. For the parameter  $s_{ik}$  (i.e., set to one if the actor  $k$  appears in scene  $i$ ) and  $x_{ijk}$  (i.e., set to one if the actor  $k$  appears in scene  $i$ , scene  $j$ ), they are mainly generated based on the actor  $k$ . Numerical experiments are performed on a PC (Intel (R), 3.4 GHz; Memory, 128 G). The models are implemented by the CPLEX12.6, and the methods are coded by C# (Visual Studio 2015).

6.2. *Performances of the Proposed Solution Methods.* This paper designs a TSBM and PSOBM to solve the proposed integer linear programming model. Compare the performance of the proposed solution with the results obtained by directly using the CPLEX solver. Table 1 shows the comparison of the experimental results of TSBM, PSOBM, and CPLEX solvers under the same parameters. In the ‘‘Case id’’ column of Table 1, the first three values represent the scene, the actor, and the location, and the fourth value represents the index of the different cases at the same scale (scene, actor, and location). The difference between the three cases in each group size is mainly due to the difference in the length of the scene. In this paper,  $OBJ_T$ ,  $OBJ_P$ , and  $OBJ_C$  show the objective value of TSBM, PSOBM, and CPLEX. And  $Gap = (OBJ_T - OBJ_C)/OBJ_C$ ,  $Gap1 = (OBJ_P - OBJ_C)/OBJ_C$ .

It can be seen from Table 1 that under the small-scale problems, the TSBM has a shorter solution time than CPLEX, and the average difference rate is 0.18%. The PSOBM has a longer solution time, and the average difference rate is

TABLE 2: Comparison between TSBM and PSOBM.

Case id	TSBM		PSOBM		Gap2
S-A-P	$OBJ_T$	$T_T$ (s)	$OBJ_P$	$T_P$ (s)	
15-5-4-1	36680	51	38800	53	5.78%
15-5-4-2	22320	74	31300	39	40.23%
15-5-4-3	32800	113	38430	137	17.16%
20-5-6-1	30320	122	46330	102	52.80%
20-5-6-2	35260	127	53840	81	52.69%
20-5-6-3	32090	214	47350	163	47.55%
30-10-10-1	227730	677	341930	813	50.15%
30-10-10-2	225000	616	380560	898	69.14%
30-10-10-3	213140	654	241730	946	13.41%
30-20-10-1	468200	1295	572380	1547	22.25%
30-20-10-2	439530	1240	668980	1435	52.20%
30-20-10-3	515170	1293	594190	1636	15.34%
40-20-15-1	897070	2176	1092090	2807	21.74%
40-20-15-2	692040	2112	984500	2749	42.26%
40-20-15-3	763560	2034	978900	3006	28.20%
				Avg.	35.39%

7.92%. Therefore, in the small-scale movie scene scheduling problem, the TSBM can obtain good results in a shorter time than the PSOBM. Based on the above analysis, it can be concluded that for small-scale problems, the proposed TSBM is a better way to obtain a satisfactory solution.

**6.3. Comparison between TSBM and PSOBM.** For small-scale problems, CPLEX, TSBM, and PSOBM can obtain good results. But the CPLEX takes a long time. In addition, the result of TSBM and PSOBM takes a short time. Therefore, for large-scale problems, the validity of the TSBM and PSOBM needs to be further studied experimentally. In this test, the result of the comparison between TSBM and PSOBM is listed in Table 2. And  $Gap2 = (OBJ_P - OBJ_T)/OBJ_T$ .

The results in Table 2 show that the average difference between the experimental results obtained by the TSBM and the PSOBM is 35.39%. And as the size of the experimental case increases, the time required for problem solving continues to increase. Experimental results show that TSBM is more suitable for solving medium-scale movie scenes scheduling problem.

## 7. Conclusions

This paper mainly studies the optimization problem of movie scenes scheduling by considering transfer costs and certain shooting duration. The purpose is to minimize the total cost of actors and transfer costs included in the shooting scene of the crew, considering the different wage of the actors, the duration of the scene shooting, and the transfer costs between scenes. An integer linear programming model is built, and TSBM and PSOBM are developed to solve the proposed model. Being compared with other related research work, the major contribution of this study includes two aspects:

(1) Being different from the general movie scenes scheduling problem, the proposed model in this paper considers the transfer costs and certain shooting duration.

(2) We propose different methods of TSBM and PSOBM for the movie scenes scheduling problem with different scale.

The numerical experiments indicate both TSBM and PSOBM are suitable for solving small-scale problems. For large-scale problems, the TSBM can get better results.

For future study, we will consider finding a better optimization method in TSBM to extend our work.

## Data Availability

The data used in this research is generated randomly. Detailed data generation method is introduced in Section 6.1.

## Conflicts of Interest

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

## References

- [1] T. C. Cheng, J. E. Diamond, and B. M. Lin, "Optimal scheduling in film production to minimize talent hold cost," *Journal of Optimization Theory and Applications*, vol. 79, no. 3, pp. 479–492, 1993.
- [2] A.-L. Nordström and S. Tufekci, "A genetic algorithm for the talent scheduling problem," *Computers & Operations Research*, vol. 21, no. 8, pp. 927–940, 1994.
- [3] A. Fink and S. Voß, "Applications of modern heuristic search methods to pattern sequencing problems," *Computers & Operations Research*, vol. 26, no. 1, pp. 17–34, 1999.
- [4] F. Bomsdorf and U. Derigs, "A model, heuristic procedure and decision support system for solving the movie shoot scheduling problem," *OR Spectrum*, vol. 30, no. 4, pp. 751–772, 2008.

- [5] M. Garcia de la Banda, P. J. Stuckey, and G. Chu, "Solving talent scheduling with dynamic programming," *INFORMS Journal on Computing*, vol. 23, no. 1, pp. 120–137, 2011.
- [6] Y. Kochetov, "Iterative local search methods for the movie scene scheduling problem," in *Proceedings of the 1st International Symposium and Balkan Conference on Operational Research*, pp. 282–288, Thessaloniki, Greece, September 2011.
- [7] X. Liang, Z. Zhang, H. Qin, S. Guo, and A. Lim, "A Branch-and-Bound Algorithm for the movie scene scheduling problem," in *Proceedings of the 27th International Conference on Modern Advances in Applied Intelligence, Part I*, vol. 8481, pp. 208–217, Springer, New York, NY, USA, 2014.
- [8] H. Qin, Z. Zhang, A. Lim, and X. Liang, "An enhanced branch-and-bound algorithm for the talent scheduling problem," *European Journal of Operational Research*, vol. 250, no. 2, pp. 412–426, 2016.
- [9] T. C. Cheng, B. M. Lin, and H.-L. Huang, "Talent hold cost minimization in film production," *Journal of Industrial and Management Optimization*, vol. 13, no. 1, pp. 223–235, 2017.
- [10] S.-Y. Wang, Y.-T. Chuang, and B. M. T. Lin, "Minimizing talent cost and operating cost in film production," *Journal of Industrial and Production Engineering*, vol. 33, no. 1, pp. 17–31, 2016.
- [11] P. Gregory, A. Miller, and P. Prosser, "Solving the rehearsal problem with planning and with model checking," *College of Science and Engineering School of Computing Science*, vol. 103, no. 3, pp. 441–446, 2004.
- [12] N. Sakulsom and W. Tharmmaphornphilas, "Scheduling a music rehearsal problem with unequal music piece length," *Computers & Industrial Engineering*, vol. 70, no. 1, pp. 20–30, 2014.
- [13] L. Zhen, "Tactical berth allocation under uncertainty," *European Journal of Operational Research*, vol. 247, no. 3, pp. 928–944, 2015.
- [14] H. Hu, J. Mo, and C. Ma, "Optimization of vehicle routing with pickup based on multibatch production," *Discrete Dynamics in Nature and Society*, vol. 2018, Article ID 2804589, 9 pages, 2018.
- [15] Z. Xue, C. Zhang, W.-H. Lin, L. Miao, and P. Yang, "A tabu search heuristic for the local container drayage problem under a new operation mode," *Transportation Research Part E*, vol. 62, no. 1, pp. 136–150, 2014.
- [16] L. Shen, "A tabu search algorithm for the job shop problem with sequence dependent setup times," *Computers & Industrial Engineering*, vol. 78, pp. 95–106, 2014.
- [17] H. Qin, W. Ming, Z. Zhang, Y. Xie, and A. Lim, "A tabu search algorithm for the multi-period inspector scheduling problem," *Computers & Operations Research*, vol. 59, pp. 78–93, 2015.
- [18] B. Peng, Z. Lü, and T. C. E. Cheng, "A tabu search/path relinking algorithm to solve the job shop scheduling problem," *Computers & Operations Research*, vol. 53, no. 53, pp. 154–164, 2015.
- [19] H. Li and B. Alidaee, "Tabu search for solving the black-and-white travelling salesman problem," *Journal of the Operational Research Society*, vol. 67, no. 8, pp. 1061–1079, 2016.
- [20] X. Li and L. Gao, "An effective hybrid genetic algorithm and tabu search for flexible job shop scheduling problem," *International Journal of Production Economics*, vol. 174, pp. 93–110, 2016.
- [21] O. Shahvari and R. Logendran, "An enhanced tabu search algorithm to minimize a bi-criteria objective in batching and scheduling problems on unrelated-parallel machines with desired lower bounds on batch sizes," *Computers & Operations Research*, vol. 77, pp. 154–176, 2017.
- [22] C.-T. Wei, H. Zuo, C.-B. Jiang, and S.-f. Li, "Modeling multilevel supplier selection problem based on weighted-directed network and its solution," *Discrete Dynamics in Nature and Society*, vol. 2017, Article ID 8470147, 12 pages, 2017.
- [23] J. Kennedy and R. Eberhar, "Particle swarm optimization," in *Proceedings of the IEEE International Conference*, pp. 1942–1948, IEEE, 1995.
- [24] H. Garg and S. P. Sharma, "Multi-objective reliability-redundancy allocation problem using particle swarm optimization," *Computers & Industrial Engineering*, vol. 64, no. 1, pp. 247–255, 2013.
- [25] N. Hamta, S. M. T. Fatemi Ghomi, F. Jolai, and M. Akbarpour Shirazi, "A hybrid PSO algorithm for a multi-objective assembly line balancing problem with flexible operation times, sequence-dependent setup times and learning effect," *International Journal of Production Economics*, vol. 141, no. 1, pp. 99–111, 2013.
- [26] Z. Yuguang, A. Bo, and Z. Yong, "A PSO algorithm for multi-objective hull assembly line balancing using the stratified optimization strategy," *Computers & Industrial Engineering*, vol. 98, pp. 53–62, 2016.
- [27] B. Yao, B. Yu, and P. Hu, "An improved particle swarm optimization for carton heterogeneous vehicle routing problem with a collection depot," *Annals of Operations Research*, vol. 242, no. 2, pp. 303–320, 2016.
- [28] A. Khatami, S. Mirghasemi, A. Khosravi, C. P. Lim, and S. Nahavandi, "A new PSO-based approach to fire flame detection using K-Medoids clustering," *Expert Systems with Applications*, vol. 68, pp. 69–80, 2017.
- [29] 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.
- [30] L. Zhen, E. P. Chew, and L. H. Lee, "An integrated model for berth template and yard template planning in transshipment hubs," *Transportation Science*, vol. 45, no. 4, pp. 483–504, 2011.

## Research Article

# Simulation Optimization of Discrete Logistics Processes: A Case Study on Logistics of an E-Commerce Enterprise in Shanghai

Xiang Xu , Yue Li, and Ruiyun Tang 

*School of Management, Shanghai University, Shang Da Road 99, Shanghai 200444, China*

Correspondence should be addressed to Ruiyun Tang; [try19970814@163.com](mailto:try19970814@163.com)

Received 23 November 2018; Revised 29 December 2018; Accepted 3 January 2019; Published 20 January 2019

Guest Editor: Xinchang Wang

Copyright © 2019 Xiang Xu 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 rapid development of online shopping in recent years, logistics distribution has received much attention from enterprises and online consumers. Logistics distribution involves many factors and complex processes; conventional qualitative methods are unable to provide an effective analysis. Thus, this paper sets a framework to solve the above problem. A case study of an E-commerce enterprise in Shanghai on logistics distribution is proposed to discretize the whole process and minimize the total costs. Then the AnyLogic software is used to simulate and optimize the system from three aspects, including routes selection, warehouses quantity, and warehouses layout. Finally, this paper analyzes the simulation results, which would provide some valuable references for practical logistics.

## 1. Introduction

In recent years, E-commerce industry has developed rapidly. Online shopping is almost necessary to everyone. Logistics distribution is the last link of online shopping whose importance is rising as society demand increases. Whether goods can be delivered to consumers in time affects the consumers satisfaction of this shopping directly. Especially in the annual “Double Eleven” shopping festival, parcel quantities in various regions have increased rapidly and many delivery points have exploded. As shown in Figure 1, the logistics orders of Double Eleven in 2018 have set a new record. Logistics industry has also transformed into the rapid development stage. It is estimated that distribution costs account for over 50% of the total operational costs. This problem needs to be solved urgently. Besides, logistics distribution as one of the three major contents (distribution, storage, and management) in the logistics field includes scheduling management, distribution tools, distribution routes, delivery time, natural environment, human resources, and so on. More are getting to the importance of logistics distribution. Therefore, in this paper, much attention is paid to optimize and analyze the logistics distribution, shorten the delivery time, improve the distribution efficiency, and reduce the distribution costs. The

distribution tools are necessary in the whole process which occupy considerable resources under demands uncertainty. To lower the logistics costs dramatically, reducing this consumption is significant through optimization [1]. Hence this paper considers the vehicle as the breakthrough point to analyze the costs optimization in system.

Meanwhile, the research on vehicle is not a theoretical problem in logistics distribution process. Lots of factors are contained including vehicles quantity, distribution terminal, delivery time, unloading time, and demand changes. Conventional qualitative methods are not insufficient to solve it. In recent years, computational technology including hardware and software has developed rapidly. This technology characterizes reflecting on complicated processes or behaviors to solve problems through simulation. Simulation is a new subject that has gradually formed with the development of computer technology. It was firstly proposed in the early 20th century and was mainly utilized in water conservancy research. Simulation is the process of experimental research on the system by establishing and using the real system model. Similar to the application of algorithm on theoretical issues, simulation has significant effects on practical problems, particularly the complex and practical problems like logistics distribution. By discretizing and dividing the whole

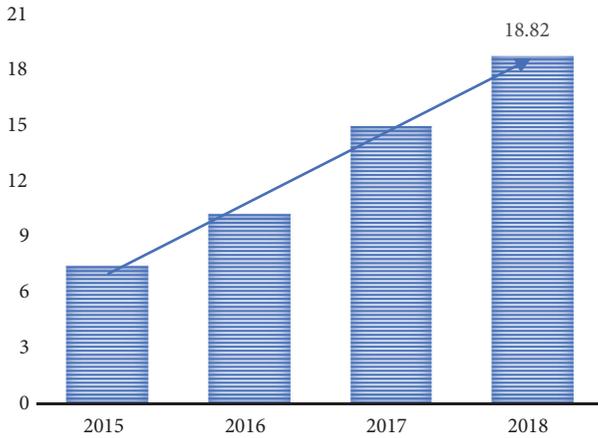


FIGURE 1: Parcel quantities during the “Double Eleven”.

process into different parts, an integrated model is established and analyzed for every part to obtain the system data. According to the obtained data, optimal results are calculated [2]. The methods have more practical value than the theoretical algorithm. Better simulation methods have been presented in recent years as the computer technology develops. Simulation on the application of logistics distribution will have broader prospects.

This paper studies the discrete logistics processes which include many stochastic variables and factors. The method of mathematical modeling is not suitable. Therefore, the simulation is used to optimize the logistics distribution system and get the practical results.

For the simulation research of logistics distribution problems, the GPSS language (The General Purpose Systems Simulator) was firstly presented by American Geoffrey Gordon in 1961, which is a solution to discrete events, particularly the queuing phenomenon [3]. Considering the combination of continuous system and discrete system, simulation language of hybrid system occurred after the 1970s, for instance, SLAM language (Simulation Language for Alternative Modeling), which translates the program into FORTRAN language and then compiles it into machine codes with high execution rate. The language can be used in discrete systems, continuous systems, and hybrid systems consisting of both [4]. KV has established an interactive model to support the logistics planning on container operations, which can provide the evaluation of the ports performance, optimize the utilization rate of ports, and shorten the turnaround time of ships [5]. Ila et al. proposed an eight-step simulation model development process (SMDP) to the design, implementation, and evaluation of logistics and supply chain study by adopting discrete events [6]. Meng et al. modeled the problem of free delivery in E-commerce companies with uncertain demands and discussed the influence of uncertain factors on the optimal solution [7]. Geng et al. established a self-organized elastic supply chain model based on MAS and designed the local fitness function, neighbor structure, and community interaction rules with the enterprises as agent. The results indicate that the system has an aggregation effect

and its evolution can be controlled by MAS parameters [8]. Cui et al. transformed one step of RUL estimate of simulation model in discrete events logistics system into two steps. An instance validated the effectiveness and testified the performance of the two-step RUL estimation which is better than the one-step estimation [9]. Li et al. established a nonlinear model for vehicle routing planning problems subject to time windows which considered the influence of road irregularities on fresh fruits and vegetables. Compared to the results with the conventional models, the new model is superior to the prior one [10]. Lin et al. constructed a method of using transit signals first in the case of emergency evacuation among a sudden disaster without a specialized-transit channel. The bus signal priority dominated by four factors which could evacuate personnel and lessen time as much as possible. Paramics is utilized to simulate this method which is better than the other methods [11]. Thies et al. studied the effects of resources sharing on potential savings in logistics industry through a model of discrete instances and the resource utilization in installation stage of offshore wind energy generator based on Agent. The simulation results show that weather has a significant influence on the installation time, use time, and resources utilization rate. Meanwhile, the resources sharing has a potential saving on the installation of offshore wind energy generator [12]. Li et al. adopted the method of dynamic traffic network analysis and discussed the optimization of regional traffic organization. Then they simulated it with self-developed software and proposed the optimized model and algorithm to prove the method on availability and feasibility [13]. Wei et al. identified that, in a discrete-time VMI supply chain system composing of one retailer and one manufacturer, production fluctuations can be interestingly stable even if the retailer subsystem is fluctuant. Simulation experiments are used to verify the theoretical results on inventory and production fluctuation [14]. Teodor Gabriel et al. proposed an innovative classification method for the use of simulation in a complicated system of multimodal transportation with multi-participant. This method identifies the main findings, trends, and future routes in multi-dimension of multimodal transportation [15]. Mandi et al. presented a grouping formula based on the branch pricing to study the opportunity-constrained vehicle routing problem with stochastic demands and enhanced the solution quality by simulation experiments and sensitivity analysis [16]. Hu et al. proposed an urban traffic model (AUTM) for predicting and avoiding the traffic congestion. This model is used to the simulation of large-scale practical cases in different cities under different congestion conditions and has satisfactory results [17]. In response to the increasingly complicated logistics systems studies on discrete or continuous process, the number of simulation software arises with the rapid development of computer technology in related fields [18], which has developed a tendency to visualization, modularity, and intelligence. Modeling speed is accelerated and simulation effect is improved through prefabrication of various components. Many simulation platforms in this area are commonly seen on the market, for instance, Areal, Witness, Flexsim, Promodel, Automod, and so on.

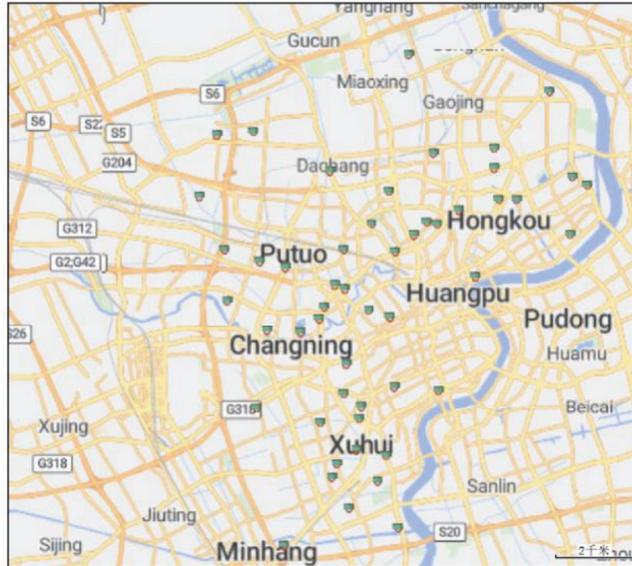


FIGURE 2: Distribution of self-raising points in main urban area of Puxi.

This paper considers adopting a simulation tool, AnyLogic software, which is developed early in this century to visualize modeling with a wide application scope. Complicated logistics distribution problems are discretized and simulated from the perspective of different processes. Moreover, one E-commerce enterprise in Shanghai is studied as a case. This paper starts with continuous changes of warehouse quantities and demands and optimizes the route selection, warehouses quantity, and warehouses layout. Ultimately, some optimization suggestions are raised based on simulation results of the software.

## 2. Application of AnyLogic Simulation Platform in Logistics Distribution Field

AnyLogic, a commercial simulation software released in 2000 by the AnyLogic Company, is a powerful simulation platform which can be applied in a wide range of fields, including logistics simulation, supply chain simulation, virus pervasion, road traffic, pedestrian evacuation, military simulation, and so on. This platform can also be used in discrete events modeling, agent-based modeling, and dynamics system modeling. This paper combines the AnyLogic technology with a case of an E-commerce enterprise in Shanghai on logistics distribution to propose the optimization suggestions.

The detailed introduction on the distribution case of an E-commerce enterprise in Shanghai is shown as follows.

E-commerce enterprises usually distribute goods to customers in two steps: (1) deliver goods from large-scale warehouses to distribution stations; (2) deliver goods from distribution station to customers with numerous manpower and material resources. Thus, two-part costs occur. The second costs are much more than the first one due to the large number of involving personnel. Moreover, higher risks of traffic accidents and loss of goods are generated. Hence,

to save costs and reduce risks, the E-commerce company in Shanghai has established commodity self-raising points in various regions. The company only dispatches vehicles transferring goods to the distribution points and customers pick up the goods themselves so that the costs of the second part can be completely saved. At the same time, the injury of the delivery personnel and the loss of the goods are dramatically reduced.

According to statistics, 51 self-raising points have been established in the main urban area of Puxi by the E-commerce company. The distribution map is obtained through the AnyLogic platform as follows in Figure 2.

Firstly, one large-scale warehouse is considered to construct in Northwest of Shanghai outside the main urban area, which is responsible for goods distribution to self-raising points in main urban area of Puxi. The location is shown in Figure 3.

The AnyLogic simulation is used to model and calculate the cost and the time requirement of completing 51 self-raising points on vehicles, which are salient criterion for assessing the solution.

Many practical factors need to be considered before starting the simulation. For instance, the time requirements for delivery beyond the limits are causing the compensation to the customer called tardiness cost which increases the total costs. Besides, the number of goods from every self-raising point affects the total delivery time and some time-sensitive delivery requirements for subsequent points.

Because of the different situations every day, the above two factors of every self-raising point can generate the orders quantity and the delivery time requirement through a random function as the fundamentals of the simulation. Thus, the following four items are included in every self-raising point: (1) name; (2) location; (3) goods quantity; (4) time limits.



FIGURE 3: Single warehouse location.

While large-scale warehouses only involve delivery, only the following two items need to be considered: (1) name; (2) location.

Two solutions are considered in this paper for route planning: (1) the shortest route solution, which calculates the shortest route between two locations as the real path for vehicles travelling; (2) the time-limited precedence solution, which considers firstly to deliver the goods with time-sensitive requirements and then the shortest route solution is adopted. Solution one can reduce the vehicles travelling costs and increase the tardiness costs while solution two is completely the opposite. Ultimately, the simulation results are used to compare the two solutions on the total costs and the total time.

Setting up three agents for the simulation of this problem, they are as follows: (1) self-raising point agent; (2) warehouse agent, (3) distribution vehicle agent.

The distribution vehicle agent is the main activity target, including controlling the vehicle from the warehouse, searching for the closest self-raising point and unloading the goods, continuously searching for a new point, and returning to the warehouse until all goods are unloaded.

The time to accomplish the delivery of the self-raising points is calculated to compute the travelling costs. The time span consists of the vehicle travelling time from one point to another and the unloading time at the terminal.

The calculation formula of total costs is as follows:

$$C_{total} = C_{run} + C_{delay} \quad (1)$$

where

$$C_{run} = T_{run} * perMinCost \quad (2)$$

$$C_{delay} = T_{delay} * perMinCost \quad (3)$$

$$C_{total} = C_{run} + C_{delay} \quad (4)$$

$$T_{run} = t_{arrive} - t_{start} \quad (5)$$

$$T_{delay} = t_{arrive} - t_{limit} \quad (if \ t_{limit} < t_{arrive}) \quad (6)$$

$$T_{delay} = 0 \quad (if \ t_{limit} > t_{arrive}) \quad (7)$$

Meanwhile, the total time of entire system is calculated to judge if the delivery solution meets the criterion.

The calculation formula is

$$T_{total} = \sum_{i=1}^n T_{run}^i + T_{unload}^i \quad (8)$$

where

$$T_{unload} = orderNum * perMinUnloa \quad (9)$$

The logical structure to implement the functions of the distribution vehicle agents is composed of four states, six transitions, and one selection structure, as shown in Figure 4.

The following operation interface can be obtained through the above analysis and modeling, which is shown in Figure 5.

Simulation to the self-raising points of the E-commerce enterprise vehicle distribution can be performed by selecting the route planning mode and clicking the running button. Relevant data is obtained.

### 3. Simulation Performance and Optimization Analysis

This paper considers the logistics distribution settings under single warehouse, double warehouses, and three warehouses. The raw data of these three settings are all in Table 1. In a sharp increase of orders for special events like “Double Eleven”, we have increased the order number by ten times on the basis of the original data, which makes the research more realistic. Finally, we compare the results of simulation and provide suggestions for improvements.

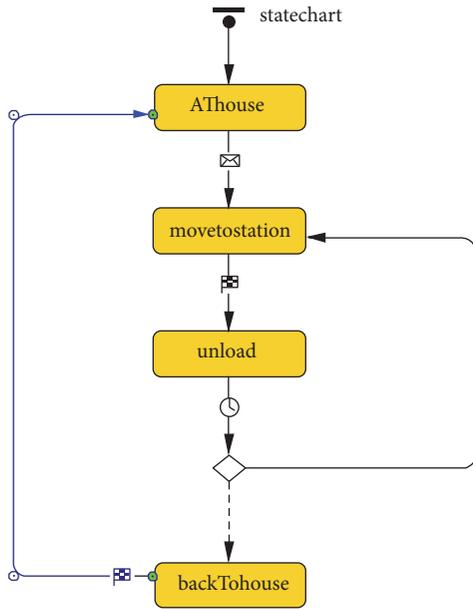


FIGURE 4: Logical structure of vehicle agent.

3.1. Comparison of Two Solutions in Single Warehouse Mode. Only one delivery at all distribution points is accomplished in this mode. Assuming the shortest route solution is selected to the next point, the following results are obtained by running the AnyLogic:

TotalRunCost:41657.899;  
 TotalDelayCost:1074320.791;  
 TotalCost:1115978.69.

Meanwhile, some critical data on the sequence of every distribution point, time nodes, and time length during distribution processes are shown in Table 2.

According to Table 2, obviously the total time accomplishing the whole process (among vehicle departing and returning to the warehouse) is 1541.58 virtual time with single warehouse (the shortest route solution).

As shown in Table 2, some goods have time-restricted requirements that must be delivered before time limits, or tardiness costs occurred. The time-limited precedence solution is utilized to minimize the tardiness costs when the next station route needs to be selected. In other words, the goods with time requirements should be delivered first and then the shortest route solution is adopted to deliver goods without time requirements.

Through AnyLogic platform, the results of distribution costs in single warehouse (time-limited precedence solution) are as follows:

TotalRunCost:49796.137;  
 TotalDelayCost:59418.424;  
 TotalCost:109214.56.

Meanwhile, 1622.96 virtual times are required to accomplish the entire distribution procedure.

Comparisons of the results on the shortest route solution and the time-limited precedence solution in single warehouse mode are shown in Table 3.

As presented in Table 3, the time-limited precedence solution causes the increments in total traveling costs by 19.54% and the abatements in total tardiness costs by 94.47%. The total costs are also reduced by 90.21%, because this solution can reduce the tardiness costs of the self-raising points with time limits effectively. Nevertheless, without adopting the shortest route solution, the vehicle travelling costs are raised and the total distribution time also increases by 5.28%. Because the amount of cost reduction is far larger than the increments, the total costs are reduced. In summary, under the circumstances of generating order number randomly, to the self-raising points with time-restricted requirements, the time-limited precedence solution has lower tardiness costs and total costs, but the traveling costs and distribution time increased merely.

Moreover, the total time span to complete the distribution is long whatever the two solutions are. In practice, reducing the delivery time to the customers significantly is an important problem the E-commerce companies face. This paper considers adding one warehouse, that is, double warehouses mode to solve this problem.

3.2. Simulation Comparison of Double Warehouses Mode and Single Warehouse Mode. The distribution in double warehouses mode is delivering goods to all the self-raising points simultaneously with two warehouses. According to the warehouse layout of the E-commerce enterprise, the distribution warehouse is added to the west of Shanghai interplaying with the original one. The distribution of double warehouses and self-raising points is shown in Figure 6.

The simulation consequences of double warehouses in the shortest route solution by AnyLogic are shown as follows:

TotalRunCost:42067.155;  
 TotalDelayCost:377939.599;  
 TotalCost:420006.754.

It takes 848.60 virtual time to accomplish the entire distribution process.

Similarly, following consequences of double warehouses can be obtained in the time-limited precedence solution:

TotalRunCost:51287.608;  
 TotalDelayCost:45760.399;  
 TotalCost:97048.007.

It takes 890.88 virtual time to accomplish.

The consequences of two solutions are compared in Table 4.

As shown in the Table 4, in double warehouses mode with the time-limited precedence solution, the total traveling costs and delivery time increase by 21.92% and 4.98%, the total tardiness costs and the total costs reduce by 87.89% and 76.89%. The reasons are the same as the single warehouse mode that without adopting the shortest route solution results in an increase in the travelling costs and entire delivery time. Nevertheless, the tardiness costs are significantly reduced so that the total costs are reduced.

In order to comprehend the influence of the warehouse quantity on the distribution costs and the delivery time, above results are compared to gain Tables 5 and 6.

By comparing Tables 5 and 6, the double warehouses model can significantly reduce the total time compared to

TABLE 1: Raw data.

Name	Latitude	Longitude	Number of orders	Product category	Delivery Time
s1	31.1431986	121.4210407	100	1	100
s2	31.15728268	121.4120836	50	1	150
s3	31.15546308	121.4371571	10	2	0
s4	31.16333719	121.4147434	90	1	200
s5	31.1343072	121.4481408	30	1	250
s6	31.16970695	121.4252906	20	2	0
s7	31.16707584	121.4416896	30	2	0
s8	31.18410185	121.4266127	40	2	0
s9	31.18935785	121.4276737	40	2	0
s10	31.19520877	121.4184094	100	1	300
s11	31.1988346	121.4461166	30	2	0
s12	31.18874702	121.371578	20	2	0
s13	31.20861033	121.4197887	50	1	350
s14	31.2231975	121.3954163	30	2	0
s15	31.27327301	121.4338336	40	2	0
s16	31.26069097	121.4462738	30	2	0
s17	31.26820805	121.4564257	100	1	400
s18	31.28828157	121.4429477	20	2	0
s19	31.29730166	121.4110734	40	2	0
s20	31.24911288	121.4890998	60	1	450
s21	31.27418872	121.4633177	40	2	0
s22	31.27321697	121.4688453	30	2	0
s23	31.2796735	121.4802926	80	1	500
s24	31.30568695	121.4669172	20	2	0
s25	31.23087848	121.4437099	40	2	0
s26	31.28454415	121.5018165	90	1	550
s27	31.26859412	121.5399282	20	2	0
s28	31.28451296	121.5115345	30	2	0
s29	31.29881095	121.499463	100	1	600
s30	31.30787625	121.4993919	90	1	650
s31	31.29473544	121.5409558	30	2	0
s32	31.2911163	121.5489509	80	1	700
s33	31.33387484	121.5288115	40	2	0
s34	31.31553869	121.3699688	30	2	0
s35	31.31371912	121.3505742	20	2	0
s36	31.35066483	121.4538375	40	2	0
s37	31.1259537	121.3866134	20	2	0
s38	31.18216029	121.4056214	40	2	0
s39	31.22972111	121.4052614	100	1	750
s40	31.22416167	121.3773722	30	2	0
s41	31.23514491	121.4078012	40	2	0
s42	31.23383979	121.4318426	30	2	0
s43	31.24339301	121.418824	20	2	0
s44	31.24539919	121.4137998	30	2	0
s45	31.23813096	121.355908	40	2	0
s46	31.25328122	121.3872342	30	2	0
s47	31.26148927	121.4184197	20	2	0
s48	31.25603512	121.3732177	30	2	0
s49	31.26155358	121.3545372	30	2	0
s50	31.28576633	121.3412785	40	2	0
s51	31.19722065	121.4698344	40	2	0

TABLE 2: Simulation results in single warehouse mode (the shortest route solution).

Name	Arrive Time	Left Time	Run Length	Unload Length	delayLength	runCost	Delay Cost	House Name
WareHouseA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	WareHouseA
s50	8.04	28.04	8.04	20.00	0.00	804.00	0.00	WareHouseA
s35	35.18	45.18	7.14	10.00	0.00	714.00	0.00	WareHouseA
s48	59.11	74.11	13.94	15.00	0.00	1394.00	0.00	WareHouseA
s46	76.46	91.46	2.34	15.00	0.00	234.00	0.00	WareHouseA
s49	97.33	112.33	5.87	15.00	0.00	587.00	0.00	WareHouseA
s45	119.04	139.04	6.71	20.00	0.00	671.00	0.00	WareHouseA
s40	143.81	158.81	4.77	15.00	0.00	477.00	0.00	WareHouseA
s14	163.05	178.05	4.24	15.00	0.00	424.00	0.00	WareHouseA
s41	182.12	202.12	4.07	20.00	0.00	407.00	0.00	WareHouseA
s39	204.91	254.91	2.79	50.00	0.00	279.00	0.00	WareHouseA
s43	259.37	269.37	4.46	10.00	0.00	446.00	0.00	WareHouseA
s44	271.22	286.22	1.85	15.00	0.00	185.00	0.00	WareHouseA
s42	292.76	307.76	6.55	15.00	0.00	655.00	0.00	WareHouseA
s25	311.13	331.13	3.36	20.00	0.00	336.00	0.00	WareHouseA
s13	337.75	362.75	6.62	25.00	0.00	662.00	0.00	WareHouseA
s10	367.06	417.06	4.31	50.00	67.06	431.00	13412.	WareHouseA
s9	419.23	439.23	2.18	20.00	0.00	218.00	0.00	WareHouseA
s8	441.26	461.26	2.03	20.00	0.00	203.00	0.00	WareHouseA
s38	465.79	485.79	4.52	20.00	0.00	452.00	0.00	WareHouseA
s4	490.17	535.17	4.38	45.00	290.1	438.00	58032	WareHouseA
s2	536.73	561.73	1.57	25.00	386.7	157.00	77346	WareHouseA
s1	566.45	616.45	4.71	50.00	466.4	471.00	93290	WareHouseA
s3	621.10	626.10	4.65	5.00	0.00	465.00	0.00	WareHouseA
s6	633.17	643.17	7.07	10.00	0.00	707.00	0.00	WareHouseA
s7	648.96	663.96	5.79	15.00	0.00	579.00	0.00	WareHouseA
s11	671.47	686.47	7.52	15.00	0.00	752.00	0.00	WareHouseA
s51	691.79	711.79	5.32	20.00	0.00	532.00	0.00	WareHouseA
s20	724.95	754.95	13.16	30.00	274.9	1316.00	54990	WareHouseA
s23	762.95	802.95	8.00	40.00	262.9	800.00	52590	WareHouseA
s22	806.11	821.11	3.16	15.00	0.00	316.00	0.00	WareHouseA
s21	823.64	843.64	2.53	20.00	0.00	253.00	0.00	WareHouseA
s17	845.93	895.93	2.29	50.00	445.9	229.00	89186	WareHouseA
s16	898.68	913.68	2.75	15.00	0.00	275.00	0.00	WareHouseA
s15	917.43	937.43	3.75	20.00	0.00	375.00	0.00	WareHouseA
s18	942.59	952.59	5.16	10.00	0.00	516.00	0.00	WareHouseA
s19	959.43	979.43	6.84	20.00	0.00	684.00	0.00	WareHouseA
s47	989.60	999.60	10.17	10.00	0.00	1017.00	0.00	WareHouseA
s24	1014.17	1024.17	14.57	10.00	0.00	1457.00	0.00	WareHouseA
s30	1030.21	1075.21	6.03	45.00	380.2	603.00	76042	WareHouseA
s29	1077.77	1127.77	2.56	50.00	477.7	256.00	95554	WareHouseA
s26	1131.39	1176.39	3.62	45.00	581.3	362.00	116278	WareHouseA
s28	1178.76	1193.76	2.37	15.00	0.00	237.00	0.00	WareHouseA
s31	1200.60	1215.60	6.84	15.00	0.00	684.00	0.00	WareHouseA
s32	1217.66	1257.66	2.06	40.00	517.6	206.00	103532	WareHouseA
s27	1264.99	1274.99	7.33	10.00	0.00	733.00	0.00	WareHouseA
s33	1288.87	1308.87	13.88	20.00	0.00	1388.00	0.00	WareHouseA
s36	1325.59	1345.59	16.72	20.00	0.00	1672.00	0.00	WareHouseA
s34	1371.66	1386.66	26.07	15.00	0.00	2607.00	0.00	WareHouseA
s12	1419.80	1429.80	33.13	10.00	0.00	3313.00	0.00	WareHouseA
s37	1446.56	1456.56	16.76	10.00	0.00	1676.00	0.00	WareHouseA
s5	1470.34	1485.34	13.78	15.00	1220	1378.00	244068	WareHouseA
WareHouseA	1541.58	1541.58	56.24	0.00	0.00	5624.00	0.00	WareHouseA

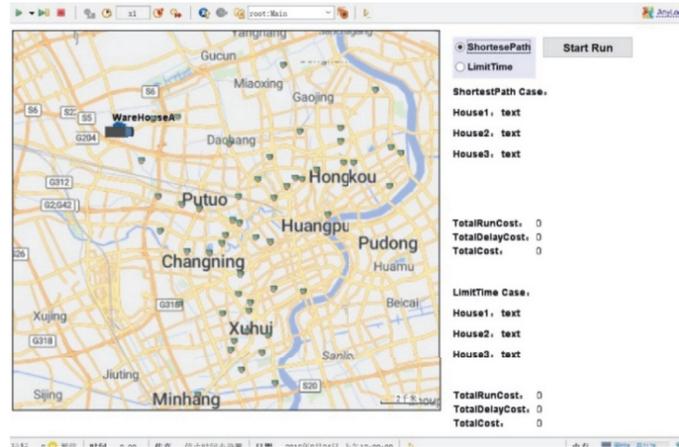


FIGURE 5: Operation interface of simulation platform.



FIGURE 6: Distribution of double warehouses and self-raising points.

TABLE 3: Comparison results of two solutions in single warehouse mode.

	ShortestPath	LimitTime	Compare Result
TotalRunCost	41657.899	49796.137	+19.54%
TotalDelayCost	1074320.791	59418.424	-94.47%
TotalCost	1115978.69	109214.561	-90.21%
FinishTime	1541.580	1622.960	+5.28%

TABLE 4: Comparison of two solutions in double warehouses mode.

	ShortestPath	LimitTime	Compare Result
TotalRunCost	42067.155	51287.608	+21.92%
TotalDelayCost	377939.599	45760.399	-87.89%
TotalCost	420006.754	97048.007	-76.89%
FinishTime	848.600	890.880	+4.98%

TABLE 5: Comparison results of two modes in shortest route solution.

	SingleHouse	DoubleHouse	Compare Result
TotalRunCost	41657.899	42067.155	+1%
TotalDelayCost	1074320.791	377939.599	-64.82%
TotalCost	1115978.690	420006.754	-62.36%
FinishTime	1541.580	848.600	-44.95%

TABLE 6: Comparison results of two modes in time-limited precedence solution.

	Single House	Double House	Compare Result
TotalRunCost	49796.137	51287.608	+3%
TotalDelayCost	59418.424	45760.399	-22.99%
TotalCost	109214.561	97048.007	-11.14%
FinishTime	1622.960	890.880	-45.11%

the single warehouse model in the two solutions. Particularly under the circumstances of the shortest route solution, the

tardiness costs and the total costs are reduced dramatically, because the two warehouses distribute goods simultaneously

TABLE 7: Comparison between normal order number and ten times of the shortest route solution in double warehouses mode.

	NomalOrders	10timesOrders	Compare Result
TotalRunCost	42067.155	42067.155	+0%
TotalDelayCost	377939.599	8148314.73	+2055.98%
TotalCost	420006.754	8190381.884	+1850.06%
FinishTime	848.600	6428.600	+657.55%

TABLE 8: Comparison between normal order number and ten times of the time-limited precedence solution in double warehouses mode.

	NomalOrders	10timesOrders	Compare Result
TotalRunCost	51287.608	51287.608	+0%
TotalDelayCost	45760.399	2758085.139	+5927.23%
TotalCost	97048.007	2809372.747	+2794.83%
FinishTime	890.880	6870.880	+671.25%

to different self-raising points which reduces the time. In the shortest route solution, the delay time is so long that reducing delivery time can have great effects, while, in the time-limited precedence solution, the delay time is tiny without evident efficiency. In conclusion, the double warehouses mode has more effects than the single one in tardiness costs and total costs in the shortest route solution.

Moreover, in special shopping festivals, for instance, 11.11, 618, and so on, goods quantities increase dramatically on every self-raising point of the E-commerce companies. Therefore, this article will enlarge ten times of the order number for each point to study this problem with AnyLogic.

Simulation results of double warehouses in the shortest route solution are shown (the order number is magnified by ten times):

TotalRunCost:42067.155;  
TotalDelayCost:8148314.73;  
TotalCost:8190381.884.

Meanwhile, the overall distribution time is 6428.60 virtual time.

Simulation time of double warehouses in the time-limited precedence solution is as follows (the order number is magnified by ten times):

TotalRunCost:51287.608;  
TotalDelayCost:2758085.139;  
TotalCost:2809372.747.

Meanwhile, the overall distribution time is 6470.88 virtual time.

Compared with the normal order number, two tables are obtained as in Tables 7 and 8.

As shown in Tables 7 and 8, the tardiness costs, the total costs, and the time requirements are many times larger than the normal one during the special shopping festival. The double warehouses model is completely incapable of meeting the practical demands. Therefore, this paper considers adding another warehouse to reduce the time and costs in three warehouses mode.

3.3. *Three-Warehouse Mode and the Influence of Distribution Area Adjustment on the Simulation Results.* Three warehouses are utilized to be distributed simultaneously in three-warehouse mode which is also the fundament for multiple warehouse distribution research. In this paper, the warehouse location of the E-commerce enterprise in Shanghai is taken as example. We expand the double one to three-warehouse model. The distribution of warehouses and self-raising points are shown in Figure 7.

The simulation consequences of three warehouses in the shortest route solution can be obtained by AnyLogic as follows:

TotalRunCost:45398.807;  
TotalDelayCost:132487.919;  
TotalCost:177886.726.

Meanwhile, the overall distribution time is 641.76 virtual time.

Similarly, the simulation consequences of three warehouses in the time-limited precedence solution are obtained as follows:

TotalRunCost:52468.394;  
TotalDelayCost:15795.939;  
TotalCost:68264.333.

Meanwhile, the overall distribution time is 676.48 virtual time.

Comparing the mentioned consequences with the two warehouses in two solutions, a significant improvement has occurred. However, by observing the three warehouses and distribution sites in Figure 7, the new warehouse WarehouseC is close to the WarehouseB. Therefore, the location of WarehouseC is adjusted to the southwest of Shanghai so that the nearby self-raising points have a relatively close warehouse to save the delivery time and total costs. The distribution of the adjusted warehouse and self-raising points are shown in Figure 8.

As shown in Figure 8, the WarehouseC is located in the southwest of Shanghai and these three warehouses are distributed evenly after adjustment of the location.

With the AnyLogic, the simulation consequences of three warehouses (adjusted location) in the shortest route solution can be obtained:

TotalRunCost:42057.032;  
TotalDelayCost:49689.094;  
TotalCost:91746.126.

Meanwhile, the overall distribution time is 697.08 virtual time.

Similarly, the simulation consequences of three warehouses (adjusted location) in the time-limited precedence solution can be obtained:

TotalRunCost:49418.701;  
TotalDelayCost:0;  
TotalCost:49418.701.

Meanwhile, the overall distribution time is 747.00 virtual time.

Comparison of the simulation results in three warehouses mode without adjusting, Tables 9 and 10, is obtained.

As shown in Tables 9 and 10, after the adjustment of warehouse location, the traveling costs, the tardiness costs, and the total costs are reduced, particularly in the



FIGURE 7: Distribution of three warehouses and stations.



FIGURE 8: Distribution of three warehouses (adjusted location) and stations.

TABLE 9: Comparison between original location and adjusted location of three warehouses in the shortest route solution.

	Original	Changed	Compare Result
TotalRunCost	45398.807	42057.032	-7.36%
TotalDelayCost	132487.919	49689.094	-62.5%
TotalCost	177886.726	91746.126	-48.42%
FinishTime	641.760	697.080	+8.62%

TABLE 10: Comparison between original location and adjusted location of three warehouses in the time-limited precedence solution.

	Original	Changed	Compare Result
TotalRunCost	52468.394	49418.701	-5.81%
TotalDelayCost	15795.939	0	-100%
TotalCost	68264.333	49418.701	-27.61%
FinishTime	676.480	747.000	+10.42%

time-limited precedence solution. All the time requirements in initialization data are met as the warehouse location adjusts and the tardiness costs are not generated. In this paper, the overall distribution time increases, which is primarily due to comprehensive conditions of self-raising points. If differences are shown, then the overall distribution time may be decreased.

#### 4. Conclusions

This paper uses AnyLogic simulation software to model and simulate the vehicle distribution process. Then the results are analyzed and optimized on three factors including routes selection, warehouses quantity, and warehouses layout. As

shown in the simulation consequences, the time-limited precedence solution can dramatically reduce the tardiness costs and the total costs; increasing the warehouses quantity can significantly lessen the overall delivery time; vehicle travelling costs, tardiness costs, and total costs can also be reduced by distribution of warehouse locations reasonably, which also have an influence on the overall delivery time. What can also be observed from the results is that the method studying logistics distribution by AnyLogic is feasible, which can visualize complicated problems and improve operability effectively. More optimal methods and algorithms like heuristic can be used in future research. The optimization module also can be contained.

## Data Availability

All the data used to support the findings of this study are included in our manuscript and are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] L. Zhen, "Tactical berth allocation under uncertainty," *European Journal of Operational Research*, vol. 247, no. 3, pp. 928–944, 2015.
- [2] L. Zhen, E. P. Chew, and L. H. Lee, "An integrated model for berth template and yard template planning in transshipment hubs," *Transportation Science*, vol. 45, no. 4, pp. 483–504, 2011.
- [3] P. F. I. Casas and J. Casanovas, "JGPSS, an open source GPSS framework to teach simulation," in *Proceedings of the 2009 Winter Simulation Conference, WSC 2009*, vol. 1-4, pp. 256–267, IEEE, Austin, Tex, USA, December 2009.
- [4] C. D. Pegden, A. Alan, and B. Pritsker, "SLAM: Simulation language for alternative modeling," *Simulation*, vol. 33, no. 5, pp. 145–157, 1979.
- [5] K. V. Ramani, "An Interactive Simulation Model for the Logistics Planning of Container Operations in Seaports," *Simulation*, vol. 66, no. 5, pp. 291–300, 1996.
- [6] I. Manuj, J. T. Mentzer, and M. R. Bowers, "Improving the rigor of discrete-event simulation in logistics and supply chain research," *International Journal of Physical Distribution & Logistics Management*, vol. 39, no. 3, pp. 172–201, 2009.
- [7] Q.-C. Meng, T. Zhang, M. Li, and X.-X. Rong, "Optimal order strategy in uncertain demands with free shipping option," *Discrete Dynamics in Nature and Society*, Art. ID 578280, 6 pages, 2014.
- [8] L. Geng, R. B. Xiao, and X. Xu, "Research on MAS-based supply chain resilience and its self-organized criticality," *Discrete Dynamics in Nature and Society*, vol. 2014, Article ID 621341, 14 pages, 2014.
- [9] Y. Cui, J. Shi, and Z. Wang, "Discrete Event Logistics Systems (DELS) simulation modeling incorporating two-step Remaining Useful Life (RUL) estimation," *Computers in Industry*, vol. 72, pp. 68–81, 2015.
- [10] P. Q. Li, J. He, D. Y. Zheng, Y. S. Huang, and C. H. Fan, "Vehicle routing problem with soft time windows based on improved genetic algorithm for fruits and vegetables distribution," *Discrete Dynamics in Nature and Society*, vol. 2015, Article ID 483830, 8 pages, 2015.
- [11] C. Y. Lin and B. W. Gong, "Transit-based emergency evacuation with transit signal priority in sudden-onset disaster," *Discrete Dynamics in Nature and Society*, vol. 2016, Article ID 3625342, 13 pages, 2016.
- [12] T. Beinke, A. Ait Alla, and M. Freitag, "Resource Sharing in the Logistics of the Offshore Wind Farm Installation Process based on a Simulation Study," *International Journal of e-Navigation and Maritime Economy*, vol. 7, pp. 42–54, 2017.
- [13] S. B. Li, G. M. Wang, T. Wang, and H. L. Ren, "Research on the Method of Traffic Organization and Optimization Based on Dynamic Traffic Flow Model," *Discrete Dynamics in Nature and Society*, vol. 2017, Article ID 5292616, 9 pages, 2017.
- [14] Y. Wei, F. Chen, and H. Wang, "Inventory and production dynamics in a discrete-time vendor-managed inventory supply chain system," *Discrete Dynamics in Nature and Society*, vol. 2018, Article ID 6091946, 15 pages, 2018.
- [15] T. G. Crainic, G. Perboli, and M. Rosano, "Simulation of intermodal freight transportation systems: a taxonomy," *European Journal of Operational Research*, vol. 270, no. 2, pp. 401–418, 2018.
- [16] M. Noorizadegan and B. Chen, "Vehicle routing with probabilistic capacity constraints," *European Journal of Operational Research*, vol. 270, no. 2, pp. 544–555, 2018.
- [17] W. B. Hu, H. Wang, Z. Y. Qiu, L. P. Yan, C. Nie, and B. Du, "An urban traffic simulation model for traffic congestion predicting and avoiding," *Neural Computing Applications*, vol. 30, no. 6, pp. 1769–1781, 2018.
- [18] 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

# Pricing Decisions on Reward-Based Crowdfunding with Bayesian Review System Facing Strategic Consumers

Yuting Chen,<sup>1</sup> Rong Zhang,<sup>2</sup> and Bin Liu <sup>1</sup>

<sup>1</sup>*School of Economics and Management, Shanghai Maritime University, Shanghai 201306, China*

<sup>2</sup>*Research Center of Logistics, Shanghai Maritime University, Shanghai 201306, China*

Correspondence should be addressed to Bin Liu; [liubin@shmtu.edu.cn](mailto:liubin@shmtu.edu.cn)

Received 31 October 2018; Revised 14 December 2018; Accepted 19 December 2018; Published 20 January 2019

Academic Editor: Lu Zhen

Copyright © 2019 Yuting Chen 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.

Crowdfunding marks a popular and sustainable means by which small and microentrepreneurs obtain financial resources for their innovative project. Consumers increasingly rely on online reviews to make purchase decisions. However, the crowdfunding nowadays lacks a form type of review system. This paper is designed to extend research on the optimal pricing decision with review system for the reward-based crowdfunding. Firstly, a Bayesian analysis is established to construct consumers' belief update process in presence of review system. Secondly, we take the strategies without the review system as a benchmark to explore the impacts of review system under preannounced pricing and responsive pricing. Finally, through the equilibrium analysis, we find that the review system has a positive impact on the creator under responsive pricing policy. The fraction of favorable review has a large effect on the profit of preannounced pricing. When the fraction is about 80%, the profit is the maximum. Generally speaking, the review system will make more profit for the creator.

## 1. Introduction

In recent years, crowdfunding has emerged as a new and sustainable channel for organizations or enterprises to receive money from a pool of individuals for different types of projects, especially for developing innovative products. Rewards-based crowdfunding constructs a funding context that enables project creators to obtain financial resources from the general public and sell the crowdfunding product at the same time [1]. According to the latest data, global crowdfunding experienced an accelerated growth in 2016 to reach \$19.96 billion (from \$6.1 billion in 2013) raised by some active platforms across the world. One of the leading rewards-based crowdfunding platforms is Kickstarter. In 2015 alone, 77000 projects were successfully funded there, raising around \$700 million from the general public from nearly every country on the planet. Reward-crowdfunding is regarded as a marketing and financing tool. The backers can participate in the project (hereafter called period 1) at a lower price and obtain the crowdfunding product after the project is successful (period 2). Besides, the consumers can directly purchase the product on the market in period 2 when the

crowdfunding project is successful. Our work develops a two-period model to discuss the optimal dynamic pricing decision for the creators and explore the effective way to reduce consumers uncertainty.

Launching a new crowdfunding product involves uncertainty; the chief uncertainty is the quality of the product. Some consumers may postpone or abandon their pledging and purchase decisions because they do not know the true quality of the crowdfunding product. Online product reviews have become an important information source for consumers to mitigate the uncertainty about the quality of a product. With the ubiquity of online shopping, the importance of social influence has never been greater; an increasing number of consumers make purchase decisions refer to online product reviews. According to a previous survey conducted by Deloitte and Touche, about 43% of consumers were reinforced of their original purchase decision by reviews, the same survey also found that 43% of consumers changed their intentions about which product to buy, and 9% of consumers even abandoned their purchase decision because of the negative reviews for the products. More recently, the advance of the Internet and online communities has

dramatically magnified the influence of reviews, as a product review online become an important information source for the consumers. While consumers initially are confuse about the true quality of the crowdfunding product online, they learn about it through some form of online reviews and adjust their estimates of its quality along the way and making their purchase decisions accordingly. Recently, Kickstarter encloses the Facebook Like button to describe crowdfunding projects and shows a counter reflecting the number of Facebook users who have previously “liked” this project. For subsequent visitors to the webpage, the button thus becomes a quality signal. It is also a form of informal online review. But there are few crowdfunding platforms containing online review systems. In this paper, we investigate how online reviews provided by previous consumers drive the dynamic pricing for crowdfunding products.

We characterize the consumers’ behaviors as follows. Consumers arrive at the market according to a Poisson process and face the decision of either participating in the crowdfunding (period 1) with unknown quality or waiting until the crowdfunding product is successful (period 2). The consumers differ in their base valuations for the observable attributes of the crowdfunding product, which determines their willingness to participate in the crowdfunding. These base valuation parameters are assumed to be independently and identically drawn from a known distribution, which is described in Section 3. Consumers who participate in the crowdfunding product in period 1 experience its true quality. In presence of online review system, those purchasers report whether they “liked” or “disliked” the product, i.e., if their ex-post utility was positive or negative, respectively. An arriving consumer in period 2 observes the history of purchasing decisions and reviews made by former buyers, combines this information with his prior estimate, infers the associated product quality, and makes his own purchasing decision. The sequence of reviews in period 2 affects the evolution of the observable information so that the dynamic of the market responds over time.

It is typical to assume that fully rational consumers update their estimate for the unknown quality of the product through a Bayesian analysis. Two alternative classes of dynamic pricing policies employed in our study come from the spirit of the seminal paper by Besanko and Winston [2]. One is preannounced pricing, which required the project creators announce the full price path from the beginning of period 1. The other is responsive pricing which means the creators announce only the first period price (crowdfunding pricing) and delay the second period price (market pricing) announcement until the beginning of the second period.

This paper illustrates how the review system affects the profits of crowdfunding creators and proposes the optimal dynamic pricing policies when considering strategy consumers’ behaviors and the heterogeneity in consumers’ valuations. The rest of this paper proceeds as follows. Section 2 presents related literature review. The basic Bayesian analysis model of the online review system is presented in Section 3. In Sections 4 and 5, the preannounced pricing policy and responsive pricing policy are analyzed respectively, and then we study the effect of an online system using the scenario

without review system as the benchmark. Section 6 concludes.

## 2. Literature Review

Some existing empirical and analytical works have studied crowdfunding pricing decision, including pricing and product design decisions in the context of crowdfunding [3]. Several papers have empirically investigated the consumers’ behavior caused by asymmetric quality information and observational learning from early contributions (see, e.g., [1, 4]). Chakraborty and Swinney [5] discuss how the creators may signal the quality of their projects through funding targets and how the creators’ behavior can be different under the objective of profit-maximization versus success-maximization. Buysere et al. [6] studies that crowdfunding is a group behavior in which many individuals contribute their resources to help others or organize an activity. Du et al. [7] focus on the contingent policies that creators can apply to the dynamic pledging progress after the project design has been determined. The authors demonstrate the importance of contingent policies and analyze three implementable policies. Chen et al. [8] propose an optimal pricing strategy for green crowdfunding products facing the mixed market of strategic and myopic consumers when substitutes exist. Araman and Caldentey [9] deal with the detailed modeling of the firm’s optimal timing to stop the voting and start or abandon launching the product. In a different set of new product launching, Marinési and Girotra [10] focus on forward-looking customer behavior and the intended use of the acquired information. Wessel et al. [11] assess the effects of nongenuine social information on consumers’ decision-making in the context of reward-based crowdfunding. Xu et al. [12] investigate a firm’s optimal funding choice when launching an innovative product to the market with both market uncertainty and word-of-mouth (WoM) communication. Our paper yields new insights into how crowdfunding review system improves consumers’ beliefs about the quality of crowdfunding product and presents the optimal dynamic pricing policy for creators.

The review systems have been studied both empirically and analytically in the literature. Many of these models focus on how the nature of a product in their market affects review outcome (e.g., positive or negative), and in turn how the reviews affect consumers’ willingness-to-pay and therefore the product demand. Banerjee [13] and Bikhchandani et al. [14] study the update process as the scenario where a consumer observes a signal and the decisions of the consumers who made a decision before him. Here, consumers are rational and update their beliefs in a Bayesian way. Ellison and Fudenberg [15, 16] consider that consumers exchange information about their experienced utility and use simple decision rules to choose between actions. Kwark et al. [17] view the reviews as information mitigating the uncertainty in a product quality to consumers’ needs and investigate how this additional information affects upstream product competition. Hu et al. [3] consider the opposite type of social behavior, where consumers value a product more if it is more popular in the population. Roma et al. [18] consider

an entrepreneur who designs a reward-based crowdfunding campaign when the campaign provides a signal about the future demand for the product and subsequent venture capital is needed. Crapis et al. [19] present a review learning trajectory based on mean-field approximations and studied the price optimization questions in the presence of online review. Jing [20] analyze the impacts of social learning (SL) on the dynamic pricing and consumer adoption of durable goods in a two-period monopoly. Jiang and Guo [21] analyze firms' review system design and product pricing strategies. The authors formally model two review system design decisions—what rating scale cardinality to use and whether to offer granular review reports. Another stream of research has modeled product reviews as free advertising and analyzed how sellers should adjust their own marketing mix strategies in the presence of the reviews [22–24]. Fasheng Xu et al. [12] explore the interaction between social learning and network externalities. We focus on the dynamic pricing decision on reward-based crowdfunding in presence of review system and take the crowdfunding time into consider. Existing learning models assume that consumers update information signals from review system in a Bayesian framework. Ifrach et al. [25] analyze a Bayesian model where both the quality of the product and the reviews can assume only two possible values, and they provided conditions for learning. Papanastasiou et al. [26] study the social learning by a Bayesian model and investigates how social learning affects the interaction between a dynamic pricing monopolist and a forward-looking consumer population, within a simple two-period model. Based on the above researchers, we extend the model to study the updating process in presence of review system and how online reviews improve the optimal profit of crowdfunding creators. This paper is a first attempt at understanding the online crowdfunding review system when considering consumer arrival to market according to the Poisson process and taking into account the consumers' heterogeneous valuations that declined over time.

Previous economics and marketing literature has studied humans' behaviors and two classes of pricing policies. Heterogeneous consumers are present throughout the entire season and optimally select the timing of their purchases to maximize individual surplus. Stokey [27] and Lansberger and Meilijson [28] assume that a monopolist commits to a price strategy for the entire selling horizon in order to maximize expected revenue. Besanko and Winston [2] propose a contingent subgame perfect pricing strategy for the monopolist. They consider a seller facing a finite pool of rational consumers that aim to maximize their individual utilities by optimally timing their purchases. Aviv and Pazgal [29] take into account fashion-like seasonal goods in the presence of forward-looking customers. The literatures that investigate dynamic pricing typically employ either preannounced pricing (e.g., [28]) or responsive pricing [2]. Responsive price plans generate value because they allow the firm to react optimally to updated information (e.g., [30]). The general agreement in the literature is that a company will prefer a preannounced pricing when facing forward-looking consumers (e.g., [31]). Yin et al. [32] and Whang [33] deal with demand learning by preannounced pricing. Cachon

and Swinney [31] examine the firm's quantity and salvage-pricing decisions under responsive pricing. This paper is also related to the pricing of experience goods, whose quality can be determined only upon consumption (e.g., [34, 35]). Our work extends the above studies, especially Papanastasiou and Savva [36]'s model of social learning and Aviv and Pazgal [29] contingent Pricing Model, to deal with the optimal pricing of crowdfunding experience goods with online review system. We assume that consumers in our model face uncertainty regarding the intrinsic quality of a new and innovative crowdfunding product.

### 3. The Benchmark Model

We consider a creator posting a crowdfunding product on a platform and discuss the creator's review system choices and pricing strategies over two periods. The first period is known as the crowdfunding period, in which consumers raise money for the project. In the second period, the crowdfunding product sells on the market. The people purchased in the first period also gain product in the second period. At the beginning of the first period, the creator makes its pricing decision. Because the creator will charge different prices in the different period, first period consumers may choose to purchase in the first period or wait to purchase in the second period. If the creator chooses to host a review system, consumers who purchase in the first period may post their product reviews, and consumers in the second period can learn from these reviews before making their purchasing decision. The total number of consumers in both periods is normalized to one. Each consumer only demands one unit of the crowdfunding product during the selling cycle, and there is no activity after the second period.

*3.1. Model Assumption.* The sales season is  $[0, H]$ , which is divided into two periods,  $[0, T]$  and  $[T, H]$ . The crowdfunding product is launched at time  $t=0$ . Consumers arrive thereafter according to a Poisson process with rate  $\lambda$ , which is independent of the product's quality and consumers' preference parameters. Consumer  $k$ 's valuation of the crowdfunding product comprises two components: observable attributes (like color or style) and unobservable attributes (like quality or fit). The consumer  $k$  has a base valuation  $V_k$  for the observable attributes of crowdfunding product and is heterogeneous in terms of their preferences for different product attributes. Without loss of generality, we assume that base valuation  $\{V_k\}$  is independent and identically distributed (i.i.d.) random variables drawn from a known distribution function  $F$ . We denote the corresponding density function by  $f$ . The valuation  $V_k$  is discounted by a known exponential function  $e^{-\alpha x}$ , where  $\alpha \geq 0$  represents the sensitive degree of waiting and  $x$  is denoted as the waiting time. To reflect these, we use a valuation function developed by Aviv and Pazgal [29]:

$$V_k(t) = V_k e^{-\alpha x}. \quad (1)$$

Crowdfunding platforms like Kickstarter, Jingdong, and so on always require creators set a target for their crowdfunding projects. The project is deemed to be successful only if the

total funding exceeds the target. Consumers will participate in the crowdfunding in the first period by a lower price  $p_1$  and gain products in second period if the crowdfunding is successful. On the contrary, the platform will return money to consumers. Thus, consumers face a risk ( $\mu$ ) of whether the project will be successful and whether they can receive the product when taking part in the crowdfunding project in first period. Because of the risk of the crowdfunding, consumers' valuation of participating in the crowdfunding in the first period is lower than the valuation after the project is proved to be successful. Consumers' valuation for participating in the first period can be expressed by  $\mu V_k(t)$  ( $0 < \mu < 1$ ). In the second period, the market price  $p_2$  is higher than first period.

Each consumer has his private information about his arrival time and his own valuation for the product. The creator knows the product's real attributes and consumers' valuation distribution  $F$ , but he does not know individual consumers' valuation. Other parameters  $\lambda$ ,  $\alpha$ ,  $T$ ,  $H$ ,  $\mu$  are known to the creator and all consumers. We assume that all of consumers are strategic, and they will make their decisions by comparing the expected surplus in both periods.

**3.2. Consumers' Belief Updating Process.** When the creator hosts an online review system, second period consumers can discover more about the product from review and update their beliefs on product valuation and different dimensions of quality, such as comfort or durability. We use  $s \in Z$  to denote the rate scales. For example,  $s=2$  represents the case of two rating levels such as "like" and "dislike". After consumption, consumers know the true quality of the product and how well the product fits their tastes. Park and Nicolau [37] showed that people perceive extreme ratings (positive or negative) as more useful and enjoyable than moderate ratings. We consider that a consumer who purchased the product reports that "like" if his utility is nonnegative and the true quality and their preference; he will report "dislike" if his utility is negative. We assume that the review rate level follows a symmetric binomial distribution  $B(\theta, n)$ , where  $n$  represents the review number. Samples  $X_i$  are independent and identically distributed (i.i.d.). Through the advertisement or introduction on the crowdfunding platform, consumers have prior valuations about the crowdfunding product during the first period. To simplify the analysis, we assume that consumers' prior belief about the product follow a uniform distribution  $U[0, 1]$ .

$$q(\theta) = \begin{cases} 1 & 0 < \theta < 1, \\ 0 & \text{other.} \end{cases} \quad (2)$$

Thanks to the review system, first period consumers rate the product based on their consumption. By observing the previous reviews, second period consumers will update their beliefs on product valuation and product fit. We use a Bayesian model to analyze how the review system interacts with strategic consumers.  $q(\cdot)$  is the consumers' prior belief about the product and  $q(x | \cdot)$  is the posterior distribution of consumers' belief.

According to the analysis above, the prior belief distribution is

$$P(x | \theta) = C_n^x \theta^x (1 - \theta)^{n-x}. \quad (3)$$

So  $\theta$  and samples  $X_i$ 's joint probability is

$$h(x, \theta) = q(\theta) P(x | \theta) = C_n^x \theta^x (1 - \theta)^{n-x}, \quad (4)$$

where  $x=0,1,2,\dots,n$ ,  $0 < \theta < 1$ . The marginal density of  $X_i$  is

$$\begin{aligned} m(x) &= \int_{\Theta} f(x | \theta) dF^{\pi}(\theta) = \int_{\Theta} P(x | \theta) q(\theta) d\theta \\ &= \int_0^1 C_n^x \theta^x (1 - \theta)^{n-x} d\theta \\ &= C_n^x \frac{\Gamma(x+1) \Gamma(n-x+1)}{\Gamma(n-2)}. \end{aligned} \quad (5)$$

Thus, according to Bayes' rule, the posterior distribution of consumers' belief can be expressed as follows:

$$\begin{aligned} q(x | \theta) &= \frac{h(x, \theta)}{m(x)} \\ &= \frac{\Gamma(x+1) \Gamma(n-x+1)}{\Gamma(n-2)} \theta^x (1 - \theta)^{n-x}, \end{aligned} \quad (6)$$

where  $\Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x} dx$  ( $s > 0$ ) is called Gamma function and  $\Gamma(n+1) = n!$ . We denote  $\gamma$  as the fraction of favorable review. Then  $x = \gamma n$ . The posterior distribution is the Beta distribution,  $\beta(\gamma n + 1, n - \gamma n + 1)$  according to the expression above.

## 4. Preannounced Pricing

The creators must decide the pricing policy at the beginning of the sale season. When consumers are strategic, they may postpone the purchase in hope of more information. With such uncertainty, many consumers may end up not buying and lead to crowdfunding failure, reducing the creators' profitability. Recent studies show that a way to mitigate this negative effect caused by strategic consumers is to use a posted or preannounced pricing policy (Correa et al. [38]). In this section, we investigate the optimal pricing decision in preannounced pricing policy. We denote  $p_i$  as the product price in period  $i$ , where  $i=1,2$ . In this policy, the creators announce  $p_1$  and  $p_2$  in the first period and consumers make their first period purchase decision by comparing expect surplus. In presence of review system, consumers remaining in the market update their valuation by the previous online reviews and make their second period purchase decision. Our focuses are on analyzing how the review system impact on the creator's profit and discovering the optimal design of pricing mechanism with preannounced prices.

**4.1. Benchmark: Preannounced Pricing in Absence of Review System.** We discuss the preannounced pricing without review system as a benchmark. In this section, we consider a

two-period pricing problem in which the creator announces the fixed price path  $\{p_1, p_2\}$  at the beginning of the first period. The strategic consumers compare their surpluses in both periods by considering waiting cost and crowdfunding risk. As shown in Figure 1, a consumer arrives at the crowdfunding platform and notices the project at time  $t$  ( $0 < t < T$ ); his valuation of attending this crowdfunding project is  $\mu V_k e^{-\alpha(T-t)}$ . Moreover, the valuation of this product will be  $V_k e^{-\alpha(T-t)}$  if the consumer waits until the second period and purchase the product at time  $T$ . If a consumer notices the product at time  $t$  ( $T \leq t \leq H$ ), the valuation is  $V_k$  because of no waiting. In order to describe the consumers' purchasing decision specifically, a threshold function is provided in Theorem 1.

**Theorem 1.** *The threshold function  $\chi(t)$  for preannounced pricing without review system can be expressed as follows:*

$$\chi(t) = \begin{cases} \frac{p_1 - p_2}{\mu - 1} e^{\alpha(T-t)} & 0 < t < T, \\ p_2 & T \leq t \leq H. \end{cases} \quad (7)$$

*Proof.* See the appendix.  $\square$

The consumer  $k$  will take part in the crowdfunding or purchase the product if  $V_k \geq \chi(t)$  in both periods. In addition, if the consumer visits the platform before time  $T$  and  $V_k < \chi(t)$ , he will revisit the platform at time  $T$  and purchases the product if the valuation is larger than the threshold at this time. However, in the preannounced pricing policy, the strategic consumers arriving at  $[0, T]$  will weigh their valuation and threshold according to the fixed price path  $\{p_1, p_2\}$ . Thus, they can decide to purchase or postpone or even abandon the product immediately. According to Theorem 1, there are four types of decisions about the product. The first one is that the consumers arriving in the first period and  $V_k \geq \chi(t)$ ; this type of consumers will buy it in the first period. The second type is that the consumers arrive in the first period, but their valuation is lower than the threshold  $V_k < \chi(t)$ ; this type of consumers will postpone their purchasing decision until the second period. And the third type is that the consumers notice the product in the second period; if  $V_k \geq \chi(t)$ , they will buy one immediately. The rest of the consumers will not buy anyone during the selling season. Then we discuss the expected number of each type as follows:

As the assumption above, consumers arrive at the platform and notice the project according to a Poisson process with a rate of  $\lambda$ . This process is independent of the valuation or the pricing policy. And consumers are heterogeneous in their common valuation as well as in their arrival time. The common valuation is drawn from a distribution  $F$ .

The condition that consumers take part in the crowdfunding in the first period should be satisfied:

$$\begin{aligned} 0 < t < T, \\ V_k \geq \chi(t). \end{aligned} \quad (8)$$

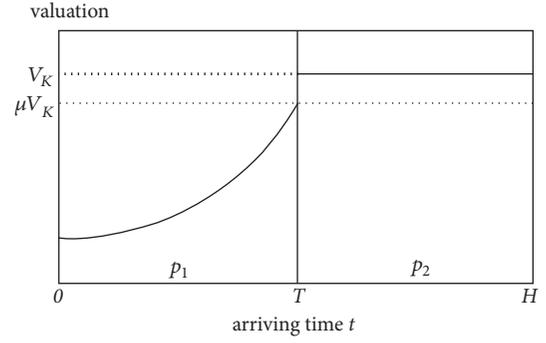


FIGURE 1: The consumer's valuation without review system.

Thus, the expected population of the first type without review system can be expected as

$$n_1^{pb} = \lambda \int_0^T F\left(\frac{p_1 - p_2}{\mu - 1} e^{\alpha(T-t)}\right) dt. \quad (9)$$

The case that consumers arrive in the first period and purchase the product in the second period is satisfied:

$$\begin{aligned} 0 < t < T, \\ V_k < \chi(t), \\ V_k e^{-\alpha(T-t)} \geq p_2. \end{aligned} \quad (10)$$

The expected population of the second type is

$$n_2^{pb} \equiv \lambda \int_0^T \left[ F(p_2 e^{\alpha(T-t)}) - F\left(\frac{p_1 - p_2}{\mu - 1} e^{\alpha(T-t)}\right) \right] dt. \quad (11)$$

The third kind of consumers arrives in the second period and purchases the product as long as their valuation is larger than the price.

$$\begin{aligned} T < t < H, \\ V_k \geq p_2. \end{aligned} \quad (12)$$

And the expected number of the third type of consumers is

$$n_3^{pb} \equiv \lambda \int_T^H F(p_2) dt. \quad (13)$$

According to the analysis above, the total profit of creator in preannounced pricing policy without review system is

$$\begin{aligned} \max_{(p_1, p_2)} \quad \pi^{pb} &= (p_1 - c) n_1^{pb} + (p_2 - c) (n_2^{pb} + n_3^{pb}) \\ \text{s.t.} \quad p_1 n_1^{pb} &\geq K. \end{aligned} \quad (14)$$

Here,  $p_1, p_2$  are decision variables and  $c$  is a constant.  $n_1^{pb}, n_2^{pb}, n_3^{pb}$  are the expected number of each kind of consumer, which is a function of  $p_1, p_2$  and impacted by  $\alpha, \mu, F$ , and so on.  $K$  is a fixed cost and  $K$  is the target of the crowdfunding.

When the funds reach the target value  $K$ , the crowdfunding is successful. Otherwise, the platform will return all funds to the funders.

The creator's optimal pricing strategy without review system is as follows.

**Proposition 2.** *Suppose that  $F$  follows a uniform distribution; the optimal pricing strategy of preannounced pricing in absence of online review system is*

$$\begin{aligned} p_1^* &= c, \\ p_2^* &= \frac{c(T - Te^{\alpha T} - H\alpha + T\alpha + T\mu - T\mu e^{\alpha T} + T\alpha\mu - T\alpha\mu)}{2(T\alpha - H\alpha + T\mu - T\mu e^{\alpha T} + H\alpha\mu - T\alpha\mu)}. \end{aligned} \quad (15)$$

*Proof.* See the appendix.  $\square$

It is obvious that the first period price should be low enough to ensure successful crowdfunding. And when the risk of crowdfunding is small  $\mu \rightarrow 1$ , the prices of first period and second period tend to be the same under preannounced pricing without review system. In absence of review system, with the risk of crowdfunding, the valuations of product in both periods tend to be consistent. The consumers waiting until the second period also are confused about the unobservable attribute of the product.

**4.2. Preannounced Pricing in Presence of Review System.** The consumers tend to make purchase decision refer to online product reviews. Smith [39] states that 60% of consumers consider ratings and reviews important when researching products. The consumers update their valuations of the crowdfunding product through the online reviews from previous consumers. Through the review system, the consumers will reduce uncertainty to the crowdfunding product.

In presence of review system, the consumers purchased in first period post their reviews on the platform after experiencing the product. If the consumer's utility is larger than the price he paid, he will announce his review as "Like"; otherwise, he will announce "Dislike". As discussed in the previous section, the reviews from consumers follow a binomial distribution. To simplify to the model, we assume that the consumers experience the product and post the review as soon as they receive the product. The rest of the consumers will update their valuations through those reviews. Similarly, as shown in Figure 2, the consumer  $k$ 's valuation of participating in crowdfunding in the first period is  $\mu V_k e^{-\alpha(T-t)}$  (where  $\mu$  is the risk of crowdfunding). However, if the consumer postpones his purchasing decision until the second period, the valuation of the product will be  $V_k e^{-\alpha(T-t)} + \theta$  where  $\theta$  means the belief updating from the review system.  $\pi(x | \theta)$  is deduced from the Bayesian model in the previous section. If a consumer arrives at the platform in the second period, his valuation of the product will be  $V_k + \theta$ .

Thanks to the review system, the consumers remaining in the market gain more information about the product.

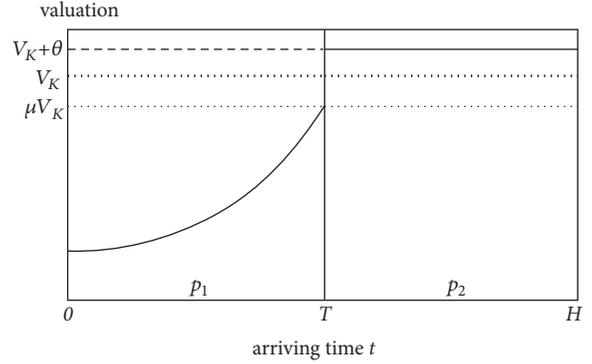


FIGURE 2: The consumer's valuation with the review system.

However, more consumers will tend to postpone their purchase to the second period, which leads to gathering less information and failing in crowdfunding. Theorem 3 expresses the equilibrium of consumers' decision and the creators' pricing strategy.

**Theorem 3.** *The threshold function  $\phi(t)$  for preannounced pricing with review system is*

$$\phi(t) = \begin{cases} \varphi(t) & 0 < t < T, \\ p_2 & T \leq t \leq H. \end{cases} \quad (16)$$

where  $\varphi(t)$  is the unique solution to the implicit equation:

$$\begin{aligned} \mu \varphi e^{-\alpha(T-t)} - p_1 &= E(\max\{\varphi e^{-\alpha(T-t)} + \theta - p_2, 0\}) \\ &= \int_{p_2 - \varphi e^{-\alpha(T-t)}}^1 (\varphi e^{-\alpha(T-t)} + \theta - p_2) q(x | \theta) dx. \end{aligned} \quad (17)$$

*Proof.* See the appendix.  $\square$

For any given pricing path  $\{p_1, p_2\}$ , it is optimal for consumers to make their purchasing decision according to a threshold function  $\phi(t)$ . If consumers' valuation of time  $t$   $V(t) \geq \phi(t)$ , they will purchase the product (or attend in the crowdfunding in the first period). If a consumer arrives in the first period and the valuation  $V_k < \varphi(t)$ , but  $V_k e^{-\alpha(T-t)} \geq p_2$ , he will postpone the purchase until the second period. The left-hand side of the equation represents the consumer's surplus of participating in the crowdfunding in the first period. The right-hand side of the equation represents the expected surplus of purchasing a product in the second period.  $\pi(x | \theta)$  is deduced from the Bayesian analysis in the previous section. According to the Bayesian analysis of review system, the posterior distribution of consumers' belief is a Beta distribution,  $\beta(\gamma n + 1, n - \gamma n + 1)$ , and  $q(x | \theta) = (\Gamma(x + 1)\Gamma(n - x + 1)/\Gamma(n + 2))\theta^x(1 - \theta)^{n-x}$ .

Based on the analysis above, the consumers can divide into four kinds. If the consumer would like to participate in the crowdfunding in the first period and obtain the product

in the second period, his valuation and arriving time should be satisfied:

$$\begin{aligned} 0 < t < T, \\ V_k \geq \varphi(t). \end{aligned} \quad (18)$$

The expected number of consumers joining in the crowdfunding in the first period is

$$n_1^{ps} = \lambda \int_0^T F(\varphi(t)) dt. \quad (19)$$

The condition that the consumer will postpone his purchase until the second period should be satisfied:

$$\begin{aligned} 0 < t < T, \\ V_k < \varphi(t), \\ V_k e^{-\alpha(T-t)} \geq p_2. \end{aligned} \quad (20)$$

The expected population of consumers postponing their purchase until the second period is

$$n_2^{ps} \equiv \lambda \int_0^T [F(p_2 e^{\alpha(T-t)}) - F(\varphi(t))] dt. \quad (21)$$

The expected population of consumers arriving in the second period and buying the product immediately is

$$n_3^{ps} \equiv \lambda \int_T^H F(p_2) dt. \quad (22)$$

Thus, the creator's total profit in both periods with the review system can be expressed in the following maximization problem:

$$\begin{aligned} \max_{(p_1, p_2)} \quad & \pi^{ps} = (p_1 - c)n_1^{ps} + (p_2 - c)(n_2^{ps} + n_3^{ps}) \\ \text{s.t.} \quad & p_1 n_1^{ps} \geq K, \end{aligned} \quad (23)$$

where  $K$  is a fixed cost of this production process and also the funding target of the crowdfunding.  $n_1^{ps}$  is the demand in the first period, and  $n_2^{ps} + n_3^{ps}$  is the demand in the second period.

## 5. Responsive Pricing

Under the responsive pricing, a creator announces first period price  $p_1$  according to the actual product attribute and estimated demand. The second period  $p_2$  pricing decision is constrained by the realized demand of first period and impacted by the review rate level generated by the first period in present of the review system. In short,  $p_2$  here is endogenous and posts at the beginning of second period according to sales feedback in first period. With responsive pricing, a creator can adjust their pricing strategy in response to market in time. We consider the subgame that begins after the first period price is fixed. Similarly, our study begins with the optimal responsive pricing strategy without a review system.

*5.1. Benchmark: Responsive Pricing in Absence of Review System.* Under the responsive policy, the creator announces the first period price at the beginning of the first period. The consumers participating in the crowdfunding would get the product until time  $T$ . The second period price  $p_2$  is announced at time  $T$  according to the demand in the first period. Actually, there are some intervals between the end of the first period and the beginning of the second period. To simplify the model, we ignore the intervals. The second period price  $p_2$  and the purchase decision by consumers form a Nash equilibrium in the subgame. In this section, we take the pricing without review system as the benchmark as well.

First, we consider the creator's profit of second period. Similar to the preannounced pricing, the profit in the second period is

$$\pi_2^{rb} = (p_2 - c)(n_2^{rb} + n_3^{rb}). \quad (24)$$

where  $n_2^{rb}$  and  $n_3^{rb}$  are the expected number of second and third type consumers, respectively. To analyze the demand for the second period, the purchase strategies of consumers should be considered first. The consumers' purchase strategies are described by the threshold function as well, as shown in Theorem 4.

**Theorem 4.** *The threshold function  $\psi(t)$  for responsive pricing without review system is*

$$\psi(t) = \begin{cases} \nu_1(t) & 0 < t < T, \\ \nu_2(t) & T \leq t \leq H, \end{cases} \quad (25)$$

where  $\nu_1(t)$  and  $\nu_2(t)$  are unique solutions to the implicit equations

$$\begin{aligned} \mu \nu_1 e^{-\alpha(T-t)} - p_1 &= E(\max\{\nu_1 e^{-\alpha(T-t)} - p_2^*, 0\}), \\ \nu_2 - p_2^* &= 0. \end{aligned} \quad (26)$$

*Proof.* See the appendix.  $\square$

Under the responsive pricing policy without review system,  $\psi(t)$  is denoted as the threshold function. A consumer  $k$  arrives at the platform in the first period; he will take part in the crowdfunding if  $V_k \geq \nu_1(t)$ ; he will not purchase a product until the second period if  $V_k < \nu_1(t)$  and  $V(t) \geq \nu_2(t)$ . If the consumer  $k$  arrives in the second period, he will purchase the product immediately if  $V_k \geq \nu_2(t)$ . In the first equation, the left-hand side is the consumer surplus of taking part in the crowdfunding in the first period; the right-hand side is the expected surplus of buying the product in the second period. In the second equation,  $\nu_2 - p_2^*$  means the surplus of the consumer arriving in the second period. The second equation is the critical condition of consumers buying a product in the second period immediately.  $p_2^*$  is deduced from  $\pi_2^{rb}$ . In order to analyze  $p_2^*$ , we consider the expected demands of the four types firstly.

Similar to the preannounced pricing, the condition that the first period consumer will fund the crowdfunding in the first period is

$$\begin{aligned} 0 < t < T, \\ V_k \geq v_1(t). \end{aligned} \quad (27)$$

Thus, the expected demand of the first period with responsive pricing in absence of review system is

$$n_1^{rb} = \lambda \int_0^T (F(v_1(t))) dt. \quad (28)$$

Similarly, the condition that consumers will postpone their purchase until the second period should be satisfied:

$$\begin{aligned} 0 < t < T, \\ V_k < v_1(t), \\ V_k e^{-\alpha(T-t)} \geq v_2(t). \end{aligned} \quad (29)$$

And the expected demand of this type of consumers is

$$n_2^{rb} = \lambda \int_0^T [(F(v_2(t) e^{\alpha(T-t)}) - F(v_1(t)))] dt. \quad (30)$$

The condition that the second period consumer will buy the product immediately should be satisfied:

$$\begin{aligned} T < t < H, \\ V_k \geq p_2. \end{aligned} \quad (31)$$

The expected number of the third type is

$$n_3^{rb} = \lambda \int_T^H F(v_2(t)) dt. \quad (32)$$

$n_2^{rb} + n_3^{rb}$  is the demand of the second period. The optimal second period price  $p_2^*$  is derived from

$$\frac{\partial \pi_2^{rb}}{\partial p_2} = 0. \quad (33)$$

Here,  $p_2^*$  is a function of  $v_1$  and  $v_2$ , we denote the optimal second period price deduced from the subgame is  $p_2^*(v_1, v_2)$ . From the implicit equations above, we can obtain the threshold functions  $v_1$  and  $v_2$ .

The optimal pricing decision ( $p_1^*$ ,  $p_2^*$ ) and the maximum income of the creator ( $\pi^{rb}$ ) are derived from the following nonlinear programming:

$$\begin{aligned} \max_{p_1, p_2} \quad & \pi^{rb} = (p_1 - c)n_1^{rb} + (p_2 - c)(n_2^{rb} + n_3^{rb}) \\ \text{s.t.} \quad & p_1 n_1^{rb} \geq K. \end{aligned} \quad (34)$$

**Proposition 5.** Suppose that  $F$  follows a uniform distribution. In the absence of review system, the optimal pricing decision

of the creator with common valuation following a uniform distribution is

$$\begin{aligned} p_1^* &= \frac{c(1 + 2T\alpha + 2H\alpha(\mu - 1) - 2\mu - 2T\alpha\mu + e^{T\alpha}(2\mu - 1))}{3 + 4T\alpha + 4H\alpha(\mu - 1) - 4\mu - 4T\alpha\mu + e^{T\alpha}(4\mu - 3)}, \\ p_2^* &= \frac{1}{6}c \left( 3 + \frac{1}{\mu} + \frac{2}{4\mu - 3} \right. \\ &\quad \left. - \frac{(H - T)\alpha(\mu - 1)}{\mu((H - T)\alpha(\mu - 1) + (e^{T\alpha} - 1)\mu)} \right. \\ &\quad \left. - \frac{8(H - T)\alpha(\mu - 1)}{(4\mu - 3)(3 + 4(H - T)\alpha(\mu - 1) - 4\mu + e^{T\alpha}(4\mu - 3))} \right). \end{aligned} \quad (35)$$

*Proof.* See the appendix.  $\square$

With the review system, the optimal pricing is impacted by the crowding time, the risk sensitive degree, and the sensitive degree of waiting. The larger  $\mu$  and  $\alpha$  are, the higher price and profit are.

**5.2. Responsive Pricing in Presence of Review System.** In presence of review system, the first period consumers announce their review on the platform, and the rest of the consumers will update their valuation in the second period. We assume that all consumers taking part in the crowdfunding in the first period will post their review at the same time as they get the product. Suppose that, in the second period, consumers and creator are able to obtain all reviews for the consumers in the first period. Thus, the creator will respond to the review level of the previous consumers.

Similar to the strategy without review system, we consider the second period profit first. The profit is

$$\pi_2^{rs} = (p_2 - c)(n_2^{rs} + n_3^{rs}). \quad (36)$$

**Theorem 6.** The threshold function  $\psi'(t)$  for responsive pricing with review system is

$$\psi'(t) = \begin{cases} v'_1(t) & 0 < t < T, \\ v'_2(t) & T \leq t \leq H \end{cases} \quad (37)$$

where  $v'_1(t)$  and  $v'_2(t)$  are unique solutions to the implicit equations

$$\begin{aligned} \mu v'_1 e^{-\alpha(T-t)} - p_1 &= E(\max\{v'_1 e^{-\alpha(T-t)} + \theta - p_2^*, 0\}) \\ &= \int_{p_2^*(v_2) - \varphi e^{-\alpha(T-t)}}^1 (\varphi e^{-\alpha(T-t)} + \theta - p_2^*) f(x, \theta) d\theta, \\ p_2^* &= v'_2. \end{aligned} \quad (38)$$

*Proof.* To see the appendix.  $\square$

In presence of review system, the consumers purchasing strategies are based on the threshold function  $\psi'(t)$ . When  $0 < t < T$  and  $V_k \geq v'_1(t)$ , the consumer will take part in the crowdfunding in the first period; if here  $V_k < v'_1(t)$  and  $V_k e^{-\alpha(T-t)} < v'_2(t)$ , the consumer will not buy a

product until the second period. If  $T \leq t \leq H$  and  $V_k \geq v'_2(t)$ , the consumer arriving in the second period would buy the product. In the first equation, the left-hand side is the consumer surplus of attending in the crowdfunding in the first period; the right-hand side is the expected surplus of postponing the consumer's purchase.  $p_2^*(v'_1)$  and  $p_2^*(v'_2)$  are optimal second period prices derived by the second period profit.  $\theta$  is deduced by Bayesian analysis in the previous section. The expected demands of each type are

$$n_1^{rs} = \lambda \int_0^T (F(v'_1(t))) dt. \quad (39)$$

$$n_2^{rs} = \lambda \int_0^T [(F(v'_2(t) e^{\alpha(T-t)}) - F(v'_1(t)))] dt. \quad (40)$$

$$n_3^{rs} = \lambda \int_T^H F(v'_2(t)) dt. \quad (41)$$

Substitute  $n_2^{rs}$  and  $n_3^{rs}$  into  $\pi_2^{rs}$ . From  $\delta\pi_2^{rs}/\delta p_2 = 0$ , we can get  $p_2^*(v'_1, v'_2)$ . Despite the implicit equations above,  $v'_1$  and  $v'_2$  are deduced and expressed by  $p_1, \mu, \alpha$ , and so on. The maximum profit and optimal pricing decision can be gotten by the following nonlinear programming:

$$\begin{aligned} \max_{p_1, p_2} \quad & \pi^{rs} = (p_1 - c)n_1^{rs} + (p_2 - c)(n_2^{rs} + n_3^{rs}) \\ \text{s.t.} \quad & p_1 n_1^{rs} \geq K. \end{aligned} \quad (42)$$

## 6. Equilibrium Analysis

In this section, we take a closer look at the impacts of the review system and pricing policy by the equilibrium analysis. In order to explore the expected profit performance impacted by the sensitive degree of waiting and the good review level, a numerical study is conducted. Suppose that the common valuation  $\{V_k\}$  follows a uniform distribution; the Passion process rate  $\lambda = 1$ ; the risk of participating in the crowdfunding  $\mu = 0.5$ . The length of the sale season  $H = 1$ . To simplify the parameter, we set  $\rho = e^\alpha$ . The comparison of profits under preannounced pricing and responsive pricing and the impact of the review system is shown by Figures 3–7.

(1) *Under Preannounced Pricing.* In the preannounced pricing policy, creator announces  $p_1$  and  $p_2$  at the beginning of the first period. The strategic consumers will make their purchasing decision through the threshold function based on their common valuation and sensitive degree of waiting. The consumers funding in the crowdfunding will gain the product at the beginning of the second period. In presence of review system, the consumers of the first period will post their review on the platform; the rest of consumers update their belief for the product by the reviews. Thus, the rest consumers learn more about the product's unobservable attribute and their uncertainty of the product is mitigated, while, in absence of the review system, the consumers in the second period are still confused about the product's attribute. To visualize the impact of the review system, we compare the profits in Figure 3.

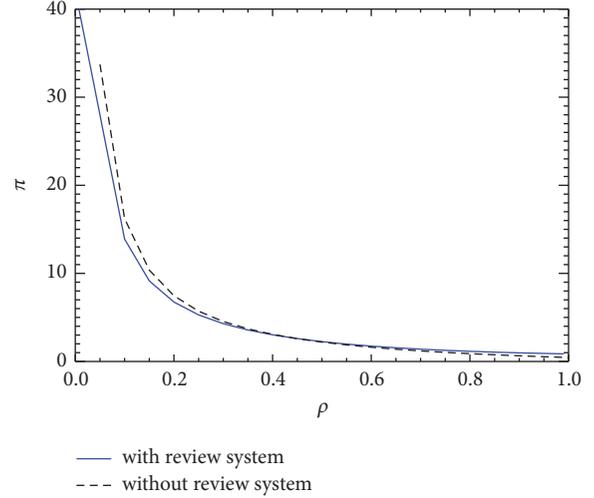


FIGURE 3: Comparison under preannounced pricing.

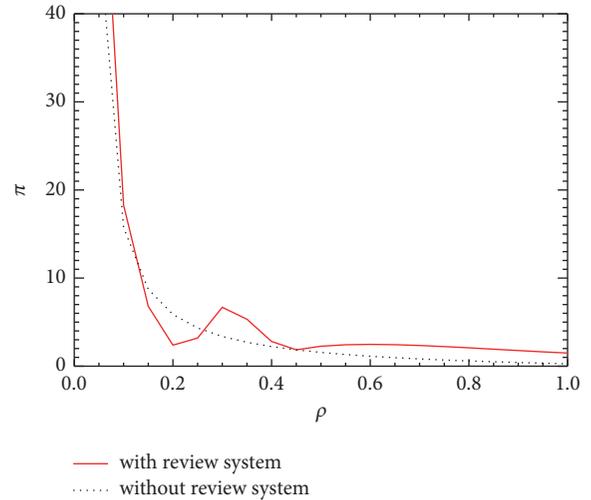


FIGURE 4: Comparison under responsive pricing.

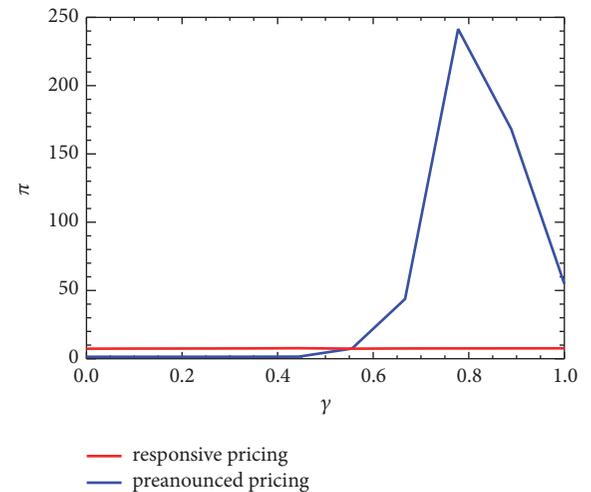


FIGURE 5: Preannounced pricing versus responsive pricing 1.

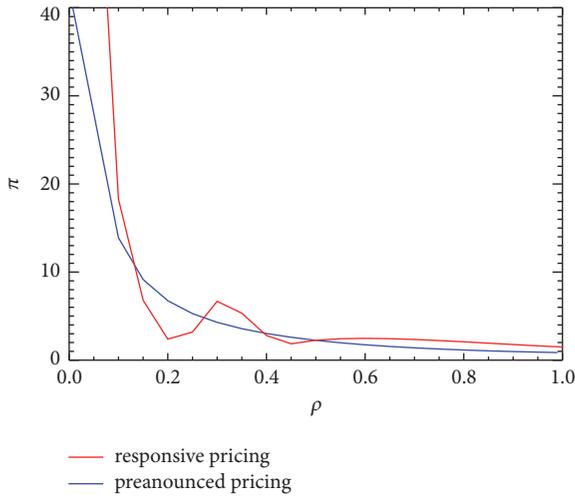


FIGURE 6: Preannounced pricing versus responsive pricing 2.

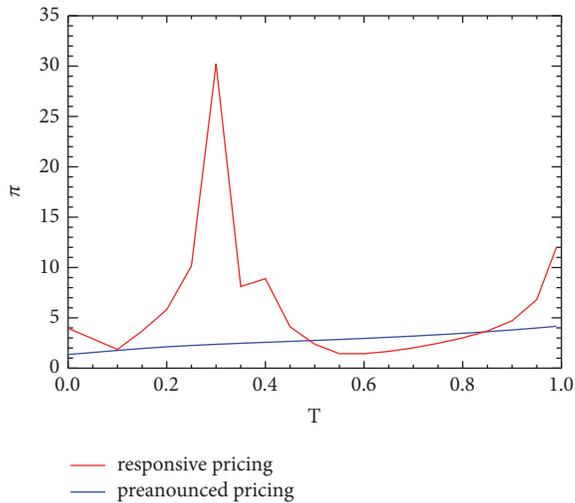


FIGURE 7: Preannounced pricing versus responsive pricing 3.

Under the preannounced policy, as shown in the Figure 3, the advantage of the review system is not obvious.  $\rho = e^\alpha$  is small, which means that the consumers are not sensitive about waiting, and the valuation is higher than large  $\rho$ . That is why the profit is decreased by the increasing  $\rho$ . When  $\rho$  is small, as shown in Figure 3, the preannounced pricing without review system makes more profit. In absence of review system, the consumers cannot learn more information about the product. When the sensitive degree coefficient  $\rho$  is small, the utility difference of purchasing in the first and second period is small, while, in presence of review system, if  $\rho$  is small, the consumers in the second period will learn more information about the product and the valuation in the second period is higher than in the first period. Thus, with the review system, consumers tend to purchase in the second period if they are not sensitive for waiting, which may lead to failure in the crowdfunding. When the consumers are sensitive to waiting, the profit of creator will be lower either with the review system or without review system. But, when

$\rho$  is larger enough, the valuation in both periods is lower, but the profit in review system is a little larger, because the valuation of the product in the second period is larger with the review system. In summary, under preannounced pricing, the impact of the review system is not significant. Whether to disclose the review system on the crowdfunding platform mainly depends on the sensitive degree of waiting.

(2) *Under the Responsive Pricing.* Under the responsive pricing, creator announces the first period price at the beginning of the first period and announces the second period price based on the demand in the first period or review rating level at the beginning of the second period. With the review system, the consumers of the first period post their review online and the rest of the consumers get more information about the product and the creator learn more about the valuation of the consumers. Thus, both consumers and creator can respond to the reviews online under responsive pricing policy. The impact of the review system is shown in Figure 4.

Observe that the review system has superiority under the responsive pricing policy. As shown in Figure 4 the profit is larger with the review system than without review system when the sensitive degree of waiting is large or small. If the consumers are not sensitive for waiting, the creator can gain considerable profits in both cases; and here the advantage of review system is not obvious. When  $\rho$  is approximately 0.2, the policy without review system will make more profit for the creator. If the sensitive degree is larger, the profit of the creator will be decreased, but the policy with the review system is better than without it. Generally speaking, under the responsive pricing policy, the review system has a positive impact on improving the creator's profit.

However, most of the crowdfunding platform still only provide the single selling period for the product. The crowdfunding platform always takes a certain percent of the funds from the creator as the fee. Thus, the more profit creator will get the more platform will get. According to the results, providing two periods and disclosing the review system for the crowdfunding product can improve the profit for the creator. We suggest that the platform ought to provide the second period selling channel for the creator, which can also provide a platform for the consumers to follow the crowdfunding product in case those consumers regret no funding in the first period. At the same time, the first period consumers have an effective way to feedback their experience and the second period consumers are familiar with the product through those previous reviews. Besides, the review system helps the creators better satisfy consumers' preference, which benefits both consumers and creators.

(3) *Pricing Policy Comparison.* Preannounced pricing and responsive pricing have their own advantages. It is necessary to adopt different pricing policies in the different social environment, which will make more profit for the creator and platform. To analyze the superiority of each policy with the review system, a numerical study is required. Next, we discuss the impacts of the review rating level and the sensitive degree of waiting, respectively.

(a) *The Review Rating Level.* As mentioned above, we set  $\gamma$  as the fraction of favorable review which is a standard to measure the level of review ratings. The fraction of favorable review reflects not only the true quality of the product but also the consumer's valuation distribution through the product. According to the previous analysis, consumers' review follows a binomial distribution, and the posterior distribution of valuation is the Beta distribution. The profit of different policies with increasing  $\gamma$  is shown in Figure 5.

Observe that the profit of responsive pricing policy almost makes no difference with increasing favorable review. If the fraction of favorable review is small, responsive pricing seems to be a better choice, but if  $\gamma$  is larger, preannounced pricing is dominant. As shown in Figure 5, when the product is satisfied with the consumers' preference, the profit of preannounced pricing is dozens of times than responsive pricing. This result is consistent with previous studies (e.g., Papanastasiou et al. [36] and Aviv and Pazgal [29]). The creator's profit under the preannounced pricing policy increase with the increasing  $\gamma$ , especially, when  $\gamma$  is about 0.8, is the most. As Vogt and Fesenmaier's study [40] showed, the positive reviews have higher positive influence than negative reviews. However, the profit of preannounced pricing policy does not increase monotonically with  $\gamma$ ; when  $\gamma$  is too high, the profit begins to decrease. Park and Nicolau [37], Kahneman and Tversky [41], and Hu, Pavlou, and Zhang [42] suggest that rational consumers tend to focus on negative reviews more seriously. While positive reviews satisfy people's enjoyment, they make little impact on usefulness. Additionally, it is important to recognize that when online reviews are not helpful for consumers to make an online purchase decision, individuals would not be likely to revisit the website [37]. Those studies are consistent with our result and show that the fraction of favorable review is not the higher the better.

(b) *The Sensitive Degree of Waiting.* With the increasing of consumers sensitive degree of waiting, the profit of the creator comes down. In presence of review system, two policies have their own advantages in different cases. The result of equilibrium analysis is shown in Figure 6.

In presence of review system, when the consumers are extremely sensitive or little sensitive about waiting, the responsive pricing is dominant. When  $\rho$  is about 0.2 and 0.45, the preannounced pricing is better. Responsive pricing policy can set the second period price in response to the consumers' valuation, demand, and review levels. Generally speaking, with the impact of  $\rho$ , the responsive pricing has more advantages, but it is not obvious.

(c) *The Length of Funding Time ( $T$ ).* In this part, we discuss the profit impacted by the length funding time (the length of the first period  $T$ ) under preannounced pricing and responsive pricing, respectively. If the funding time is too short, the crowdfunding will fail because of lacking funders; if the funding time is too long, the total profit may be lower because the first period price is less than the second period price and the consumers purchasing in the first period has to wait a long time to receive the product. Figure 7 shows the impacts of  $T$  in different pricing policies.

Obviously, the length of funding time has fewer impacts on the preannounced pricing; the curve is gentle. However, the curve of responsive pricing is volatile; when  $T$  is about 0.3, the responsive pricing is a dominant strategy. If the length of funding time in the range of [0.5, 0.85], the preannounced pricing can make more profit.

## 7. Conclusion

In recent years, crowdfunding has been regarded as a sustainable and viable alternative channel for entrepreneurs to fund their early stage businesses. However, the uncertainty of the crowdfunding product leads to the lower success rate of crowdfunding and the lower profit of the creator. Hofstede [43] shows that people feel anxious and threatened when they are faced with uncertainty and they rely on their shared beliefs about the current situation. Thus, it is necessary to disclose a review system on the crowdfunding platform so that the latter consumers will mitigate the uncertainty. In this paper, we discuss the impact of the review system and explore the optimal pricing decision of the reward-based crowdfunding. We propose the Bayesian analysis to describe the consumers' valuations update process with the review system. Besides, in order to investigate the performance of the review system, the model without the review system is provided as the benchmark. What is more, the preannounced pricing and the responsive pricing are compared in our model. Thanks to the equilibrium analysis, the following results are discovered. First, under the preannounced pricing policy, the review system does not have the obvious effect on the creator's profit. When the sensitive degree of waiting is low, the strategy without review system makes more profit. Second, under the responsive pricing policy, the strategy with the review system has a dominant advantage in most situations. Third, under the responsive pricing, the impact of the fraction of favorable review is not obvious and the curve of responsive pricing is gentle, while the curve of preannounced pricing is volatile. Finally, the length of funding time has some effect on the responsive pricing; when the cycle time ( $T$ ) is at the range of 0.25 to 0.35 of the total season, the profit is maximum.

According to the conclusion above, we provide the following suggestions. Firstly, the creator should estimate the product's attributes accurately and do some investigation about the crowdfunding product, which can help to understand the consumers' valuation, demand, and sensitive degree of waiting. Next, generally speaking, the review system has a positive impact on improving the creator's profit. Thus, the crowdfunding platform can gain more profit and hence the consumers' loyalty to the platform through the review system as well. We suggest that the platform should disclose an online review system to provide more utilities for consumers, creators, and itself as well. Besides, the consumers will mitigate their uncertainty of the crowdfunding product. Finally, the second period selling channel should be opened on the crowdfunding platform, which can provide a way for the consumers not purchasing in the first period to follow the product and help the rest of the consumers to obtain first-hand reviews from the platform.

However, there are some limitations to this study. Firstly, we only take the strategic consumers into account, but the real market is mixed with strategic and myopic consumers. Secondly, the review rating is diversification, but we only follow Park and Nicolau [37]'s research to consider the extreme review ratings. Finally, in order to simplify the model, the interval between the first and second period is ignored. Those limitations provide some direction for further researches and the model of review system can be used in other areas.

## Appendix

See Theorem 1.

*Proof of Theorem 1.* The consumer  $k$  will take part in the crowdfunding or purchase the product if  $V_k \geq \chi(t)$  in both periods. In addition, if the consumer visits the platform before time  $T$  and  $V_k < \chi(t)$ , he will revisit the platform at time  $T$  and purchases the product if the valuation is larger than the threshold at this time. However, in the pre-announced pricing policy, the strategic consumers arriving at  $[0, T]$  will weigh their valuation and threshold according to the fixed price path  $\{p_1, p_2\}$ .

The consumer's surplus in period 1  $s_1$  and in period 2  $s_2$  is

$$\begin{aligned} s_1 &= \mu V_k e^{-\alpha(T-t)} - p_1 \\ s_2 &= V_k e^{-\alpha(T-t)} - p_2 \end{aligned} \quad (\text{A.1})$$

At time  $T$  ( $0 < T < 1$ ), the condition consumers take part in the crowdfunding in the first period should be satisfied:

$$\begin{aligned} 0 &< t < T, \\ s_1 &\geq s_2. \end{aligned} \quad (\text{A.2})$$

So,  $V_k \geq ((p_1 - p_2)/(\mu - 1))e^{\alpha(T-t)}$ . Similarly, we obtain that  $\chi(t) = p_2$  when  $T < t < H$ .  $\square$

See Proposition 2.

*Proof of Proposition 2.*

$$n_1^{pb} = \lambda * T * \int_0^T \frac{p_1 - p_2}{(\mu - 1)} * e^{\alpha(T-t)} dt$$

$$\begin{aligned} n_2^{pb} &= \lambda * T * \left( \int_0^T p_2 * e^{\alpha(T-t)} dt \right) - n_1^{pb} \\ n_3^{pb} &= \lambda * p_2 * (H - T) \\ \pi^{pb} &= (p_1 - c) n_1^{pb} + (p_2 - c) (n_2^{pb} + n_3^{pb}) \end{aligned} \quad (\text{A.3})$$

Let the first-order condition of  $\pi^{pb}$  be 0.  $\partial \pi^{pb} / \partial p_1 = \partial \pi^{pb} / \partial p_2 = 0$ .

$$\begin{aligned} p_1 &= c \\ p_2 &= \frac{c(T - e^{T\alpha}T - H\alpha + T\alpha + T\mu - e^{T\alpha}T\mu + H\alpha\mu - T\alpha\mu)}{2(-H\alpha + T\alpha + T\mu - e^{T\alpha}T\mu + H\alpha\mu - T\alpha\mu)} \end{aligned} \quad (\text{A.4})$$

$\square$

See Proposition 5.

*Proof of Proposition 5.*

$$\begin{aligned} \pi_2^{rb} &= (p_2 - c)(n_2^{rb} + n_3^{rb}) \\ &= \left( p_2 - \frac{p_1 - p_2}{\mu - 1} \right) * \int_0^T e^{\alpha(T-t)} dt + p_2 * (H - T) \\ \partial_{p_2}(\pi_2^{rb}) &= 0 \\ p_2 &= \frac{-p_1 + e^{T\alpha}p_1 - cH\alpha + cT\alpha - c\mu + ce^{T\alpha}\mu + cH\alpha\mu - cT\alpha\mu}{2(-H\alpha + T\alpha - \mu + e^{T\alpha}\mu + H\alpha\mu - T\alpha\mu)} \quad (\text{A.5}) \\ v_1 &= \frac{p_1 - p_2}{\mu - 1} e^{\alpha(T-t)} \\ n_1^{rb} &= \int_0^T v_1 dt \\ \pi^{rb} &= (p_1 - c)n_1^{rb} + \pi_2^{rb} \end{aligned}$$

Let

$$\partial_{p_1} \pi^{rb} = 0 \quad (\text{A.6})$$

We obtain that

$$\begin{aligned} p_1^* &= \frac{c(1 + 2T\alpha + 2H\alpha(-1 + \mu) - 2\mu - 2T\alpha\mu + e^{T\alpha}(-1 + 2\mu))}{3 + 4T\alpha + 4H\alpha(-1 + \mu) - 4\mu - 4T\alpha\mu + e^{T\alpha}(-3 + 4\mu)} \\ p_2^* &= \frac{1}{6}c \left( 3 + \frac{1}{\mu} + \frac{2}{-3 + 4\mu} - \frac{(H - T)\alpha(-1 + \mu)}{\mu((H - T)\alpha(-1 + \mu) + (-1 + e^{T\alpha})\mu)} \right. \\ &\quad \left. - \frac{8(H - T)\alpha(-1 + \mu)}{(-3 + 4\mu)(3 + 4(H - T)\alpha(-1 + \mu) - 4\mu + e^{T\alpha}(-3 + 4\mu))} \right) \end{aligned} \quad (\text{A.7})$$

$\square$

## Data Availability

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

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors gratefully acknowledge the support from the National Science Foundation of China through Grant 71571117 and Human and Social Science Foundation of Education Committee of China through Grant no. 18YJA630143.

## References

- [1] E. R. Mollick and V. Kuppaswamy, *After the campaign: Outcomes of crowdfunding*, 2014, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2376997](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376997).
- [2] D. Besanko and W. L. Winston, "Optimal price skimming by a monopolist facing rational consumers," *Management Science*, vol. 36, no. 5, pp. 555–567, 1990.
- [3] M. Hu, J. Milner, and J. Wu, "Liking and following and the newsvendor: Operations and marketing policies under social influence," *Management Science*, vol. 62, no. 3, pp. 867–879, 2016.
- [4] S. M. Freedman and G. Z. Jin, "Learning by Doing with Asymmetric Information," in *evidence from Prosper.com (No. w16855)*. National Bureau of Economic Research, Learning by Doing with Asymmetric Information, evidence from Prosper.com (No. w16855). National Bureau of Economic Research, 2011.
- [5] S. Chakraborty and R. Swinney, "Signaling to the crowd: private quality information and rewards-based crowdfunding," *Social Science Electronic Publishing*, p. 44, 2016.
- [6] K. De Buysere, O. Gajda, R. Kleverlaan, and D. Marom, *A Framework for European Crowdfunding*, from <http://Crowdfundingnetwork.edu>, 2012.
- [7] L. Du, M. Hu, and J. Wu, "Contingent stimulus in crowdfunding," [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2925962](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2925962).
- [8] Y. Chen, R. Zhang, and B. Liu, "Joint Decisions on Production and Pricing with Strategic Consumers for Green Crowdfunding Products," *International Journal of Environmental Research and Public Health*, vol. 14, no. 9, 2017.
- [9] V. F. Araman and R. Caldentey, SSRN, 2016, <https://ssrn.com/abstract=272351>.
- [10] S. Marinesi and K. Girotra, *Information acquisition through customer voting systems*, 2013, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2191940](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2191940).
- [11] M. Wessel, F. Thies, and A. Benlian, "The emergence and effects of fake social information: Evidence from crowdfunding," *Decision Support Systems*, vol. 90, pp. 75–85, 2016.
- [12] F. Xu, X. Guo, G. Xiao, and F. Zhang, "Crowdfunding or Bank Financing: Effects of Market Uncertainty and Word-of-Mouth Communication," 2018 <https://ssrn.com/abstract=3209835> or <http://dx.doi.org/10.2139/ssrn.3209835>.
- [13] A. V. Banerjee, "A simple model of heed behavior," *The Quarterly Journal of Economics*, vol. 107, no. 3, pp. 797–817, 1992.
- [14] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of Political Economy*, vol. 100, no. 5, pp. 992–1026, 1992.
- [15] G. Ellison and D. Fudenberg, "Rules of thumb for social learning," *Journal of Political Economy*, vol. 101, no. 4, pp. 612–643, 1993.
- [16] G. Ellison and D. Fudenberg, "Word-of-mouth communication and social learning," *The Quarterly Journal of Economics*, vol. 110, no. 1, pp. 93–125, 1995.
- [17] Y. Kwark, C. Jianqing, and S. Raghunathan, "Online product reviews: Implications for retailers and competing manufacturers," *Information Systems Research*, vol. 25, no. 1, pp. 93–110, 2014.
- [18] P. Roma, E. Gal-Or, and R. R. Chen, "Reward-Based Crowdfunding Campaigns: Informational Value and Access to Venture Capital," in *Information Systems Research*, vol. 29, pp. 525–777, 3 edition, 2018.
- [19] D. Crapis, B. Ifrach, C. Maglaras, and M. Scarsini, "Monopoly pricing in the presence of social learning," *Management Science*, vol. 63, no. 11, pp. 3586–3608, 2017.
- [20] B. Jing, "Social learning and dynamic pricing of durable goods," *Marketing Science*, vol. 30, no. 5, pp. 851–865, 2011.
- [21] Y. Jiang and H. Guo, "Design of consumer review systems and product pricing," *Information Systems Research*, vol. 26, no. 4, pp. 714–730, 2015.
- [22] Y. Chen and J. Xie, "Third-party product review and firm marketing strategy," *Marketing Science*, vol. 24, no. 2, pp. 218–240, 2005.
- [23] B. Jiang, K. Jerath, and K. Srinivasan, "Firm strategies in the 'mid tail' of platform-based retailing," *Marketing Science*, vol. 30, no. 5, pp. 757–775, 2011.
- [24] D. Kuksov and Y. Xie, "Pricing, frills, and customer ratings," *Marketing Science*, vol. 29, no. 5, pp. 925–943, 2010.
- [25] B. Ifrach, C. Maglaras, and M. Scarsini, "Bayesian social learning with consumer reviews," p. 28.
- [26] Y. Papanastasiou, N. Bakshi, and N. Savva, *Scarcity strategies under quasi-Bayesian social learning*, 2014, [http://www.yiangos-papanastasiou.com/StockOuts\\_2016.pdf](http://www.yiangos-papanastasiou.com/StockOuts_2016.pdf).
- [27] N. L. Stokey, "Intertemporal price discrimination," *The Quarterly Journal of Economics*, vol. 93, no. 3, pp. 355–371, 1979.
- [28] M. Landsberger and I. Meilijson, "Intertemporal price discrimination and sales strategy under incomplete information," *The RAND Journal of Economics*, pp. 424–430, 1985.
- [29] Y. Aviv and A. Pazgal, "Optimal pricing of seasonal products in the presence of forward-looking consumers," *Manufacturing & Service Operations Management*, vol. 10, no. 3, pp. 339–359, 2008.
- [30] W. Elmaghraby and P. Keskinocak, "Dynamic pricing in the presence of inventory considerations: research overview, current practices, and future directions," *Management Science*, vol. 49, no. 10, pp. 1287–1309, 2003.
- [31] G. P. Cachon and R. Swinney, "Purchasing, pricing, and quick response in the presence of strategic consumers," *Management Science*, vol. 55, no. 3, pp. 497–511, 2009.
- [32] R. Yin, Y. Aviv, A. Pazgal, and C. S. Tang, "Optimal markdown pricing: implications of inventory display formats in the presence of strategic customers," *Management Science*, vol. 55, no. 8, pp. 1391–1408, 2009.
- [33] S. Whang, "Demand uncertainty and the Bayesian effect in markdown pricing with strategic customers," *Manufacturing & Service Operations Management*, vol. 17, no. 1, pp. 66–77, 2015.

- [34] D. Bergemann and J. Välimäki, "Market diffusion with two-sided learning," *The RAND Journal of Economics*, vol. 28, no. 4, pp. 773–795, 1997.
- [35] N. Vettas, "On the informational role of quantities: Durable goods and consumers' word-of-mouth communication," *International Economic Review*, vol. 38, no. 4, pp. 915–944, 1997.
- [36] Y. Papanastasiou, K. Bimpikis, and N. Savva, "Crowdsourcing exploration," *Management Science*, vol. 64, no. 4, pp. 1477–1973, 2017.
- [37] S. Park and J. L. Nicolau, "Asymmetric effects of online consumer reviews," *Annals of Tourism Research*, vol. 50, pp. 67–83, 2015.
- [38] J. Correa, R. Montoya, and C. Thraves, "Contingent preannounced pricing policies with strategic consumers," *Operations Research*, vol. 64, no. 1, pp. 251–272, 2016.
- [39] A. Smith, "Civic engagement in the digital age," *Pew Research Center*, vol. 25, pp. 307–332, 2013.
- [40] C. A. Vogt and D. R. Fesenmaier, "Expanding the functional information search model," *Annals of Tourism Research*, vol. 25, no. 3, pp. 551–578, 1998.
- [41] D. Kahneman and A. Tversky, "Prospect theory: an analysis of decision under risk," *Econometrica*, vol. 47, no. 2, pp. 263–292, 1979.
- [42] N. Hu, P. A. Pavlou, and J. J. Zhang, "Why do online product reviews have a J-shaped distribution? Overcoming biases in online word-of-mouth communication," *Communications of the ACM*, vol. 52, no. 10, pp. 144–147, 2009.
- [43] G. Hofstede, "Culture's consequences: International differences in work-related values," 1980, <http://ci.nii.ac.jp/ncid/BA07769656>.

## Research Article

# Research on Supply Chain Coordination Based on Block Chain Technology and Customer Random Demand

Yongfei Li,<sup>1</sup> Bill Wang,<sup>2</sup> and Dong Yang<sup>3</sup> 

<sup>1</sup>*School of Modern Post, Xi'an University of Posts & Telecommunications, China*

<sup>2</sup>*Business Information Systems Department, Auckland University of Technology, New Zealand*

<sup>3</sup>*Glorious Sun School of Business and Management, Donghua University, China*

Correspondence should be addressed to Dong Yang; yangdong@dhu.edu.cn

Received 4 November 2018; Accepted 1 January 2019; Published 20 January 2019

Academic Editor: Lu Zhen

Copyright © 2019 Yongfei 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.

Based on disruptive innovation and Stackelberg noncooperative game theory, the paper focuses on supply chain coordination under the combined effects of block chain technology and random demand. Firstly, both a decentralized and a centralized supply chain decision models are built in a single-cycle newsvendor random demand situation. Then, through revenue sharing contract the study designs a brand-new supply chain coordination model which is Del trust, decentralized, and traded anonymously. Furthermore, the numerical comparative analysis on the optimal decision and supply chain coordination are conducted. It is found that the whole supply chain revenue can achieve and even beyond the performance level of the centralized supply chain with effectively expanding sales market and reducing supply chain risk. When the retail price is stable and supply chain is coordinated with revenue sharing mechanism, decentralized supply chain can achieve minimum optimal revenue. Coordination results have effect on short-term revenues of block chain members only. Implications and suggestions for future research in supply chain coordination are provided.

## 1. Introduction

Because of the uncertainty of competition and changing of the market, customer demand is becoming more and more random, which must be highly paid attention to and valued as one significant factor by supply chain members [1–4]. Random demand can lead to the difficulty of supply chain coordination and risk management among supply chain members; therefore, how to effectively manage the random demand has become a core challenge for supply chain decision-makers and researchers to consider [2–6]. Among random demand problems, this paper focuses on the supply chain coordination between a single manufacturer and single retailer under single cycle newsvendor random demand model. The two parties have developed a certain degree of trust mechanism while the core, centralized enterprise has a leading position [3–7]. However, the existence of these trust mechanism and centralized enterprises may generate a variety of difficulties including higher cost, lower efficiency, poor chronology, unsafe data storage, and lower robustness of

supply chain. These create new challenges to current supply chain techniques and coordination theory [8–11].

Block chain technology possesses the technical features including decentralization, distrust, agent elimination, chronology, anonymity, group maintenance, opening-sourcing, being programmable, dispersed, unchangeable encrypted data, safety, and reliability. Moreover, it is hopeful that block chain technology can remodel and change the current forms of human social activities thoroughly, and the technology has been implemented and achieved positive practice and application in many fields including banking, finance, security, insurance, express, notarization, medicine, music, crowd funding, dispersed data, intelligent manufacturing, and the Internet of Things [8–11]. Decentralized trust based point-to-point transaction, cooperation, and coordination can be achieved by block chain technology with the methods of encrypting data, time stamping, distributed consensus, and economic incentives. Thus, information chain can be altered to value chain in order to offer a feasible and coordinative way to address the current and universal supply

chain issues including high cost by centralized organizations, low efficiency, unsafely data storage, and poor robustness [8–11].

Furthermore, in the real practice, supply chains are mostly decentralized supply chains which may lead to double marginalization, bullwhip effect, and other negative effects; thus a rational supply chain coordination mechanism needs to be built to bond all involved members [7, 12–15]. Revenue sharing contract can bond the above parties together and help decentralized supply chains achieve the performance of centralized supply chains [3, 7, 13]. Especially, with the support of block chain technique, commodity pricing power belongs to none of the parties, so revenue sharing contract needs to be implemented to coordinate the order and wholesale price, expand market share, and enhance the advantages of block chain. Therefore, in order to improve whole profit of supply chain and reduce risks of high cost, low efficiency, poor chorology, and unsafely data storage which resulted by trust mechanism and centralized enterprises, application of Revenue sharing contract, and coordination issues of interests relationship among supply chain members under single cycle newsvendor random demand model, based on consideration of synergy brought by block chain technique and random demand, should be more valuable for scholars to focus on and study in future.

The remainder of the paper is organized as follows. Following the literature review in Section 2, the research model is introduced in Section 3. Section 4 discusses the supply chain coordination strategy with calculation for manufacturer and retailer based on block chain concept. After the analysis and comparison of relevant statistics from the models in Section 5, our interpretations of the results and research conclusion are presented in Section 6.

## 2. Literature Review

Pasternack applied Return Contract to study supply chain coordination problem of single cycle newsvendor random demand model facing by suppliers and dealers for the first time [7]. Following that, researches on this issue were conducted by many scholars. Cachon and Lariviere addressed the conditions of contract coordination through comparison and analysis of revenue sharing contract and price discount, quantity discount, and Return Contract [1]. They also stated that in single cycle newsvendor random demand model formed by single supplier and single buyer and other situation, Revenue Contract can achieve not only the Pareto optimum, but also the random revenue distribution of supply chain. He [9] investigated a supply chain channel coordination problem under stochastic demand and both sales effort and retail price sensitive with standard newsvendor setting. Chen and Bell adopted a two buyback prices' agreement which were unsold inventory and customer returns between a manufacturer and a retailer in a decentralized supply chain where the retailer simultaneously determines the retail price and order quantity while experiencing customer returns and price dependent stochastic demand. They found the two buyback prices' agreement could achieve perfect supply

chain coordination and win–win for both manufacturers and retailers [10].

Li and Su studied the coordination problem of competitive disadvantage supply chain under the situation of random demand and revocable strategy [2]. Govindana and Popiuc researched a two- and three-echelon reverse supply chain coordination by revenue sharing contract to explore the implications of recycling from an efficiency perspective for all participants in the process [11]. In their mode, the customer was willing to return obsolete units as a function of the discount offered by the retailer in exchange for recycling devices with a remanufacturing value. The results showed that performance measures and total supply chain profits improved through coordination.

Zhao and Zhu studied a revenue-sharing mechanism to examine how to coordinate a remanufacturing supply chain between a remanufacturer and a retailer considering stochastic remanufacturability rate and random demand [12]. The research results showed that a revenue-sharing contract can increase profit for the remanufacturer, the retailer, and the whole supply chain. Their case study demonstrated that the government subsidy could benefit from the proposed revenue-sharing mechanism and the profit increase for the whole supply chain.

Giri and Chakraborty coordinated a supply chain considering instantaneous/noninstantaneous supplies from single-vendor to single-buyer whose replenishment time interval was not dependent upon the input with stochastic demand and uncertain yield [13]. They further extended to consider a revenue sharing contract to align incentives of decentralized partners within the supply chain. The coordinated policy was beneficial in comparison to the individual policies made by the supply chain partners.

Furthermore, Stackelberg noncooperative gaming is applied to study the supply chain coordination and risks sharing issues formed by single supplier and single dealer [2, 14–16].

However, these researches are related to the supply chain coordination problem of single cycle newsboy random demand model in which certain degree's trust mechanism is possessed in the supply chain and one of the parties is core enterprise in the supply chain performing the leading role. And the existence of trust mechanism and centralized enterprises may usually lead to multiple issues for the supply chain including high cost, low efficiency, poor chorology, unsafely data storage, and poor robustness and other problems.

In the area of block chain, Christidis and Devetsikiotis stated that a block chain could enable trustless networks and worked as a distributed data structure which was replicated and shared among the members of a network [17]. Block chains could solve the double-spending problem which was similar to the bilateral effect in a supply chain. Khandelwal et al. claimed that a block chain was a distributed ledger of transactions which were maintained by a network of trustless nodes [18]. Block chains could previously run only through a trusted intermediary, operating in a decentralized fashion, without the need for a central authority while achieving the same functionality with the same amount of certainty [19–21].

Based on the literatures above, block chain technology and interruptive innovation theory, in the situation of random demand and rebuy strategy, can be combined together. Our study applies Stackelberg noncooperative gaming theory and revenue sharing contract into analyzing the brand-new coordination mechanism of supply chain under single cycle newsboy random demand model with the features of Del trust, decentralization, and traded anonymous.

In the current research, we mainly contribute to the extant literature in the following two ways. First, the coordination problem in the situation of customer random demand is necessary but the research in this topic is still scant. Most extant papers are related to the trust mechanism to certain degree possessed by all parties in the supply chain with real-name registered trade with one party as the core enterprise and holding the dominant position. Our study focuses on supply chain coordination problems related to Del trust, decentralization, and traded anonymous. This innovative perspective can enrich and supplement the previous study. Second, most previous studies with the results for coordination have more influences on short-term revenue of other members in the supply chain and lead to poor robustness. Our study combines block chain technology and random demand situation while the coordination result has few influences on the short-term revenue of other members of block chain (other members); thus our research can improve the robustness of supply chain significantly.

### 3. Model Description

Assuming there are multiple manufacturers  $M$  ( $M_1 M_2 \dots M_n$ ) and multiple retailers  $R$  ( $R_1 R_2 \dots R_m$ ) in the supply chain, in which supply chain members are dispersed in areas and regions, no core enterprises within the network, the position, power, and responsibility of enterprises at every level are equal; any damage or loss of the enterprises of any level has no effect on the overall system operations, namely, decentralized and distributed management phenomenon with robustness. And no trust requirement for data exchanging among levels in the supply chain and all the operation regulations and data are transparent and visible to all the members. Therefore, enterprises at different levels cannot and are not capable of cheating others within the regulated rules and time. Because there is no trust is required among enterprises at different levels, all the trades do not need to disclose identity; namely, all the enterprises involved are anonymous. Every retailer  $R$  in the supply chain will face the random demand sales market of single cycle newsboy and will start to place orders before the selling season of each single cycle.

The stochastic demand model is mainly embodied by the one-cycle newsvendor model used by predecessors. In the newsvendor model, the purchasing quantity is determined by the demand, and the demand quantity is a random variable  $X$  [6, 22].

The distribution function of the stochastic demand variable  $X$  is a function  $F(x) = P(X < x)$ , and the distribution of the stochastic demand variable function means that if  $X$  is a stochastic variable, then  $Y = g(X)$  is also a stochastic variable,

and the distribution law of  $Y$  is the distribution of the function of the stochastic demand variable  $X$ .

The impact from the current decisions of retailer  $R$  on short future can be ignored and one-cycle schedule is only needed for one time. Additionally, since all the products will be delivered before demand, all inventory that is either ordered or manufactured can be used to satisfy demand directly with no waiting. Assuming  $X$  as the market and customer demand variation of the retailer  $R$  within the sales season, and its Probability Distribution Function and Cumulative Distribution Function are  $f(x)$  and  $F(x) = P(X \leq x)$ ,  $F(\cdot)$  is continuous differentiable and strictly increasing,  $F(0) = 0$ ,  $\bar{F}(x) = 1 - F(x)$ , the demand average value of customer demand variation  $X$  is  $u = E(x) = \int_0^{\infty} xf(x)dx$ , and set  $\omega$  is the unified wholesaling price of manufacturer in the supply chain,  $p$  is the unified sales price of retailer in the supply chain, and both will remain unchangeable in the short-term.

Lack of trust between supply chain members leads to low competitiveness of supply chain. To satisfy the survival and development need of the members themselves and the whole supply chain and to improve the overall competitiveness of the supply chain, the contract for trading should be signed by supply chain members which contains honest trade contact with high cost of breaking contract and confidentiality breach, agreement of applying block chain technology, and related ideas. Namely, all the trades can only be transacted after the implementation of one-off random orders which will be opened to all the supply chain members involved through the recordable and traceable internal sharing platform in the form of certain key and code, reaching the consensus of all the members.

The key means to ensure the safety of trades and code can be helpful to the result that other members in the supply chain can see the trade but cannot know which are the two parties involved, and even if the vendor and purchaser only know the information about itself but no information about the other side. Tracing will be activated by certain one-off decrypted method which will be applied after the internal sharing platform received the application from the breached party with the code, identifying code, and other methods. After it is activated, all the members can view the detail trade information and the event of default of the two parties and deal with the event according to contract rules. To ensure the safety and privacy of the trade among members, only the information about breach of contract applied will be decrypted, instead of other information.

Once the trade is settled, it cannot be denied and regretted; otherwise, high penalty of which the loss will overweight the gain must be paid. Namely, the penalty will be multiple times compared to the trade amount which is too high to break the contract for both sides, so there is few contact breaking issues. And all trades must be kept secret by all supply chain members; otherwise, high penalty is required to be paid. Namely, any member cannot give away the secret, so there are few secret divulging issues.

Assuming, in the supply chain, a retailer  $R1$  needs to purchase some products from manufacturer  $M1$  now,  $Q1$  is

the order amount of retailer R1, and the expected production of manufacturer M1 which refers to M1 expects to manufacture products according to the order amount from R1. The ordered products should be open to other members  $M_2 \dots M_n, R_2 \dots R_m$  in the supply chain, and consensus needs to be reached by these members. Once the trade is settled, it cannot be denied; otherwise, the high penalty must be paid. Assume  $g_r$  and  $g_m$  refer to the penalty cost that will be paid by the defaulting retailer  $R_1$  and the penalty cost of outstock for the goodwill (penalty of outstock) of manufacturer  $M_1$ , respectively, and  $g = g_r + g_m$ ; because of the none-deniability of the trade between retailer  $R_1$  and manufacturer  $M_1$ , namely,  $g_r$  and  $g_m$  are too big, so both sides will not breach the contract normally. Moreover,  $c_m$  and  $c_r$  refer to the production unit cost for the manufacturer and the marginal cost for the retailer. And  $c_r$  refers to the cost to acquire product unit, instead of the cost during products selling and based on block chain idea  $c_r \rightarrow 0$ ; for the products that have not been sold within the sales season can be returned to manufacture from retailer, with the unit refund price as  $\alpha\omega$  and  $\alpha$  as the refund discount coefficient. In the short term after the trade, the overall supply chain performance will only be influenced by the trade; at the same time, revenue of other members in the supply chain will remain the same in the short-term, assuming as  $\pi(M_2 \dots M_n, R_2 \dots R_m)$ .

According to "bilateral marginal revenue", if retail  $R_1$  places the optimal orders based on its own interest, the the overall optimal revenue cannot be achieved by the order amount. And revenue sharing contract can be one of the feasible methods to address the "bilateral marginal revenue" issues. With the implementation of revenue sharing contract, manufacturer  $M_1$  will sell the products to retailer  $R_1$  with the wholesaling price  $\omega_\phi$  which is lower than its marginal cost  $c_m$ , and retailer  $R_1$  will retain the revenue of expected sales part  $\phi$ ; namely,  $\phi\pi_{r_1}(Q_1, \omega_\phi)$ ,  $\phi \in [0, 1]$ , and the remaining part  $(1-\phi)\pi_{r_1}(Q_1, \omega_\phi)$  will be returned to manufacturer  $M_1$ ,  $Q_1, \omega_\phi$  are the decision variables. In this case, the optimal revenue for R and the whole supply chain can be reached at the same time, and the revenue gained by M and R will not be lower than that before the implementation of revenue sharing contract; therefore, the coordinative status of the whole supply chain can be achieved.

#### 4. Supply Chain Coordination Based on Block Chain Idea

According to the model description, the formulas of expected revenue  $\pi_{r_1}(Q_1)$  of retailer  $R_1$ , the expected revenue  $\pi_{m_1}(Q_1)$  of manufacturer  $M_1$ , and the expected revenue of the whole supply chain  $\pi$  are as follows:

$$\begin{aligned} \pi_{r_1}(Q_1) &= (p - \omega - c_r)Q_1 - g_r \int_{Q_1}^{\infty} (x - Q_1) f(x) dx \\ &\quad - (1 - \alpha)\omega \int_0^{Q_1} (Q_1 - x) f(x) dx \end{aligned} \quad (1)$$

$$\begin{aligned} \pi_{m_1}(Q_1) &= (\omega - c_m)Q_1 - g_m \int_{Q_1}^{\infty} (x - Q_1) f(x) dx \\ &\quad - \alpha\omega \int_0^{Q_1} (Q_1 - x) f(x) dx \end{aligned} \quad (2)$$

$$\begin{aligned} \pi &= \pi(M_2, \dots, M_n, R_2, \dots, R_m) + (p - c)Q_1 \\ &\quad - g \int_{Q_1}^{\infty} (x - Q_1) f(x) dx \\ &\quad - \omega \int_0^{Q_1} (Q_1 - x) f(x) dx \end{aligned} \quad (3)$$

Within that,  $\int_0^{\infty} \{Q_1 - X, 0\}^+ f(x) dx = \int_0^{Q_1} (Q_1 - x) f(x) dx$  refers to retailer  $R_1$ 's expected 的 leftover amount, and  $\int_0^{\infty} \{X - Q_1, 0\}^+ f(x) dx = \int_{Q_1}^{\infty} (x - Q_1) f(x) dx$  refers to retailer  $R_1$ 's anticipated outstock amount.

*4.1. Decentralized Supply Chain Situation.* According to block chain ideas, the status of retailer  $R_1$  and manufacturer  $M_1$  is completely equal. However, because retailer  $R_1$  needs to purchase products from manufacturer  $M_1$ , retailer  $R_1$  is subordinated to manufacturer  $M_1$  in the real practice; namely, manufacture  $M_1$  holds the dominant status while retailer  $R_1$  stays as the passive and subordinate status, which can be seen as Stackelberg gaming problem.

According to Backward Induction,  $\partial^2 \pi_{r_1}(Q_1) / \partial Q_1^2 = -f(Q_1)[g_r + (1 - \alpha)\omega] < 0$ , so  $\pi_{r_1}(Q_1)$  is  $Q_1$ 's convex function, and retailer  $R_1$  has the only optimal order amount. The optimal order amount  $Q_1'$  of retailer  $R_1$  and expected revenue  $E\prod_{r_1}(Q_1')$  are

$$Q_1' = F^{-1} \left[ \frac{(p - \omega - c_r)}{g_r + (1 - \alpha)\omega} \right] \quad (4)$$

$$\begin{aligned} E\prod_{r_1}(Q_1') &= (p - \omega - c_r)Q_1' \\ &\quad - g_r \int_{Q_1'}^{\infty} (x - Q_1') f(x) dx \\ &\quad - (1 - \alpha)\omega \int_0^{Q_1'} (Q_1' - x) f(x) dx \end{aligned} \quad (5)$$

and manufacturer  $M_1$  and expected revenue  $E\prod_{m_1}(Q_1')$  are

$$\begin{aligned} E\prod_{m_1}(Q_1') &= (\omega - c_m)Q_1' \\ &\quad - g_m \int_{Q_1'}^{\infty} (x - Q_1') f(x) dx \\ &\quad - \alpha\omega \int_0^{Q_1'} (Q_1' - x) f(x) dx \end{aligned} \quad (6)$$

The overall expected revenue of the whole supply chain  $E \prod(Q_1')$  is

$$E \prod(Q_1') = \pi(M_2, \dots, M_n, R_2, \dots, R_m) + (p - c)Q_1' - g \int_{Q_1'}^{\infty} (x - Q_1') f(x) dx - \omega \int_0^{Q_1'} (Q_1' - x) f(x) dx \quad (7)$$

4.2. *Supply Chain Coordination.* According to revenue sharing contract, the revenue functions of retailer  $R_1$ , manufacturer  $M_1$ , and the whole supply chain are as follows:

$$\pi_{r_1}(Q_1) = \phi \left[ (p - \omega_\phi - c_r)Q_1 - g_r \int_{Q_1}^{\infty} (x - Q_1) f(x) dx - (1 - \alpha)\omega_\phi \int_0^{Q_1} (Q_1 - x) f(x) dx \right] \quad (8)$$

$$\pi_{m_1}(Q_1) = (1 - \phi) \left[ (p - \omega_\phi - c_r)Q_1 - g_r \int_{Q_1}^{\infty} (x - Q_1) f(x) dx - (1 - \alpha)\omega_\phi \int_0^{Q_1} (Q_1 - x) f(x) dx \right] + (\omega_\phi - c_m) \cdot Q_1 - g_m \int_{Q_1}^{\infty} (x - Q_1) f(x) dx - \alpha\omega_\phi \int_0^{Q_1} (Q_1 - x) f(x) dx \quad (9)$$

$$\pi = \pi(M_2, \dots, M_n, R_2, \dots, R_m) + (p - \omega_\phi - c_r)Q_1 - (1 - \alpha)\omega_\phi \int_0^{Q_1} (Q_1 - x) f(x) dx + (\omega_\phi - c_m)Q_1 - g_m \int_{Q_1}^{\infty} (x - Q_1) f(x) dx - \alpha\omega_\phi \int_0^{Q_1} (Q_1 - x) f(x) dx - g_r \int_{Q_1}^{\infty} (x - Q_1) f(x) dx \quad (10)$$

**Proposition 1.** *Constrained by revenue sharing contract, the coordinative wholesaling price  $\omega_\phi^{**}$  exists between retailer  $R_1$  and manufacturer  $M_1$ .*

*Proof.*  $\partial^2 \pi_{r_1}(Q_1)/\partial Q_1^2 = -\phi f(Q_1)[g_r + (1 - \alpha)\omega_\phi] \leq 0$ , so  $\pi_{r_1}(Q_1)$  is a concave function about  $Q_1$  and retailer  $R_1$  has the only and optimal order amount  $Q_1''$

$$Q_1'' = F^{-1} \left[ \frac{p - \omega_\phi - c_r}{g_r + (1 - \alpha)\omega_\phi} \right] \quad (11)$$

Therefore,  $\partial^2 \pi_{m_1}(Q_1)/\partial Q_1^2 = -2f(Q_1)(g_r + (1 - \alpha)\omega_\phi) + \phi f(Q_1)(g_m + \alpha\omega_\phi) \leq 0$ ; we can know that  $\pi_{m_1}(Q_1)$  is a concave function about  $Q_1$ , and  $Q_1''$ , the optimal sales quantity of  $M_1$ .

So  $\partial^2 \pi_{m_1}(Q_1)/\partial Q_1^2 = -2f(Q_1)(g_r + (1 - \alpha)\omega_\phi) + \phi f(Q_1)(g_m + \alpha\omega_\phi) \leq 0$ :

$$Q_{m_1}'' = F^{-1} \left[ \frac{p - c_m - c_r - \phi(p - \omega_\phi - c_r)}{(1 - \phi)g_r + (1 - \phi - \alpha\phi)\omega_\phi + g_m} \right] \quad (12)$$

Let  $Q_1'' = Q_{m_1}''$ ; then

$$-(1 - 2\alpha\phi)\omega_\phi^2 - (1 - \phi + g_m + g_r\phi - c_m - \alpha p + \alpha c_m + \alpha c_r + 2\alpha\phi p - 2c_r\alpha\phi)\omega_\phi + g_m(p - c_r) + g_r c_m = 0 \quad (13)$$

Let  $a = 2\alpha\phi - 1$ ,  $b = -(1 - \phi + g_m + g_r\phi - c_m - \alpha p + \alpha c_m + \alpha c_r + 2\alpha\phi p - 2c_r\alpha\phi)$ ,  $c = g_m(p - c_r) + g_r c_m$ ; then, constrained by revenue sharing contract, retailer  $R_1$  and manufacturer  $M_1$  must coordinate the wholesale price  $\omega_\phi^{**}$  to satisfy

$$\omega_\phi^{**} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (14)$$

Thus, proposition can be proved.  $\square$

Then the economic order quality is

$$Q_1^{**} = F^{-1} \left[ \frac{p - \omega_\phi^{**} - c_r}{g_r + (1 - \alpha)\omega_\phi^{**}} \right] \quad (15)$$

Constrained by revenue sharing contract, revenue expectation of retailer  $R_1$ , manufacturer  $M_1$ , and the whole supply chain are

$$E \prod_{r_1}(Q_1^{**}) = \phi \left\{ (p - \omega_\phi^{**} - c_r)Q_1^{**} - g_r \int_{Q_1^{**}}^{\infty} (x - Q_1^{**}) f(x) dx - (1 - \alpha)\omega_\phi^{**} \int_0^{Q_1^{**}} (Q_1^{**} - x) f(x) dx \right\} \quad (16)$$

$$E \prod_{m_1} (Q_1^{''*}) = (1 - \phi) \left[ (p - \omega_\phi^{''*} - c_r) Q_1^{''*} - g_r \int_{Q_1^{''*}}^{\infty} (x - Q_1^{''*}) f(x) dx - (1 - \alpha) \omega_\phi^{''*} \int_0^{Q_1^{''*}} (Q_1^{''*} - x) f(x) dx \right] + (\omega_\phi^{''*} - c_m) Q_1^{''*} - g_m \int_{Q_1^{''*}}^{\infty} (x - Q_1^{''*}) f(x) dx - \alpha \omega_\phi^{''*} \int_0^{Q_1^{''*}} (Q_1^{''*} - x) f(x) dx \quad (17)$$

$$E \prod (Q_1^{''*}) = \pi(M_2, \dots, M_n, R_2, \dots, R_m) + (p - c) \cdot Q_1^{''*} - g \int_{Q_1^{''*}}^{\infty} (x - Q_1^{''*}) f(x) dx - \omega_\phi^{''*} \int_0^{Q_1^{''*}} (Q_1^{''*} - x) f(x) dx \quad (18)$$

4.3. *Centralized Supply Chain.* The revenue function of centralized supply chain is

$$\pi(Q_1) = \pi(M_2, \dots, M_n, R_2, \dots, R_m) + (p - c) Q_1 - g \int_{Q_1}^{\infty} (x - Q_1) f(x) dx - \omega \int_0^{Q_1} (Q_1 - x) f(x) dx \quad (19)$$

As  $\partial^2 \pi(Q_1) / \partial Q_1^2 = -f(Q_1)(g + \omega) < 0$ , therefore,  $\pi(Q_1)$  is a concave function about  $Q_1$ . There is an optimal order quantity of system; namely, when  $Q_1^{''''}$ , the optimal order quantity, is

$$Q_1^{''''} = F^{-1} \left[ \frac{p - c}{g + \omega} \right] = F^{-1} \left[ \frac{p - c_m - c_r}{g_m + g_r + \omega} \right] \quad (20)$$

then, the optimal profit of the system is

$$E \prod (Q_1^{''''}) = \pi(M_2, \dots, M_n, R_2, \dots, R_m) + (p - c) Q_1^{''''} - g \int_{Q_1^{''''}}^{\infty} (x - Q_1^{''''}) f(x) dx - \omega \int_0^{Q_1^{''''}} (Q_1^{''''} - x) f(x) dx \quad (21)$$

4.4. *Discussion of Revenue Sharing Factor  $\phi$  Value of Retailer  $R_1$  and Manufacturer  $M_1$ .* Under the situation of distributed supply chain, the optimal decision of retailer  $R_1$  is partial optimal decision instead of the overall optimal decision for the whole supply chain. Under the bond of revenue sharing contract, the main purpose of the coordinative decisions

from retailer  $R_1$  and manufacturer  $M_1$ ,  $Q_1^{''''}$ , and  $\omega_\phi^{''*}$  is to improve the performance of distributed supply chain to the overall performance of the centralized supply chain; namely,  $E \prod (Q_1^{''*}) = E \prod (Q_1^{''''})$ , so,

$$Q_1^{''*} = Q_1^{''''} = F^{-1} \left[ \frac{p - c_m - c_r}{g_m + g_r + \omega} \right] \quad (22)$$

Besides, under the circumstance of no bond from the contract, the decision made by retailer  $R_1$  is similar to the overall supply chain decisions. Differently, marginal cost of centralized supply chain is  $c$ , while the marginal cost faced by retailer  $R_1$  is  $\omega_\phi^{''*}$ ; therefore, the optimal order decision  $Q_1^{''''}$  of retailer  $R_1$  under the circumstance of decentralized supply chain can be acquired by altering  $c$  to  $\omega_\phi^{''*}$  in formula (20):

$$Q_1^{''''} = F^{-1} \left[ \frac{p - \omega_\phi^{''*}}{g + \omega} \right] \quad (23)$$

Due to  $\omega_\phi^{''*} > c$ , then  $(p - \omega_\phi^{''*}) / (g + \omega) - (p - c) / (g + \omega) \geq 0$ , so  $Q_1^{''''} \geq Q_1^{''''}$ .

Now, the expected revenue of retailer  $R_1$  is

$$E \prod_{r_1} (Q_1^{''''}) = (p - \omega - c_r) Q_1^{''''} - g_r \int_{Q_1^{''''}}^{\infty} (x - Q_1^{''''}) f(x) dx - (1 - \alpha) \omega \int_0^{Q_1^{''''}} (Q_1^{''''} - x) f(x) dx \quad (24)$$

The expected revenue of manufacturer  $M_1$  is

$$E \prod_{m_1} (Q_1^{''''}) = (\omega - c_m) Q_1^{''''} - g_m \int_{Q_1^{''''}}^{\infty} (x - Q_1^{''''}) f(x) dx - \alpha \omega \int_0^{Q_1^{''''}} (Q_1^{''''} - x) f(x) dx \quad (25)$$

The precondition to make sure that both retailer  $R_1$  and manufacturer  $M_1$  can accept the contact above is to ensure the profit for both sides under the contact cannot be lower than the profit that will be gained under decentralized. Namely, the profit of retailer  $R_1$  and manufacturer  $M_1$  is the improvement of Pareto must be achieved by the contact, and inequality (26) should be satisfied at the same time:

$$E \prod_{r_1} (Q_1^{''*}) \geq E \prod_{r_1} (Q_1^{''''}), \quad (26)$$

$$E \prod_{m_1} (Q_1^{''*}) \geq E \prod_{m_1} (Q_1^{''''})$$

During the design process of revenue sharing contract, manufacturer needs to ensure the range of numerical value

TABLE 1: The EOQ under the three situations of block chain technology and random demand.

	Decentralized supply chain	Revenue sharing supply chain	Centralized supply chain
EOQ	$Q_1' = 1073$	$Q_1^{''*} = 1081, (\phi = 0.25)$ $Q_1^{''*} = 1086, (\phi = 0.50)$ $Q_1^{''*} = 1077, (\phi = 0.75)$	$Q_1''' = 1078$

of revenue sharing factor  $\phi$  be satisfied with inequality (26). Only by this way retailer  $R_1$  will accept the contact while the interest of manufacturer  $M_1$  can be maintained as well. The specific range of numerical value of  $\phi$  should be decided according to the bargaining ability between retailer  $R_1$  and manufacturer  $M_1$  and the status of retailer  $R_1$  and manufacturer  $M_1$  in the supply chain, respectively.

### 5. Statistics Analysis

Assuming that the supply chain consists of multiple manufacturers  $M$  ( $M_1M_2 \dots M_n$ ) and multiple retailers  $R$  ( $R_1R_2 \dots R_m$ ), retailer  $R_1$  will begin to place orders to manufacturer  $M_1$  every single cycle before the selling season. Market demand follows the normal distribution  $N(1000, 100^2)$ , and retail price  $P = 35$ . Based on block chain ideas, the penalty of breaking the contact for retailer  $R_1$  and the cost of goodwill resulting by outstock for manufacturer  $M_1$  are comparatively high; assuming  $g_m = g_r = 50$ , then the total penalty cost  $g = g_m + g_r = 100$ . The production cost of manufacturer  $M_1$   $c_m = 8$ , based on block chain model; the marginal cost of retailer  $R_1$  is very low; assuming that  $c_r = 0.1$ , the wholesaling price of manufacturer  $M_1$   $\omega = 20$ , the revenue sharing factor of retailer  $R_1$   $\phi \in \{0.25, 0.5, 0.75\}$ , and return discount factor  $\alpha = 0.75$ .  $\pi(M_2 \dots M_n, R_2 \dots R_m) = 500000000$ .

Therefore, the optimal order quality under the three situations of block chain technology and random demand are as shown in Table 1.

According to Table 1, based on random demand and block chain ideas, when the retail price is settled, EOQ of decentralized supply chain is the lowest. And under the circumstance of revenue sharing, the EOQ firstly increased and then decreased according to the change of sharing factor  $\phi$ . When  $\phi = 0.5$ ,  $\phi$  reached the highest point and the EOQ of centralized supply chain is higher than that of decentralized supply chain, but the revenue of centralized supply chain is lower than that of decentralized supply chain. Therefore, based on random demand and block chain ideas and with the application of the combination of repurchase strategy and revenue sharing contract, the order purchasing amount can be improved effectively; the market sharing can be enlarged, and, at the same time, the stock out amount and excessive products amount can be reduced to decrease the products return amount.

Then the optimal revenue under the three situations of block chain technology and random demand are as shown in Table 2.

According to Tables 1 and 2, based on block chain technology and random demand situation, the optimal revenue

decentralized supply chain is the lowest as 50024075. With the coordination of revenue sharing contact, when the sharing factor  $\phi = 0.75$ , the optimal overall revenue of the whole supply chain is the highest as 50026014 and that of centralized supply chain is in the middle of them as 50024210. Under the situation of decentralized supply chain, the revenue of retailers is 13792.96, which is higher than the retailers' optimal revenue (2370.02, 3355.23, and 8531.60, respectively) under the coordination of revenue sharing. And the optimal revenue of manufacturers is 10282.22, which is lower than the manufacturers' optimal revenue (21812.26, 18137.45, and 17482.08, respectively) under the coordination of revenue sharing. Although the optimal revenue of both sides is more equal, the overall revenue of the whole supply chain is lower, which is negative for the supply chain development in the long term.

When the retail price is settled, with the coordination of revenue sharing mechanism, the optimal revenue of retailers will increase and decrease according to the increase and decrease of the sharing factor, while the optimal revenue of manufacturers will change oppositely. This mechanism will be beneficial for manufacturers and the whole supply chain system, but harmful for retailers. Therefore, under this mechanism, retailers can give up certain interest to enlarge the market sales volume, reduce risks, maximize the overall revenue of the whole supply chain, and improve the overall revenue of the whole supply chain to the degree equal to or more than that of centralized supply chain.

### 6. Conclusion

Block chain technology has become a popular approach to apply into different disciplines to explore problems including customer random demand. Our research suggests that key decision-makers in supply chains should consider and value the synergy impact of block chain technology and customer random demand; then we may reduce supply chain risks, maintain supply chain robustness, and improve the core competitiveness of supply chains. Based on interruptive innovation and Stackelberg noncooperative gaming theory, this paper has built the decision models of decentralized supply chain and centralized supply chain and implements revenue sharing contact. Our research has constructed a new supply chain coordination model with features of Del trust, decentralization, and traded anonymous and explored the supply chain coordination under the combined effect of block chain technology and random demand.

The research has got some results. When assuming sales price remains, the overall optimal revenue of decentralized

TABLE 2: The optimal revenue under the three situations of block chain technology and random demand.

Supply chain types	Value of sharing factor $\phi$	Coordinative wholesaling price $\omega_\phi^{l*}$	The optimal revenue for retailers	The optimal revenue for manufacturers	The revenue for other members in the supply chain	The optimal revenue of the whole supply chain
Decentralized supply chain	-	-	13792.96	10282.22	50000000	50024075
revenue sharing supply chain	$\phi = 0.25$	$\omega_\phi^{l*} = 24.1$	2370.02	21812.26	50000000	50024182
	$\phi = 0.50$	$\omega_\phi^{l*} = 26.7$	3355.23	18137.45	50000000	50021493
	$\phi = 0.75$	$\omega_\phi^{l*} = 22.3$	8531.60	17482.08	50000000	50026014
centralized supply chain	-	-	-	-	-	50024210

supply chain under the optimal decisions is minimal and can be of benefit for retailers, but can be harmful for manufacturers and the whole supply chain, leading to negative influence on supply chain robustness in the long term. Also, the coordinative supply chain with revenue sharing mechanism can increase sales, reduce risks, and improve the overall performance of the supply chain to the degree of centralized supply chain. The coordination result may be harmful for retailers but be beneficial for other supply chain members and the whole supply chain. Through the combination of repurchase strategy, revenue sharing contact, and appropriate sharing coefficient, the revenue of supply chain members can be more equal. Further, the coordination result will only influence the short-term revenue of members in the block chain, without effecting on other members' short-term revenue, consequently improving the supply chain robustness dramatically. Therefore, our research results can provide significant implications for both supply chain coordination theory and practice.

However, several limitations should be addressed for future research. First, this paper explored supply chain coordination under the synergy effect of customer random demand and block chain technology, without considering the synergy effect of stable demand and block chain. Second, this paper assumed that there is no defect in the products and ignored the impact of products checking strategy on supply chain coordination. In the future, we will conduct the research about the supply chain coordination issue related to products quality defect and checking strategy to explore the synergy effect of customer random demand and block chain technology.

## Data Availability

The block chain and random demand data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This research was supported by National Social Science Fund of China (no. 18XGL001).

## References

- [1] G. P. Cachon, "The allocation of inventory risk in a supply chain: push, pull, and advance-purchase discount contracts," *Management Science*, vol. 50, no. 2, pp. 222–238, 2004.
- [2] Y. F. Li, Q. Su, and J. Zheng, "Supply chain coordination based on sensitive customer quality demand," *Advances in Information Sciences and Service Sciences*, vol. 4, no. 19, pp. 264–269, 2012.
- [3] Y. F. Li and W. F. Jia, "Supply chain coordination with considering defective quality products cheaply processing understochastic demand," *Journal of Residuals Science & Technology*, vol. 13, no. 6, pp. 731–737, 2016.
- [4] Y. F. Li, "Coordination and risk sharing with considering supply chain external quality fault," *Computer Modelling and New Technologies*, vol. 18, no. 12C, pp. 445–450, 2014.
- [5] G. P. Cachon, "Procuring fast delivery: Sole sourcing with information asymmetry," *Management Science*, vol. 52, no. 6, pp. 881–896, 2006.
- [6] J. Mostard, R. de Koster, and R. Teunter, "The distribution-free newsboy problem with resalable returns," *International Journal of Production Economics*, vol. 97, no. 3, pp. 329–342, 2005.
- [7] B. A. Pasternack, "Optimal pricing and return policies for perishable commodities," *Marketing Science*, vol. 4, no. 2, pp. 166–176, 1985.
- [8] G. P. Cachon and M. A. Lariviere, "Supply chain coordination with revenue-sharing contracts: strengths and limitations," *Management Science*, vol. 51, no. 1, pp. 30–44, 2005.
- [9] Y. He, X. Zhao, L. Zhao, and J. He, "Coordinating a supply chain with effort and price dependent stochastic demand," *Applied Mathematical Modelling*, vol. 33, no. 6, pp. 2777–2790, 2009.
- [10] J. Chen and P. C. Bell, "Coordinating a decentralized supply chain with customer returns and price-dependent stochastic demand using a buyback policy," *European Journal of Operational Research*, vol. 212, no. 2, pp. 293–300, 2011.
- [11] K. Govindan and M. N. Popiuc, "Reverse supply chain coordination by revenue sharing contract: a case for the personal computers industry," *European Journal of Operational Research*, vol. 233, no. 2, pp. 326–336, 2014.
- [12] S. Zhao and Q. Zhu, "Remanufacturing supply chain coordination under the stochastic remanufacturability rate and the random demand," *Annals of Operations Research*, vol. 257, no. 1–2, pp. 661–695, 2017.
- [13] B. C. Giri and A. Chakraborty, "Coordinating a vendor–buyer supply chain with stochastic demand and uncertain yield," *International Journal of Management Science and Engineering Management*, vol. 12, no. 2, pp. 96–103, 2017.
- [14] S. E. Chick, H. Mamani, and D. Simchi-Levi, "Supply chain coordination and influenza vaccination," *Operations Research*, vol. 56, no. 6, pp. 1493–1506, 2008.
- [15] I. Hwang, S. Radhakrishnan, and L. Su, "Vendor certification and appraisal: implications for supplier quality," *Management Science*, vol. 52, no. 10, pp. 1472–1482, 2006.
- [16] Y. F. Li, "Supply Chain Coordination Considering Stochastic Demand and Raw Material Quality Defects," *International Journal of u- and e-Service, Science and Technology*, vol. 9, no. 5, pp. 373–382, 2016.
- [17] K. Christidis and M. Devetsikiotis, "Blockchains and smart contracts for the internet of things," *IEEE Access*, vol. 4, no. 1, pp. 2292–2303, 2016.
- [18] S. Khandelwal, N. Gandhi, D. Choudhari, and P. Joshi, "HIBE Based Security for Content Distribution using Block Chains," *International Journal of Science Technology Management and Research*, vol. 3, no. 10, pp. 5–9, 2018.
- [19] W. Ding, "Block chain based instrument data management system," *China Instrumentation*, vol. 10, no. 1, pp. 15–17, 2015.
- [20] D. Kraft, "Difficulty control for blockchain-based consensus systems," *Peer-to-Peer Networking and Applications*, vol. 9, no. 2, pp. 397–413, 2016.
- [21] D. Wilson and G. Ateniese, "From pretty good to great: enhancing pgp using bitcoin and the blockchain," in *Proceedings of the 9th International Conference on Network and System Security*, vol. 9408 of *Lecture Notes in Computer Science*, pp. 368–375, Springer International Publishing, 2015.

- [22] J. P. Geunes, R. V. Ramasesh, and J. C. Hayya, "Adapting the newsvendor model for infinite-horizon inventory systems," *International Journal of Production Economics*, vol. 72, no. 3, pp. 237–250, 2001.

## Research Article

# Adopting a QCA Approach to Investigating the Risks Involved in Megaprojects from Auditing Perspective

Liang Ma,<sup>1</sup> Bing Zhang ,<sup>2</sup> Miao Cui,<sup>2</sup> and Ruoyu Jin<sup>3</sup>

<sup>1</sup>School of Management, Shanghai University, Shanghai 200444, China

<sup>2</sup>College of Civil Science and Engineering, Yangzhou University, Yangzhou 225127, China

<sup>3</sup>School of Environment and Technology, University of Brighton, BN2 4GJ, UK

Correspondence should be addressed to Bing Zhang; [glzhangbing@126.com](mailto:glzhangbing@126.com)

Received 7 September 2018; Revised 12 November 2018; Accepted 6 January 2019; Published 20 January 2019

Guest Editor: Xinchang Wang

Copyright © 2019 Liang Ma 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.

There is an increase of megaproject construction worldwide. At the same time, risks involved in megaprojects have also become a wide concern. Extending from the macrolevel of qualitative analysis focusing on complexity, politics, and morality, the research conducted the microscopic empirical analysis on twenty-two typical cases by adopting the quality comparative analysis (QCA) from the auditing perspective. Different from the traditional analysis method taking each causation as independent variable, the results in the study revealed that there was complex multiple concurrent causation among eight conditions; additionally, the configuration of those would be divided into six types, among which, the coverage of the three types, namely, project management risk, preliminary and construction risk, and tendering and contract management related risk, was almost eighty percent. Finally, megaproject risks in China were caused by complicated and changeable combination conditions, which would provide a new breakthrough for seeking analyzing megaproject risks through this quantitative analysis method, and indicate the researchers and practitioners to control the megaproject risks from a more systematic way.

## 1. Introduction

As a special means of economic control, megaproject auditing provides an opportunity to uncover those misconduct issues by financial and compliance audits, which detect and combat those unethical practices in those megaprojects, to ensure fundamental project principles with appropriate practices. Megaproject auditing is a cornerstone of good megaproject governance, which is an “immune system” and watchdog of megaprojects. It plays a significant role in improving the ability of megaproject management, preventing the risks, and ensuring fundamental project management principles among megaproject design, tender, construction, and settlement phase [1]. Furthermore, it helps public organization achieve accountability and integrity, as well as improving operation [1].

In China, with megaproject construction entering a new era of “the biggest investment boom in history,” the whole nation has been emphasizing the importance of megaproject auditing [2]. However, unlike some other developed countries that place significant importance to the performance

audit, China’s megaproject auditing plays an important role in oversight and insight of megaproject, and many misconducts and its risks are disclosed; i.e., auditing shows that over one-third of megaproject investment from 1991 to 1995 was ineffective [3]. For example, in the Beijing-Shanghai high-speed railway, the intercepting embezzlement funds, unstandardized procurement, the idle construction materials, and unpaid labor services were discovered with amount over 10 billion Yuan [4].

Thus, although megaproject auditing is the postmortem supervision, it could supervise and manage megaprojects effectively, prevent the emergence of misconduct risk, and further improve the investment efficiency of audit supervision. More importantly, megaproject auditing is no longer a matter of reviewing misconduct but could also analyze the cause and typology of the misconduct problems, which could gradually develop from traditional auditing to the modern performance audit. However, there is still a “black box” for systemic classification of megaproject risks through auditing, and the existing classification guidelines always focus on one

aspect, such as quality, investment, or corruption. Furthermore, to uphold auditing megaprojects, it is suggested to measure its functions and effect, especially the relationship of the megaproject risk and its effect from auditing.

It is indicated that, with the whole world entering a new “tera age” of megaprojects [5], more attention has been paid on the discussion of megaproject auditing risk. Most of the existing literature is based on the authors’ experience from observation and theoretical induction according to individual case analysis. However, due to lack of representativeness and the results are not universal, their conclusion is not suitable for a wider implication because it cannot handle multicase analysis, so individual case analysis has fallen far short of the research needs. Furthermore, to reveal the mechanism of megaproject risks, the conventional statistical analysis is not enough, nor could it explain the complex causality. Therefore, a holistic and systematic approach is needed to explore the complex interaction between those risks.

The paper contributes several ways to give a more systematic analysis on the risks of megaprojects. First, the research is based on the deep analysis of auditing reports from government instead of observation and theoretical induction, which makes the result more objective and valuable. Second, twenty-two typical cases are analyzed by QCA method, combining the qualitative case study and quantitative statistical analysis together, which makes the methodology more reasonable and advanced compared to the previous studies. Third, different from the independent relationship analysis between the single variables with result dependent variables, the research integrates complicated and changeable combination conditions, and more comprehensive influencing factors have been identified. The study revealed that there was complex multiple concurrent causation among eight conditions for megaproject risk; additionally, the configuration of those would be divided into six types, among which, project management risk, preliminary and construction risk, and tendering and contract management related risk were almost eighty percent.

The paper also effectively faces up to the practical concerns. The government and owners should pay more attention to control the risks in project process management, enhance the effectiveness of preliminary and construction management period, and design transparent and strict regulations for tendering and contract management; they should distribute more human and technology resources to these facets. The study also indicates for the contractors and designers from the industries that risks should be monitored and controlled by integrated methods since the risk types are diverse and many factors may interconnect during the life cycle of megaproject. All the entities in the market should work together to handle the various risks from megaprojects.

## 2. Literature Review

*2.1. Megaproject Risk.* Megaproject, as multiactor and multitechnology constellations [6], considered to be the most complex of all different types of projects, is a large-scale project characterized by uncertainty and ambiguity [7].

Due to its high volume of investment, great community involvement, complex decision-making processes, and many stakeholders involvement, megaproject is the wild beasts in the project world and hard to tame [8, 9], and those projects are exposed to high levels of risks [10]. Many megaprojects such as airport, bridge, and highway are usually money pits where funds are simply “swallowed up” without delivering sufficient returns as a result of unbalanced subjective beliefs and information in assessing risks and uncertainties, and taking corrective actions to effectively control and manage the identified risks at the right time [11]. So identifying and analyzing megaproject risks are currently considered a mandatory part for their great impact on megaproject [8]. With the expansion of size and scale, the associated megaproject risk would increase [12].

Identifying megaproject risks is particularly effective for risk management. Bruzelius and Rothengatter [13] proposed four megaproject risk categories, namely, cost risk, demand risk, financial market risk, and political risk; Bing et al. [14] proposed a distinction of megaproject risks between macro-, meso-, and microlevels. Little [15] developed a wider classification which included political risks, construction risks, operation and maintenance risks, legal and contractual risks, and financial risks. Within the misconduct risk for megaproject, three main illegal behavior risks are lack of supervision, imperfect laws and regulations, and lack of enterprise internal management. Risk analysis, construction on-site management, the usage of constructive funds, tendering and bidding phase, and environmental protection are the four most influences [16].

*2.2. Megaproject Auditing.* Auditing plays a key role in helping understand risks and initiate risk assessments, including identifying, assessing, and managing risks, ensuring that the audit resources are used effectively to address the areas of greatest exposure [1]. The nature of risks identified would vary according to the audit objective; audit may look “upstream” to how decisions are made within government departments and agencies or look “downstream” to how services are delivered and perhaps trick to prove a causal link between audit and performance. Megaproject auditing provides an opportunity to uncover the challenges and risks encountered in execution phases, including project governance, activities, cost estimations, and so on [10].

In China, the megaproject auditing is mainly paying attention to compliance of construction project, the cost authenticity, budget execution, and accounting statements supervision and examination; the key points of megaproject auditing includes decision making and supervision of budget and bidding procedure [17]. For example, the Guidelines on Synchronous Prevention of Auditing Risk in Megaprojects of Foshan City involved nine risk areas including tendering and bidding procedure, subcontracting, and change management. Furthermore, 38 supervision points were also listed [18]. Its work modes were mainly in three types: governmental auditing, internal auditing, and social auditing [19].

Megaproject auditing is a complicated activity, involving many units and personnel with complicated procedures, especially the variety of problems. Therefore, identifying and

assessing risk of megaproject are a basic element for auditing. However, the existing megaproject risk classifications remain insufficient to identify types of risks in practice. Specifically, those analyses often lack context and systematization. There is limited evidence of a consensus on risk classification in literature. Those classifications are merely a guide to risk identification [8]. Furthermore, the most commonly employed methodology for analyzing megaproject risk is case study, which represented over 40 percent of articles, of which 67.65 percent are a single case study [8].

### 3. Research Method and Data Description

**3.1. QCA Method.** There are some well-known difficulties in empirical research on megaproject risks via auditing. Firstly, it is difficult to obtain real data. Secondly, the public data is often fragmented and difficult to analyze systematically because of selective bias. The current research leans towards the qualitative aspects. It aims to provide a feasible way by selecting the multiple typical cases published by authorities.

Considering megaproject risks are generally complex causal system. Its cause mechanism is complex. There are numerous factors contributing to megaproject risks. Previous research [20, 21] based on qualitative discussion mostly failed to provide a comprehensive explanation. It is important to find a new and effective way to explore the gap of causation path of serious megaproject risks via auditing. The use of qualitative comparative analysis (QCA) in management research highlights that the complexity of management phenomena contains parsimonious causal paths that research can unveil. Thus, QCA can improve the understanding of management realities while preserving their holistic nature [22]. For the above reasons, this research uses QCA method to understand the influence factors and formation mechanism of megaproject risks via auditing, which could combine the advantages of qualitative and quantitative analysis. China's first megaproject auditing report was published in 2005; in the past ten years, its tasks and priorities on megaproject auditing of National Audit Office have taken place; although the identified megaproject risks via auditing have multiple orientations, the compliance auditing is still the focus and priority of the megaproject auditing [23]. Megaproject risks discovered by auditing could combine a diversity of occurrence paths and the same adoption result; there may be multiple equifinal causal chains that lead to the same compliance auditing results, which is more suitable for QCA method. Although this research adopts the multicase analysis, the case sample cannot size up to the level of the large sample; this would make it hard to get solid results by statistical method, so considering QCA is good at small sample analysis between 10 and 40 cases for deep understanding of actual phenomenon with combination of quantitative statistical analysis and qualitative analysis [24]. The aim of QCA method is to find out the causal relationships between the conditional configuration and the result through case comparison, answering which configuration of condition can lead to the expected result and which configuration could induce results in the absence with considering the interdependence of influencing factors.

QCA methods constitute promising methodological tools addressing the gap between variable-oriented and case-oriented research [25]. There are three main analysis methods in QCA, namely, crisp set QCA (csQCA), fuzzy set QCA (fsQCA), and multivalued QCA (mvQCA), among which csQCA is the first QCA technique developed and the most widely used so far [26]. csQCA is a comparative case-oriented research technique based on set theory and Boolean algebra, aiming to develop explanatory models on the basis of a systematic comparison of a limited number of cases (less than 100), and its conditions range from 2 to 13, so csQCA could integrate the best features of the case-oriented approach with the best features of the variable-oriented approach [27]. This research applies csQCA to analyze the megaproject risks via auditing, which would reveal the complex relationships of risks in the conditions of multivariate and multidimension, and achieve scientific method support for megaproject risk analysis.

**3.2. Data Selection.** In recent years, with the increasing emphasis on megaproject auditing at the national level in China, a large number of typical megaproject auditing reports have been published, especially the megaproject auditing report issued by National Audit Office; those auditing reports usually consist of three parts, that is, megaproject implementation and major achievements, major problems and risks identified, and preliminary rectification feedback, which provides an unprecedented opportunity for quantitative case analysis on megaproject risks via auditing.

Due to numerous risk factors and conditions in megaproject auditing report, to ensure the rationality and accuracy of the csQCA analysis results, this research applied the Law of Common Auditing Qualitative Description and Applicable Regulation Guide-Fixed Assets Investment Auditing (Trail), which is a summary of the National Audit Office on the auditing discover problems and risks of government projects including megaprojects. Afterwards, by the China Auditing Common Qualitative Expressions and Applicable Regulations Guide-Fixed Assets Investment Audit [28], eight classification and forms of those discovery risks were determined; more importantly, each classification lists the specific manifestation and problem characteristics; to some extent, those eight classifications are systematic analysis, induction, and classification; this is suitable as conditions for QCA analysis, so the eight categories would be selected as conditions for megaproject risks via auditing (Table 1); furthermore, considering csQCA provides a set of tools for analyzing the necessary and sufficient conditions explaining outcomes, mapping out similarities and differences between various configurations of conditions and cases [29]; the outcome of the csQCA analysis would select the proportion of illegal expenditure; the higher the indicator, the lower the compliance of its expenditure, and the greater of megaproject risk [23] and then collecting and sorting out 42 auditing reports from 2005 to 2017 on the National Audit Office website (<http://www.audit.gov.cn/>). Finally, a total of 22 cases were selected as research cases according to its integrity; this complete information includes the following: megaproject name, value amount submitted for auditing,

TABLE 1: csQCA conditions and outcome.

Measure	Name	Abbreviation
Conditions	Violation construction procedures	Pro
	Violation of funds management and accounting regulations	Cap
	Violation of tendering and bidding & contract management regulations	Bid
	Violation of quality management regulations	Qua
	Violation of construction management regulations	Con
	Violation of investment management regulations	Inv
	Violation of land requisition and immigration regulations	Req
	Violation of environmental protection and historic preservation regulations	Env
Outcome	Proportion of illegal expenditure	Res

TABLE 2: The distribution of risks of megaproject types involved.

Types	Frequency	Percentage	Cumulative Percentage
Traffic megaprojects	9	40.91	40.91
Hydraulic megaprojects	5	22.73	63.64
Government-subsidized housing	4	18.18	81.82
Post-earthquake reconstruction	2	9.09	90.91
Airports and stadiums	2	9.09	100.00

value amount of megaproject risk found, and the megaproject risks.

**3.3. Megaproject Types.** In recent years, there have been a number of state-led megaprojects with complex types, in order to better understand which types of megaproject are found to be more risky. According to 22 cases, the megaprojects are divided into five types (Table 2). Among them, traffic megaprojects include two subtypes of highway and railway projects.

It is seen that there are 9 traffic megaprojects among the 22 cases involving megaproject risks, which is over 40% of the whole cases. Hence, traffic megaprojects are a major disaster area for megaproject risks, followed by hydraulic megaprojects, which account for more than 22% of all. For the other three types of megaprojects, postearth reconstruction projects refer to the megaproject investment made by the central and local governments after the Wenchuan earthquake in 2008; its total investment exceeds 1.5 trillion Yuan, while for government-subsidized housing projects, its massive construction began in 2011, would cost about 1.3 trillion Chinese Yuan (roughly \$200 billion) furthermore, airport and stadiums are not being built across the country like other megaproject due to their financial constraint on local governments. Subject to its development time or financial constraints, fewer cases are disclosed.

#### 4. csQCA Empirical Analysis

This csQCA analysis includes three stages, which is calibrating dichotomized sets, constructing the truth table and resolving contradictions, and configuration explanations. The main task of calibrating dichotomized sets is to determine the attribution of each condition and outcome. The task in

this research is the distribution between 0 and 1. The role of constructing the truth table and resolving contradictions is a given combination of conditions associated with a given result according to consistency threshold and frequency threshold. The configuration explanations focus on simplifying selected conditional configuration by parsimonious solution and intermediate solution according to easy counterfactual and difficult counterfactual analysis. Then these three sections are carried out below.

**4.1. Calibrating Dichotomized Sets.** csQCA is based on the logic of Boolean algebra, which allow the minimal formulas that are parsimonious and its combination of conditions and outcome have values of 0 or 1 with dichotomous variables, so as to construct complex data structures using logical and holistic approaches. Considering there are two types of research data, i.e., continuous sample data and discrete sample data, the binary threshold is determined by two methods. For the discrete sample data, which focus on eight conditional variables, the threshold of dichotomous threshold is divided by whether or not there is violation of relevant provisions of the Law of Common Auditing Qualitative Description and Applicable Regulation Guide--Fixed Assets Investment Auditing (Trail); if violated, the value of the conditional variables is 1, and vice versa, while for continuous sample data, which focus on outcome, that is the proportion of illegal expenditure. It is seen that, for the 22 cases, and the minimum amount of problems auditing funding is 240 million Yuan, and the maximum amount is 56.73 billion Yuan. Furthermore, the minimum of the proportion of illegal expenditure is 0.014, the maximum amount of illegal expenditure is 0.206, and the mean of proportion of illegal expenditure is 0.07; according to the Rihoux and Ragin [25], the mechanical cut-off point (mean value) would be selected for threshold. If the proportion of illegal expenditure is over 0.07, the value of

TABLE 3: csQCA indicators and its dichotomized sets.

Abbreviation	Determines	Percentage	Value
Pro	Violating construction procedures	31.82%	1
	Not occurred	68.18%	0
Cap	Violating funds management and accounting regulations	95.45%	1
	Not occurred	4.55%	0
Bid	Violating tendering and bidding & contract management regulations	72.72%	1
	Not occurred	27.28%	0
Qua	Violating quality management regulations	77.27%	1
	Not occurred	22.73%	0
Con	Violating construction management regulations	72.70%	1
	Not occurred	27.30%	0
Inv	Violating investment management regulations	63.63%	1
	Not occurred	36.37%	0
Req	Violating land requisition and immigration regulations	33.33%	1
	Not occurred	66.67%	0
Env	Violating environmental protection and historic preservation regulations	22.72%	1
	Not occurred	77.28%	0
Res	The value of proportion of illegal expenditure is over 0.07	54.55%	1
	The value of proportion of illegal expenditure is less than 0.07	45.45%	0

outcome is 1, and vice versa; the specific settings are shown in Table 3.

It is indicated that violating funds management and accounting regulations (Cap), violating quality management regulations (Qua), violating tendering and bidding and contract management regulations (Bid), and violating construction management regulations (Con) are the four important sectors of megaproject risks via auditing, over 70% of cases violated those regulations, especially for violating funds management and accounting regulations (Cap), and almost all the cases (except one) violated funds management and accounting regulations.

#### 4.2. Constructing the Truth Table and Resolve Contradictions.

In order to use Boolean algebra as a technique of qualitative comparison, it is necessary to reconstruct a raw data matrix as a truth table [27]. The truth table includes a binary configuration combination of all the conditions influencing the results, which could reconstruct data as a list of configurations, and each configuration might correspond to a few observed cases [30], so the different combinations of input condition values and their associated outcome values are summarized in a truth table, and the number of the truth tables of these 8 conditions is 256 ( $2^8 = 256$ ). However, 256 configuration is ideal, in the actual analysis, due to the limitation of case scenario and conditional variables; there may be exhibiting contradictory configurations. Then, following the truth table, the key operation is the Boolean minimization, which is used to minimize the configurations and nonobserved cases; coverage and consistency are, respectively, used to determine whether the configuration is the necessary and sufficient condition for the outcome [31]. Coverage measures how much of the outcome is explained by each solution term and by the solution as a whole, consistency measures the degree to which solution terms and the solution as a whole

are subsets of the outcome, and their threshold is 0.8, if these values reach 0.8 or above, indicating that these 8 conditions have an important effect on the outcome. There are three types of solutions: complex, intermediate, and parsimonious solution, and each solution is based on a different treatment of the remaining combination, considering numerous studies applying intermediate solution for conditional configuration analysis [32], so this research would select intermediate solution, and Table 4 shows the values of coverage and consistency.

In Table 4, solution coverage measures the proportion of memberships in the outcome that is explained by the complete solution; solution consistency measures the degree to which membership in the solution is a subset of membership in the outcome, raw coverage measures the proportion of memberships in the outcome explained by each term of the solution, and unique coverage measures the proportion of membership in the outcome explained by each individual solution term, while “\*” means the preceding and the following term are multiplied, and “~” means not logical, which is true if the following term is false. The results of Table 4 show the value of solution coverage and solution consistency is over 0.8, which meets the configuration requirements of coverage and consistency, indicating that the analysis results have strong explanatory strengths of the csQCA. There are 6 groups of configuration which is linked together by “\*” which forms configurations for megaproject risks via auditing.

4.3. Configuration Explanations. For the parsimonious and intermediate solutions, which come from the results of easy counterfactual and difficult counterfactual analysis, to deal with the problem of limited diversity, this step would logically reduce the truth table rows to simplified combinations using parsimonious and intermediate solutions, and there are two causal conditions in configurations, which are core

TABLE 4: Values of coverage and consistency.

Solutions	Raw coverage	Unique coverage	Consistency
$\sim$ Pro*Cap*Bid*Qua*Con*Inv* $\sim$ Env	0.417	0.333	1.000
Pro*Cap* $\sim$ Bid*Qua*Con*Req* $\sim$ Env	0.167	0.167	1.000
$\sim$ Pro*Bid*Qua*Con*Inv*Req* $\sim$ Env	0.167	0.083	1.000
$\sim$ Pro*Cap*Bid* $\sim$ Qua* $\sim$ Con*Inv* $\sim$ Req*Env	0.083	0.083	1.000
Pro*Cap*Bid*Qua*Con* $\sim$ Inv* $\sim$ Req* $\sim$ Env	0.083	0.083	1.000
Pro*Cap*Bid*Qua*Con*Inv* $\sim$ Req*Env	0.083	0.083	1.000
Solution Coverage	0.917		
Solution Consistency	1.000		

TABLE 5: Configurations for megaproject risks via auditing.

Condition	Solution					
	M1	M2	M3	M4	M5	M6
Pro	⊗	●	⊗	⊗		
Cap	●	•		•	•	•
Bid	●		●	•	•	•
Qua	●	●	●	⊗		
Con	●	●	●	⊗		
Inv	●		●	●		●
Req		●	•	⊗	⊗	⊗
Env	⊗			●		●
Consistency	1.000	1.000	1.000	1.000	1.000	1.000
Raw coverage	0.417	0.167	0.167	0.083	0.083	0.083
Unique coverage	0.333	0.167	0.083	0.083	0.083	0.083
Overall solution consistency	1.000					
Overall solution coverage	0.917					

conditions and peripheral conditions, core conditions are those part of both parsimonious and intermediate solutions, while peripheral conditions are those that only appear in the intermediate solution [33], and using full circles indicate the presence of a condition, while cross-out circles indicate the absence of a condition, additionally, large circles means core conditions, while small one refers to peripheral conditions, and blank spaces indicate a “do not care” situation in which the causal condition may be either present or absent [34]. Table 5 indicates the configurations for megaproject risks via auditing.

The solution table shows that the crisp-set analysis results in 6 solutions exhibiting acceptable consistency and coverage, and those 6 solutions are named M1, M2, . . . , to M6, considering the overall raw coverage of M1, M2, and M3 is almost 80%, which means those three solutions are the main configurations of megaproject risks via auditing, and thus this research would focus on the three most common megaproject risk configurations, i.e., project management risk, preliminary and construction risk, and tendering and contract management related risk. All the three configurations involve Qua and Con, showing these two conditions are universal among megaproject risks.

4.3.1. *Project Management Risk Configuration.* The raw coverage of this configuration is 0.417 and contains 7 conditions,

and those are all core conditions, among which five are the presence of conditions, including Cap, Bid, Qua, Con, and Inv, while two of them are absence of conditions, that is, Pro and Env. It could be seen that over 40% of 22 cases are involved in this configuration. This megaproject risk configuration involves all the usual construction project management issues except schedule, because, in China, megaproject schedule delay is almost unacceptable and would induce political risk even political disaster for project managers [35]. The megaproject risks caused by this configuration is common and has great influence; there are three main reasons for this configuration. First of all, the investment amount of megaproject projects is often very large, and these megaprojects are often government-led, and control over the whole megaproject and the mobilization of large amounts of project funds are often in the hands of a small number of leaders, which provides fertility for megaproject risk on backroom trading. After all, because the whole construction period of megaproject is long, and there are numerous procedures and stakeholders involved, this hinders the supervision work, causing the megaproject construction to have many risks.

This megaproject risk configuration often starts from the megaproject beginning of tendering and bidding to project completion and payments of funds; i.e. the qualifications of the winning construction enterprises could not meet the

requirement of project grad; some megaprojects even failed to carry out tendering and bidding activities in accordance with the requirements of laws. The project quality is below the design standard, transferring, occupying, and misappropriating even defraud project funds. The whole construction process is chaotic and lacking in management; a large number of institutions are virtually nonexistent. Serious project funds management problems are often associated, for example, with the megaproject of Chongqing to Changsha expressway; there were 78.217 million fake invoices as financial vouchers, the main contractor subcontracted the project in violation of regulations without the permission of the owner, and the draft final accounts for the completion of the project exceeded the estimated amount of 1.139 billion Yuan.

*4.3.2. Preliminary and Construction Risk Configuration.* The raw coverage of this configuration is 0.167 and contains 5 conditions, and, among those conditions, the conditions of Pro, Qua, Con, and Req are the core conditions, while the condition of Cap is peripheral condition; additionally, those conditions are all the presence of conditions. Compared with the configuration of M1 (project management risk configuration), this configuration includes the conditions of Pro and Req, mainly expressed in the form of not going through the formalities according to the regulations, approval of unauthorized, starting megaproject construction without approval, and other forms. Because of the risk of those megaproject procedures, it has a certain impact on project demolition and construction; for demolition, there are some risks of compensation for land expropriation, such as expropriation and detention.

While, for construction, there are defects and risks in the preliminary, these risks would inevitably be reflected in the process of megaproject construction; the most intuitive manifestation is the construction quality risks, especially these megaprojects often have strict deadline requirements and there is a certain deadline behavior, inducing in the construction period, and the basic construction procedures have not been strictly implemented, such as there is failure to follow the drawings for construction and the relevant supervisors and others do not strictly follow the rules because of deadline behavior for construction. And these risks caused construction quality problems due to violations and deadline behavior, which makes this configuration more likely to lead to financial problems, especially overexpenditure and project management risks. For example, in guaranteeing housing projects, a total of 45 projects across the country have not been approved for land use planning, involving 1433 mu of land, and there are 5.8 billion Yuan funds for 360 projects that have problems.

*4.3.3. Tendering and Contract Management Related Risk Configuration.* The raw coverage of this configuration is 0.167 and contains 6 conditions, and there are five core conditions except Req, and, except Pro, the other five conditions of Bid, Qua, Con, Inv, and Req are all the presence of conditions. This configuration of megaproject risks via auditing is mainly caused by the uncontrollable tendering and contract management, the process of tendering of megaprojects is

in chaos, and even the tendering system is virtually nonexistent, i.e., set unseasonal terms in the bidding documents to exclude potential bidders, and turn the whole projects into several parts to avoid tendering, furthermore, there is illegal subcontracting, as a result, it is difficult to select qualified megaproject contracts, and it is inevitable that the megaproject quality cannot meet the requirements, and the investment is out of control and the other construction management problems emerge endlessly. For example, in the megaproject of expressway of Changchun to Hunchun, there are 166 subprojects of construction, design, supervision, and other procurement that have not been tendering for public bidding as required, which involved 2.051 billion Yuan, and then there were random design changes for unqualified acceptance of project quality and more than 57.523 million Yuan were paid to 34 contractors without voucher.

## 5. Conclusions

With the increasing complexity of megaproject, project risks are also increasing. A large number of researchers and government have realized that the traditional research methods on discovering megaproject risks have been difficult to reflect the large amount of problems, especially the traditional auditing on a certain risks. As a result, the corresponding suggestions proposed would have little effects on solving the mitigating megaproject risks.

Configuration analysis is a new type of analysis tool exploring internal operation essence of megaproject risks and grasping their microoperation process; thus this research would analyze the megaproject risks through auditing and sorting out 8 conditions combined with the relevant specification. Afterwards, six configurations were summarized by csQCA. The results indicated that the megaproject risks in China were not caused by a single factor, but by complicated and changeable combination conditions, which would provide a new breakthrough for seeking analyzing megaproject risks through this quantitative analysis method and systematic thinking for the industry to improve the megaproject risk controlling level.

It was also inferred that adopting the result-driven conditions factors to explore the configuration of megaproject risks via auditing would set up the relationship between complex megaproject risks and the limited paths of these problems, and revealing their intuitive connections with each other. Six configuration path analyses would bring about a turning point to solve the present situation of serious risk problems in megaprojects, which would grasp the main problem path operation mode of megaproject risks and then formulate the relevant regulations to improve the megaproject risk management. This research further identified that there were three configurations with the highest distribution among the six configurations analyzed. And the conditions of Qua and Con are the widespread risk point. The practitioners and researchers in the future may pay more attention to the above factors and points to largely enhance the megaproject management performance. However, although the cases in this research are all from the key auditing cases published by the National Audit Office of China, which are representative

to a certain extent, there is limitation in case number, and more cases are needed to further verify the scientific nature of the results.

### Data Availability

The [Excel Form File] data used to support the findings of this study are available from the corresponding author upon request.

### Conflicts of Interest

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

### Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant no. 71701179), the Humanities and Social Sciences Foundation of the Ministry of Education (Grant no. 17YJC630206), and Shanghai Pujiang Program (Grant no. 16PJ1432400).

### References

- [1] R. Date, "Supplemental guidance: the role of auditing in supplemental guidance :the role of auditing in public sector governance table of contents," *The Institute of Internal Auditors*, p. 26, 2012.
- [2] S. Xian and H. Wei, "Research on the Impact of the "Belt and Road" Strategy on China's Public Engineering Auditing Model," *Friends of Accounting*, vol. 3, pp. 108–112, 2017.
- [3] Z. Lijia, "Repositioning of audit function will curb high official corruption," *Shanghai Auditing*, 2004, <http://sjj.sh.gov.cn/sj2014/zwgk/dtyw/n388/ula2066.html>.
- [4] Z. Xiaosong, "Audit announcement warns of serious management for Megaprojects," *Chinadaily*, 2012, [http://www.chinadaily.com.cn/hqpl/zggc/2012-03-20/content\\_5471132.html](http://www.chinadaily.com.cn/hqpl/zggc/2012-03-20/content_5471132.html).
- [5] B. Flyvbjerg, "What you should know about megaprojects and why: An overview," *Project Management Journal*, vol. 45, no. 2, pp. 6–19, 2014.
- [6] J. Lehtinen, A. Peltokorpi, and K. Artto, "Megaprojects as organizational platforms and technology platforms for value creation," *International Journal of Project Management*, vol. 37, no. 1, pp. 43–58, 2019.
- [7] S. Floricel and R. Miller, "Strategizing for anticipated risks and turbulence in large-scale engineering projects," *International Journal of Project Management*, vol. 19, no. 8, pp. 445–455, 2001.
- [8] A. Sanchez-Cazorla, R. Alfalla-Luque, and A. I. Irimia-Dieguez, "Risk Identification in Megaprojects as a Crucial Phase of Risk Management: A Literature Review," *Project Management Journal*, vol. 47, no. 6, pp. 75–93, 2016.
- [9] Y. J. Zidane, A. Johansen, and A. Ekambaram, "Megaprojects-Challenges and Lessons Learned," *Procedia - Social and Behavioral Sciences*, vol. 74, pp. 349–357, 2013.
- [10] M. Huemann and F. Anbari, "Project auditing: a tool for compliance, governance, empowerment, and improvement," *Journal of Academy of Business and Economics*, vol. 7, no. 2, pp. 1–9, 2007.
- [11] P. Boateng, Z. Chen, and S. O. Ogunlana, *Megaproject Risk Analysis and Simulation: A Dynamic Systems Approach*, Emerald Publishing Limited, 2017.
- [12] Y. H. Kwak, "Perceptions and Practices of Project Risk Management: Aggregating 300 Project Manager Years," in *Proceedings of 2003 PMI® Global Congress - North America*, 2003.
- [13] N. Bruzelius, B. Flyvbjerg, and W. Rothengatter, "Big decisions, big risks. Improving accountability in mega projects," *Transport Policy*, vol. 9, no. 2, pp. 143–154, 2002.
- [14] L. Bing, A. Akintoye, P. J. Edwards, and C. Hardcastle, "The allocation of risk in PPP/PFI construction projects in the UK," *International Journal of Project Management*, vol. 23, no. 1, pp. 25–35, 2005.
- [15] R. G. Little, "The Emerging Role of Public Private Partnerships in Mega-Project Delivery," *Social Science Electronic Publishing*, vol. 16, no. 3, pp. 240–249, 2011.
- [16] S. Jianguo, Q. Wang, and S. Changzheng, "Analysis about the prevention of professional crimes and relative policy to it," *Law Science Magazine*, vol. 8, p. 110, 2009.
- [17] L. Guiwen and L. Qiang, "Innovation of public engineering auditing model based on value management," *Science and Technology Management Research*, vol. 27, no. 7, pp. 235–237, 2007.
- [18] Y. Yining and Z. Zhaohua, "Construction and exploration of synchronous prevention and standardization of integrity risk in megaprojects in foshan city," *China Standardization*, vol. 24, pp. 234–235, 2017.
- [19] Wuhan Special Office of China Audit Commission, "Combination of government audit, internal audit and social audit - exploring new audit roads for national mega construction projects," *China Auditing*, vol. 14, pp. 17–18, 2004.
- [20] B. Flyvbjerg, N. Bruzelius, and W. Rothengatter, *Megaprojects and Risk: An Anatomy of Ambition*, Cambridge University Press, 2003.
- [21] A. I. Irimia-Diéguez, A. Sanchez-Cazorla, and R. Alfalla-Luque, "Risk Management in Megaprojects," *Procedia - Social and Behavioral Sciences*, vol. 119, pp. 407–416, 2014.
- [22] A. K. Seny Kan, E. Adegbite, S. El Omari, and M. Abdellatif, "On the use of qualitative comparative analysis in management," *Journal of Business Research*, vol. 69, no. 4, pp. 1458–1463, 2016.
- [23] W. Shicheng and W. guo, "Research on evaluation index system of large-scale investment performance audit," *Audit Research*, vol. 5, pp. 42–47, 2010.
- [24] C. C. Ragin, *Redesigning social inquiry : Fuzzy Sets and beyond*, University of Chicago Press, 2008.
- [25] S.-E. Skaaning, "Assessing the robustness of crisp-set and fuzzy-set qca results," *Sociological Methods & Research*, vol. 40, no. 2, pp. 391–408, 2011.
- [26] B. Rihoux and C. C. Ragin, *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, SAGE Publications, Inc., 2009.
- [27] B. Rihoux and C. Ragin, "User's Guide to Fuzzy-Set/Qualitative Comparative Analysis 2.0, 2008," <http://www.u.arizona.edu/~cragin/fsQCA/download/fsQCAManual.pdf>.
- [28] China Audit Commission Fixed Assets Investment Audit Division, *China Auditing Common Qualitative Expressions and Applicable Regulations Guide - Fixed Assets Investment Audit*, 2014.
- [29] A. Marx and A. Dusa, "Crisp-Set Qualitative Comparative Analysis (csQCA), Contradictions and Consistency Benchmarks for Model Specification," *Methodological Innovations Online*, vol. 6, no. 2, pp. 97–142, 2011.

- [30] B. Rihoux, P. Álamos-Concha, D. Bol, A. Marx, and I. Rezsöhazy, "From niche to mainstream method? a comprehensive mapping of qca applications in journal articles from 1984 to 2011," *Political Research Quarterly*, vol. 66, no. 1, pp. 175–184, 2013.
- [31] J. Bieliauskaite, *The causes of external recognition: a crisp set analysis*, [Master, thesis], Leiden University, 2014.
- [32] M. Zhanwen, "Qualitative comparative analysis (QCA) and news communication studies," *International press*, vol. 38, no. 4, pp. 6–25, 2016.
- [33] P. C. Fiss, "Building better causal theories: A fuzzy set approach to typologies in organization research," *Academy of Management Journal (AMJ)*, vol. 54, no. 2, pp. 393–420, 2011.
- [34] C. C. Ragin and P. C. Fiss, "Net effects analysis versus configurational analysis: An empirical demonstration," *Redesigning Social Inquiry: Fuzzy Sets and beyond*, pp. 190–212, 2008.
- [35] B. Ju, L. Yongkui, L. Yujie, and L. Yun, "An approach to forming top management team of mega infrastructure projects based on improved case-based reasoning," *Journal of Systems & Management*, vol. 25, no. 2, pp. 272–281, 2016.

## Research Article

# A Metaheuristic Algorithm to Transporter Scheduling for Assembly Blocks in a Shipyard considering Precedence and Cooperating Constraints

Ning-Rong Tao,<sup>1</sup> Zu-Hua Jiang,<sup>2,3</sup> Jian-Feng Liu,<sup>4</sup> Bei-Xin Xia ,<sup>5</sup> and Bai-He Li<sup>2,3</sup>

<sup>1</sup>College of Science, Technology and Engineering, Shanghai Ocean University, Shanghai 201306, China

<sup>2</sup>Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 20240, China

<sup>3</sup>School of Mechanical Engineering, Shanghai Jiaotong University, Shanghai 200240, China

<sup>4</sup>Shanghai Waigaoqiao Shipbuilding Co. Ltd., Shanghai 200137, China

<sup>5</sup>School of Management, Shanghai University, Shanghai 200444, China

Correspondence should be addressed to Bei-Xin Xia; [bxxia@shu.edu.cn](mailto:bxxia@shu.edu.cn)

Received 3 November 2018; Accepted 31 December 2018; Published 15 January 2019

Guest Editor: Xinchang Wang

Copyright © 2019 Ning-Rong Tao 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.

Special vehicles named flat transporters are used to deliver heavy ship assembly blocks in shipyards. Because each movement of assembly blocks among workshops needs transporters and the transportations are time-consuming, the scheduling of transporters is important for maintaining the overall production schedule of assembly blocks. This paper considers an optimization transporter scheduling problem for assembly blocks. The objective is to minimize logistics time, which includes empty travel time of transporters and waiting time and delay time of block tasks. Considering time windows of ship blocks, carrying capacity of transporters, and precedence relationships of tasks, a mathematical model is proposed. A hybrid topological graph is used to denote precedence and cooperating relationships of tasks. A metaheuristic algorithm based on the hybrid topological graph and genetic algorithm and Tabu search is proposed. The performance of the algorithm was evaluated by comparing the algorithm to optimal result in small-sized instances and several strategies in large-sized instances. The results showed the efficiency and effectiveness of the proposed algorithm.

## 1. Introduction

A ship hull is constructed by hundreds of assembly blocks. Each block is over 100 tons and 15 meters (length) by 15 meters (width) by 5 meters (height) [1, 2]. Some blocks weigh even more than 500 tons and are much bigger. Normally, before being assembled into a ship hull, a block would be moved over ten times among different workshops like assembly workshop, painting workshop, outfitting workshop, and stockyards for temporary storage. Because of considerable size and weight of blocks, a specially manufactured flat transporter (hereinafter, “transporter”) is used as the transportation vehicle, which is a multiple-axle vehicle with hydraulic jack lifts so it can jack up a big and heavy block.

In the transporter scheduling problem for assembly blocks, a transportation operation includes picking up a

block, moving it from one workshop/place to another through an accessible path, and unloading it at a proper position. Each transportation operation is quite time-consuming and costly. And since a few huge blocks may extend the load capacity of any available transporter in a shipyard, several transporters should work simultaneously for transporting it, which is called **cooperating transportation**. But lack of well scheduling of transporters causes delay on executing transporting tasks and waste on waiting for preassigned cooperating transporters. It would result in traffic jam and production delay in a shipyard. Actually, in most shipbuilding companies, one of the major management issues is how to efficiently manage assembly blocks in a shipyard, which contains utilizing efficiently the scarce resources such as the block stockyards and transportation vehicles. Ship block transportation problems are crucial issues to address in

reducing the construction cost and improving the productivity of shipyards [3].

Only a few recent studies have addressed the problem of assembly block transportation scheduling. Lee et al. [4] first studied the scheduling of single-type transporters for block transportation. Actually, there are various transporters according to the deadweight. Thus, Roh and Cha [5] studied the block transportation scheduling problem with multitype transporters. Their objective was to minimize the travel distance without loading and interference between transporters while satisfying the constraints on allowable transportation weights of the transporters. Kim and Joo [6] considered a similar block transportation scheduling problem for heterogeneous transporters with different weight capacities. Their objective was to minimize logistics times, including delay time, tardy time, and empty transporter travel time. Joo and Kim [7] expanded the research of reference paper [4] by considering the scheduling problem of block transportation under a delivery restriction to determine when and by which transporter each block is delivered from its source plant to its destination plant. Wang et al. [3] proposed a greedy algorithm with the same objective as Kim and Joo's researches.

Actually, in realistic shipbuilding environments, most workshops in shipbuilding have a schedule of blocks and the door is not big enough to afford several blocks come in/out at the same time. For example, the painting workshop has several sand washing rooms in which at most two blocks can be washed together. Usually they would move out in one day, and the back one must wait for the front one has been loaded by a transporter and leaves. And when there are two blocks in a workshop or in a temporary storage yard ready for transporting, one of them may be on the way of the other because of their big sizes. This is quite often in shipbuilding. In the above studies, the release/due time and the precedence relationship between the assembly blocks was not considered. Hu et al. [8] took account of the resource available time for use and tardiness minimization in a task assignment problem. Park and Seo [9] mentioned the precedence constraints between assembly blocks in their study. Zhang et al. [10] built a model of the assignment and paths of the inbound and outbound objects to shipyards considering the order of blocks. However, the research was just about single-type transporters. It may be not conformed to most shipbuilding companies, which needs to be studied on multitype transporters [7].

The block transportation scheduling is similar to a multiple travelling salesman problem with time windows (m-TSPTW) by regarding each block as a location and transporters as travelling salesmen [6, 11], when the type of transporters is the same and all blocks are predetermined to be delivered by a specific transporter. The m-TSPTW is known as a NP-hard problem. Laporte and Osman [12], Crainic and Laporte [13], and Chao [14] surveyed well known TSP, m-TSP, and general vehicle routing problems. Zhang and Moon [15] proposed an m-TSPTW to model a container truck transportation problem and developed a cluster method and a reactive Tabu search algorithm to solve the problem. Sterzik and Kopfer [16] proposed a Tabu search algorithm for

the inland container transportation problem to control the movement of full and empty containers.

The block transportation scheduling problem can also be transferred to a scheduling problem for parallel machines with sequence-dependent setup times and precedence constraints [17]. The parallel machine scheduling problem is one of the classical problems in production systems [18]. In our problems, the machines correspond to the transporter, and the job is the transport of an assembly block. This type of scheduling also belongs to the class of NP-hard problems [9].

However, there is a major difference between the above studies and our study, which is the situation of several transporters working simultaneously for transporting one block. The problem is kind of multivehicle and one-cargo transportation problem where a cargo is simultaneously loaded by several vehicles [19]. Dohn, Rasmussen, and Larsen [20] and Drexel [21] pointed out that synchronization problems are highly relevant in routing practice. Salazar-Aguilar, Langevin, and Laporte [22] introduced a synchronized arc routing problem for snow plowing operations. The street segments with two or more lanes in the same direction are plowed simultaneously by synchronized vehicles. Salazar-Aguilar, Langevin, and Laporte [23] introduced the paint vehicle synchronization problem that several capacitated vehicles painted lines on the roads with a tank vehicle replenishing the painting vehicles. The routes and schedules for the painting and tank vehicles were optimized by synchronizing the painting and replenishment operations. Rousseau, Gendreau, and Pesant [24] solved a real-time vehicle dispatching problem where some customers were serviced with multiple resources synchronously. Derigs and Pullmann [25] introduced the multidepot multitrip VRP with order incompatibilities subjected to interroute synchronization constraints. Hu and Wei [19] studied the multivehicle and one-cargo transportation problem; however, they did not put the problem into a realistic situation; for example, they did not mention priorities of tasks, release time, and due time of tasks or weight capacity of transporters.

On the basis of the above analysis, this paper deals with a transporter scheduling problem for assembly blocks. In the problem, we consider "multivehicle and one-cargo transportation" as well as the constraints of precedence relationship and the release/due time of tasks. The objective is to minimize total logistics time including empty travel time of transporters, delay time, and waiting time of block tasks.

## 2. Mathematical Model

*2.1. Problem Description.* The transporter scheduling problem for assembly blocks in a shipyard includes assigning all of block transportation tasks to multiple transporters, sequencing the tasks for each transporter, and determining the start time to fulfill each task.

The problem is studied based on the following conditions.

(1) Since each task would be released by one workshop at a certain time and needed by another workshop at some time, which means the task cannot be operated before its release time, and it is better to be finished before the due time, each task is given a time window [release time, due time].

TABLE 1: Block transportation tasks.

Task ID	Block ID	Weight (t)	Retrieve Place	Destination Place	Release Time (min)	Due Time (min)	Priority Task
1	B1	250	P1	P3	0	120	-
2	B2	450	P2	P3	30	150	-
3	B3	300	P3	P4	100	250	-
4	B4	200	P1	P2	0	120	1
5	B5	200	P3	P2	100	250	3
6	B6	350	P2	P3	0	180	-

TABLE 2: Preprocessed block transportation tasks.

Task ID	Block ID	Weight (t)	Retrieve Place	Destination Place	Release Time (min)	Due Time (min)	Priority Task	Synchronous Task
1	B1	250	P1	P3	0	120	-	-
2	B2'	<b>225</b>	P2	P3	30	150	-	<b>7</b>
3	B3	300	P3	P4	100	250	-	-
4	B4	200	P1	P2	0	120	1	-
5	B5	200	P3	P2	100	250	3	-
6	B6	350	P2	P3	0	180	-	-
7	<b>B2'</b>	<b>225</b>	<b>P2</b>	<b>P3</b>	<b>30</b>	<b>150</b>	-	<b>2</b>

(2) The precedence relationships between blocks are considered. The precedence constraint forced the one after started until the one before is loaded by the assigned transporter.

(3) Each transporter can take no more than one block at a time and must not exceed its deadweight.

(4) Each transportation operation must not be interrupted until it is finished.

(5) For the overweight block, it needs two transporters to work synchronously. The cooperating transporters must be the same type.

Here is an example with six transportation tasks in Table 1. Considering the precedence constrains, Task 4 cannot be started until Task 1 is loaded by a transporter, and Task 5 cannot be operated until Task 3 is loaded.

The weight capacities of three transporters are 250t, 250t, and 380t, respectively. It can be found from Table 1 that B2 weights 450t, exceeding the biggest weight capacity of transporters. Thus a feasible solution should assign two transporters to deliver B2 synchronously.

In this case, we create a virtual task. As in Table 2, the original Task 2 is divided into two synchronous tasks: Task 2 and Task 7.

Figure 1 gives an assignment and sequencing plan as an example, where Task 2 needs Transporters #01 and #02 cooperating to transport B2. Based on this sequencing and assignment plan together with the data in Table 2, the start time of each task can be determined.

**2.2. Model Formulation.** A mathematical programming model can be formulated to the problem. The following notations are used:

$n$ : total number of block transportation tasks

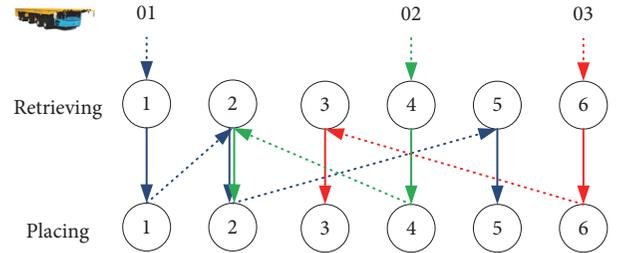


FIGURE 1: A sequencing and assignment plan for block transportation.

$m$ : total number of transporters

$w_i$ : block weight of block transportation task  $i$

$rt_i$ : release time of block transportation task  $i$

$dt_i$ : due time of block transportation task  $i$

$rp_i$ : retrieve place of block transportation task  $i$

$dp_i$ : destination place of block transportation task  $i$

$wc_k$ : weight capacity of transporter  $k$

$p_k$ : initial position of transporter  $k$

$vl_k$ : average speed of transporter  $k$  while transporting a block

$ve_k$ : average speed of transporter  $k$  while it is empty

$D = \begin{bmatrix} d(1,1) & \dots & d(1,p) \\ \vdots & \ddots & \vdots \\ d(p,1) & \dots & d(p,p) \end{bmatrix}$ : distance matrix of each pair of work places

$lt_i$ : prepare time needed for loading block in block transportation task  $i$

$ut_i$ : prepare time needed for unloading block in block transportation task  $i$

$pr_{ij}$ : precedence relationship of tasks  $i$  and  $j$ ; if  $pr_{ij} = 1$ , task  $j$  must wait at least until task  $i$  is loaded by the assigned transporter; if  $pr_{ij} = -1$ , task  $i$  must wait at least until task  $j$  is loaded by the assigned transporter; otherwise  $pr_{ij} = 0$

$sr_{ij}$ : cooperating relationship of tasks  $i$  and  $j$ ; if  $sr_{ij} = 1$ , task  $i$  must be operated with task  $j$  synchronously; otherwise  $sr_{ij} = 0$

$M$ : a big positive number

### Variables

$x_i$ : start time of picking up task  $i$

$em_i$ : empty travel time to pick up task  $i$

$dl_i$ : delay time of task  $i$

$y_i^k$ : if  $y_i^k = 1$ , task  $i$  is operated by transporter  $k$ ; otherwise  $y_i^k = 0$

$z_{ij}^k$ : if  $z_{ij}^k = 1$ , task  $j$  is operated by transporter  $k$  right before task  $i$ ; otherwise  $z_{ij}^k = 0$ .

$z_{0i}^k$ : if  $z_{0i}^k = 1$ , task  $i$  is the first task of transporter  $k$ ; otherwise,  $z_{0i}^k = 0$

**Objective Function.** The objective is to minimize total logistics time cost, defined as the weighted sum of empty travel time of transporters, delay time, and waiting time of block tasks.

$$\min f = w_1 \cdot \sum_{i \in n} em_i + w_2 \cdot \sum_{i \in n} dl_i + w_3 \cdot \sum_{i \in n} (x_i - rt_i) \quad (1)$$

**Constraints.** Each task must not be started after its release time  $rt_i$ . If the task is released, then the waiting time for being picked up is  $wa_i$ .

$$x_i \geq rt_i, \quad \text{for } \forall i \quad (2)$$

Constraints (3) give the empty travel time of each task. For task  $i$ , the empty travel time is the transporter assigned to task  $i$  moving from its last determined place to the retrieve place of task  $i$ . The last determined place of the transporter assigned to task  $i$  is  $\sum_{j \in n, j \neq i} z_{ji}^k \cdot dp_j$ , but if task  $i$  is the first task of transporter  $k$ , the last determine place is its initial place  $z_{0i}^k \cdot p_k$ .

$$em_i = d \left( \sum_{k \in m} z_{0i}^k \cdot p_k + \sum_{k \in m} \sum_{j \in n, j \neq i} z_{ji}^k \cdot dp_j, rp_i \right) \times \sum_{k \in m} \frac{y_i^k}{vl_k}, \quad \text{for } \forall i \quad (3)$$

Constraints (4) give the delay time, i.e., the time exceeding the due time of tasks. Wherein

$d(rp_i, dp_i) \times \sum_{k \in m} (y_i^k / vl_k)$  is with load travel time of task  $i$ ,

$$dl_i \geq x_i + lt_i + d(rp_i, dp_i) \times \sum_{k \in m} \frac{y_i^k}{vl_k} + ut_i - dt_i, \quad \text{for } \forall i \quad (4)$$

$$dl_i \geq 0, \quad \text{for } \forall i \quad (5)$$

For each pair of tasks, if one has a priority to the other, the latter one must be operated after the prior one is loaded by its transporter.

$$x_i - x_j + lt_i \geq M \cdot (1 - pr_{ij}), \quad \text{for } \forall i, j, i \neq j \quad (6)$$

If any two tasks are operated by the same transporter, their start time must follow Constraints (7), wherein  $d(dp_i, rp_j) \times \sum_{k \in m} (y_j^k / vl_k)$  is empty travel time from the last task  $i$  to task  $j$ .

$$x_i + lt_i + d(rp_i, dp_i) \times \sum_{k \in m} \frac{y_i^k}{vl_k} + ut_i + d(dp_i, rp_j) \times \sum_{k \in m} \frac{y_j^k}{vl_k} - x_j \leq M \cdot (1 - z_{ji}^k), \quad \text{for } \forall i, j, k \quad (7)$$

Constraints (8) ensure that each task can be executed only once.

$$\sum_{k \in m} y_i^k = 1, \quad \text{for } \forall i \quad (8)$$

Constraints (9) ensure that each transporter has at most only one first task.

$$\sum_{i \in n} z_{0i}^k \leq 1, \quad \text{for } \forall k \quad (9)$$

Constraints (10) ensure that each transporter can take only one task at a time.

$$\sum_{j \in n, j \neq i} z_{ij}^k + z_{0i}^k = y_i^k, \quad \text{for } \forall i, k \quad (10)$$

Constraints (11) ensure that each task must satisfy the assigned transporter's weight capacity.

$$\sum_{k \in m} y_i^k \cdot wc_k \geq w_i, \quad \text{for } \forall i \quad (11)$$

Constraints (12) ensure that tasks with cooperating constraint must be operated synchronously. And Constraints (13) ensure that the cooperating tasks must be operated by the same type of transporters.

$$sr_{ij} \cdot (x_i - x_j) = 0, \quad \text{for } \forall i, j \quad (12)$$

$$sr_{ij} \cdot \left( \sum_{k \in m} y_i^k \cdot wc_k - \sum_{l \in m} y_j^l \cdot wc_l \right) = 0, \quad \text{for } \forall i, j \quad (13)$$

The mathematical model is a mixed integer programming model. It can be solved by an optimization tool like CPLEX. But in realistic situations, the problem size makes it hard to

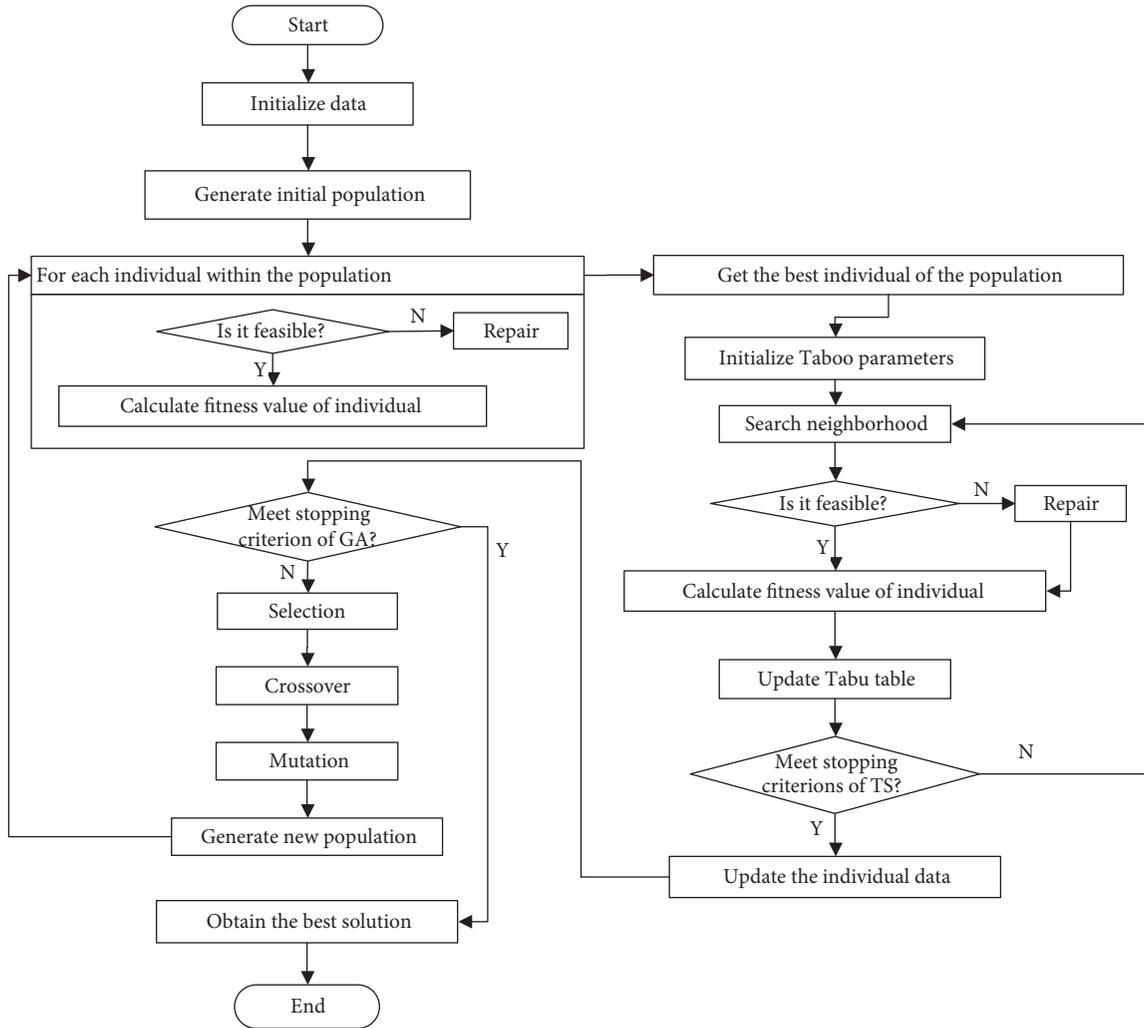


FIGURE 2: Flow chart of the proposed algorithm of block transportation scheduling.

get a good solution in reasonable computing time of CPLEX. Thus, we propose a metaheuristic algorithm to solve the transporter scheduling problem for assembly blocks.

### 3. Metaheuristic Algorithm

In this section, a basic flow of the proposed algorithm is first given. The proposed algorithm frame is based on genetic algorithm (GA), and the local search process uses Tabu search (TS). Several scheduling strategies are introduced and involved in the heuristic algorithm. Since the relationships among tasks become more complex as the problem size is growing, the algorithm cannot ensure the feasibility of solution of each chromosome. We propose a chromosome repair method through a topologic structure of task relationships.

**3.1. General Framework.** The main concept of the proposed metaheuristic algorithm is given in Figure 2.

The following are the main context of the proposed algorithm.

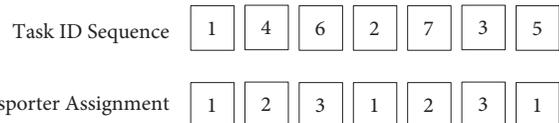


FIGURE 3: Chromosome representation in block transportation scheduling.

**3.1.1. Initial Population.** Usually, initial population can be randomly generated. The quality of initial population has some influence on the quality of algorithm. Thus, this paper takes half random generated individuals and the other half are obtained by several strategies which are described in Section 3.2.2.

**3.1.2. Encoding Scheme.** The genes of the chromosomes are designed as two arrays based on tasks assignment and sequencing information. The length of each chromosome is the number of tasks. Figure 3 is one chromosome of

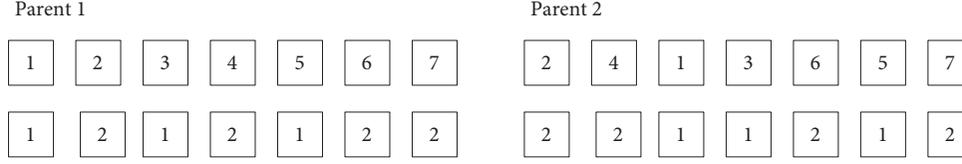


FIGURE 4: Two parents as an example.

the example shown in Table 2. It has two parts: one is the sequence of task ID, and the other one is transporter assignment information of each task. This sequence is not the order in which the tasks are executed by each transporter. For example, the first column (1,1) means task 1 is assigned to transporter 1 and it is scheduled first. The final order of tasks is obtained after decoding.

**3.1.3. Decoding Scheme.** The aim of decoding is to translate an individual to a full solution which can be evaluated by fitness function. The full solution is a worksheet telling the workers when and where each task is executed by which one or two transporters.

Based on the assignment and sequencing information in a chromosome together with precedence constraints, cooperating constraints, time window constraints etc., we can decode a chromosome by applying time scheduling strategy to the corresponding solution as follows.

*Step 1.* Obtain the data of sequencing and assignment list of tasks, and transporters' initial position and available time span, etc. Set counter  $h = 1$ ; then task  $i = \text{sequencing}(h)$ , and transporter  $k = \text{assignment}(h)$

*Step 2.* Decide the start time of task  $i$ .

*Step 2.1.* Check Matrix PR; obtain its priority-task set where  $ps_{ij} = -1$ ,  $j \in n$ .

*Step 2.2.* Obtain position  $p_k$  and release time  $et_k$  of transporter  $k$ .

*Step 2.3.* Calculate the start time  $st_i$  for task  $i$ :

$$st_i = \max \left\{ rt_i, et_k + \frac{d(p_k, rp_i)}{ve_k}, st_j + lt_j \right\}, \quad j \in tb_i \quad (14)$$

*Step 3.* Update new position  $p_k = dp_i$ ; the transporter's idle time  $et_k = st_i + lt_i + d(rp_i, dp_i)/vl_k + ult_i$ .

*Step 4.* Check Matrix SR; if  $sr_{ij} = 1$ ,  $j \in n$ , obtain task  $i$ 's cooperating task  $j$ . If Row  $i$  does not have any cooperating task, go to Step 5.

*Step 4.1.* If task  $j$  is executed before task  $i$ , then redefine the start time of these two tasks. Let  $st_i = st_j = \max(st_i, st_j)$ , with  $\Delta t = |st_i - st_j|$ ; redefine the related time, including  $et_k$ .

*Step 4.2.* If task  $j$  is executed after task  $i$ , go to Step 5.

*Step 5.* If  $h = n$ , the end, let  $h = h + 1$  and back to Step 2.

**3.1.4. Fitness Evaluation.** The objective function described in Section 2.1 is used to evaluate an individual's performance. The lower the value of the objective function, the higher the fitness of the individual.

**3.1.5. Selection.** Based on roulette selection strategy, two randomly individuals are selected from the population and the one with higher fitness value is retained. Repeat this process until obtaining required number of individuals. Suppose Figure 4 is two selected parents. Then the offspring are generated from them.

**3.1.6. Offspring Generation.** We divide the operators into two categories. Each of them contains crossover and mutation processes.

- (i) Assignment operators.
- (ii) Sequencing operators.

Assignment operators only change the assignment property of chromosomes; i.e., the sequencing of tasks is preserved in the offspring. Assignment crossover generates the offspring by exchanging the transporter assignment information of subchromosomes between two parents, as shown in Figure 5(a). The position of the subchromosome is randomly selected. And assignment mutation randomly swaps the transporter assignment information of two genes within a single parent, as shown in Figure 5(b).

Sequencing operators only change the sequence of the tasks in the parent chromosomes. The partially matched crossover (PMX), which is the best performing crossover operator for the scheduling problems [26, 27], is used. The procedures of PMX are listed as follows.

*Step 1.* Randomly select two positions on the parents.

*Step 2.* Exchange two subchromosomes between two parents.

*Step 3.* Determine the mapping relationship between two mapping sections; namely, map each gene in one mapping section with the corresponding gene in another mapping section on the same position.

*Step 4.* Repair the offspring by replacing the reduplicate genes beyond the mapping section according to the mapping relationship.

An example of PMX procedure is shown in Figure 6(a). After Step 2, some task ID may appear multiple times. So in Step 3, a mapping relationship is constructed and the repaired offspring are obtained finally.

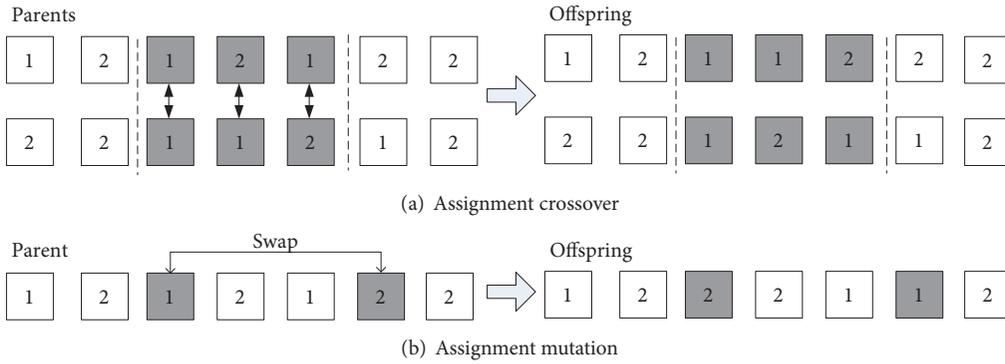


FIGURE 5: Example of assignment operators.

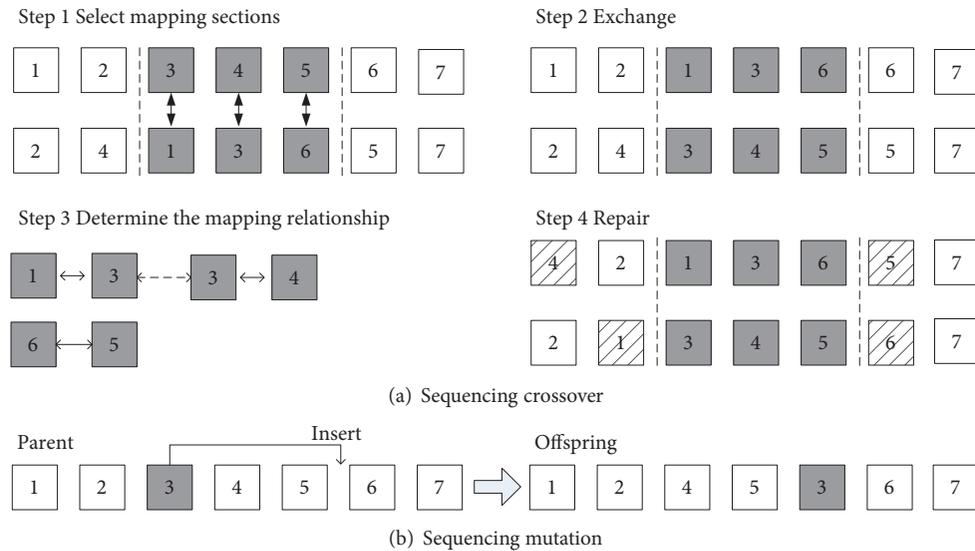


FIGURE 6: Example of sequencing operators.

And here an insertion operator acts as a mutation. As shown in Figure 6(b), it randomly selects a gene from the individual and inserts it back to a random position.

**3.1.7. Tabu Search Subalgorithm.** For the best individual in each population, use Tabu search subalgorithm to optimize it. It starts with an empty Tabu table. Two counters  $itr_1$  and  $itr_2$  are used to trace the number of total iterations and the number of iteration performed without improvement, respectively. The subalgorithm terminates when there is no improvement over the best solution obtained after a certain number of iterations ( $NonimpIter$ ), or when the total number of iterations reaches a predetermined value ( $MaxIter$ ).

Neighbourhood of an individual is searched by performing the following steps: (1) choose a gene at position; (2) transfer the transporter randomly to another one; (3) check the new individual's feasibility.

A Tabu table is a short time memory of the last several movements. It is applied to have a trace of the evolution of the search to prevent cycling. It is denoted by a  $n \times n$  matrix.

Record every transformation on the respective element in the matrix. The Tabu table is updated by keeping only the latest certain number of elements.

Suppose  $f_{best}$  is the best individual so far; set an empty Tabu table and two counters equal to 0. Repeat the following steps until one of the two counters reaches the present value.

*Step 1.* Set  $itr_1 := itr_1 + 1$ ,  $itr_2 := itr_2 + 1$ .

*Step 2.* Generate the neighbourhood of  $f_{best}$ .

*Step 3.* For each one of the neighbourhood  $f_{now}$ , calculate the fitness value. If the fitness value  $E(f_{now}) < E(f_{best})$ , then replace the individual  $f_{best}$  by  $f_{now}$  and set  $itr_2 = 0$ .

*Step 4.* Update the Tabu table by recording the neighbourhood search. Go to Step 1 until  $itr_1 = MaxIter$  or  $itr_2 = NonimpIter$ .

**3.1.8. Stopping Criterion.** The algorithm ends when the maximal number of generations ( $GAMaxIter$ ) is reached, and the best individual, together with the corresponding scheduling

and assignment solutions, is given as output. The solution after decoding the individual chromosome is the optimal solution.

**3.2. Strategies Based on Topological Relationships.** The problem actually includes three parts, which are assigning the tasks to a special transporter, sequencing the tasks for each transporter, and determining the start time of each task. The relationship constraints among tasks are presented by a topological structure. The topological map would be changed along with solving process. Then several strategies gained from experience are presented for initial solutions.

**3.2.1. Topological Description for Tasks.** Topological structure is one of the basic methods for scheduling problem. Like most scheduling problems, the precedence constraint is represented by one-way arc. Besides precedence constraints, cooperating constraints are considered in the problem. Here, we use undirected arcs to represent cooperating constraints between tasks. Thus, the relationships among tasks in Table 2 can be represented in Figure 7. For any pair of two tasks which is linked by a one-way arc, the before task should be loaded before the after one. While for any pair of two tasks which is linked by undirected arc, the two tasks must be operated synchronously. These relationships must be fully met in order to guarantee a solution is feasible.

Matrixes  $PR$  and  $SR$  are used to express the topological relationships of tasks. For the example in Table 2,  $pr_{14} = 1$  means Task 1 is prior to Task 4; oppositely,  $pr_{41} = -1$  and, for the same reason,  $pr_{35} = 1$ ,  $pr_{53} = -1$ ;  $sr_{27} = 1$  means Task 2 must be operated synchronously with Task 7, and  $sr_{72} = 1$ .

$$PR = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (15)$$

$$SR = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

**3.2.2. Strategies for Initial Solution.** The problem actually includes three parts, which are assigning the tasks to a special transporter, sequencing the tasks for each transporter, and determining the start time of each task. The topological map would be changed along with solving process. The above two matrixes present the initial relationships of the problem.

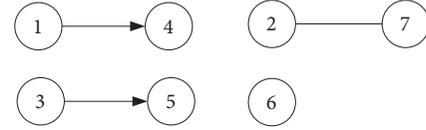


FIGURE 7: Topological map of block transportation tasks.

Referring to the matrixes, a negative element in a row of matrix  $PR$  means the cooperating task has a prior task which is not assigned yet. When Task 1 has started, delete Row 1 and make the negative elements in Column 1 to 0. Then, in the new matrix, Row 4 has no negative element. That means Task 4 is executable.

When assigning transporters, there are several rules considered here.

**Rule 1.** Select the transporter which reaches the retrieving place as early as possible.

**Rule 2.** Select the transporter which is idle firstly.

The strategies of initial solution are as follows.

**Step 1.** Choose a row without zero from  $PR$ , and add the cooperating task to executable task set.

**Step 2.** Select a task  $i$  from executable task set which has the earliest release time.

**Step 2.1.** Select a transporter which meets the task's weight constraint. If there is more than one selections, use Rule 1 or Rule 2 to choose one from them.

**Step 2.2.** Calculate the start time of task  $i$ ,  $x_i = \max(t_k + d(p_k, rp_i)/ve_k, rt_i)$ .

**Step 3.** Check Matrix  $SR$ . If task  $i$  does not have any cooperating task, go to Step 4; otherwise repeat Step 2.1 and Step 2.2 for the cooperating task  $j$ , calculate the start time  $x_j = \max(t_{k'} + d(p_{k'}, rp_j)/ve_{k'}, rt_j)$ , and let  $x_i = x_j = \max(x_i, x_j)$ .

**Step 4.** Update the transporter's idle time  $et_k = x_i + lt_i + d(rp_i, dp_i)/vl_k + ut_i$  and new position  $p_k = dp_i$ .

**Step 5.** Delete task  $i$  from executable task set and change all the negative value in column  $i$  be zero in Matrix  $PR$ ; go back to Step 1.

We take the example in Section 2.1; suppose the distance matrix  $D$  is given as follows.

$$D = \begin{bmatrix} 0 & 400 & 500 & 700 \\ 400 & 0 & 800 & 600 \\ 500 & 800 & 0 & 300 \\ 700 & 600 & 300 & 0 \end{bmatrix} \quad (16)$$

In Step 1, executable task set = {1, 2, 3, 6, 7}. We first select Task 1 and FT-1, calculate the start time of operating Task 1,

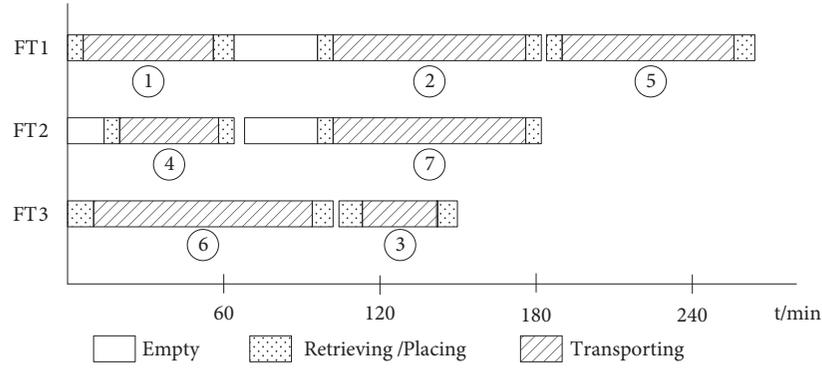


FIGURE 8: Gantt chart of a solution of block transportation scheduling.

i.e.,  $x_1$ , and get the idle time of FT-1, i.e.,  $et_1$ . Then, delete Task 1 from the set. Since  $pr_{14} = 1$ , release Task 4 and update the set= $\{2, 3, 6, 7, 4\}$ . Go to Step 2 and execute Task 6 and Task 4. Select Task-2. Since  $sr_{27} = 1$ , select Task-7 too. Assign FT-1 and FT-2 to Task-2 and Task-7. Calculate the respective start times  $x_2$  and  $x_7$  and choose  $\max(x_2, x_7)$  as the start time. Then, execute Task 3 and Task 5. Finally, a feasible solution is obtained. Figure 8 gives the Gantt chart of the solution.

3.3. Feasibility Analysis and Gene Repair. As we know, GA needs appropriate constraint handling method to get feasibility and applicability result. The most common methods are penalty function method and chromosome repair method. The penalty function method is hard to get feasible solution for complex constraints problem, while chromosome repair method is more appropriate by repair infeasible solutions. Thus, we use a chromosome repair method to repair the unfeasible offspring.

3.3.1. Feasibility Analysis. When dealing with complex constrained optimization problems, GA needs to take action to maintain the applicability and feasibility. Feasibility analysis is based on the constraints described in the mathematical model, which contains two main parts:

- (i) whether satisfying the load transporters' capacity

It is checked through Constraints (11):  $\sum_k y_i^k \cdot wc_k \geq w_i$ , for  $\forall i$ :

- (ii) whether satisfying the task relationships

An individual chromosome can be transferred into a topological diagram. The diagram is structured by nodes and directed and undirected lines. The following are the structure process:

- (i) Nodes represent tasks. In Figure 9, ①-⑦ are seven tasks.
- (ii) For two tasks with precedence relationship, use solid directed line.
- (iii) For two tasks with cooperating relationship, use solid undirected line.

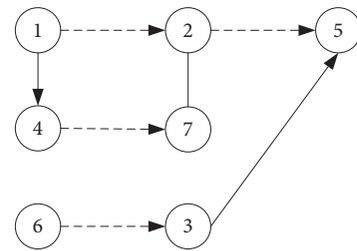


FIGURE 9: Topological diagram for the scheduling of block transportation tasks ( $n=7$ ).

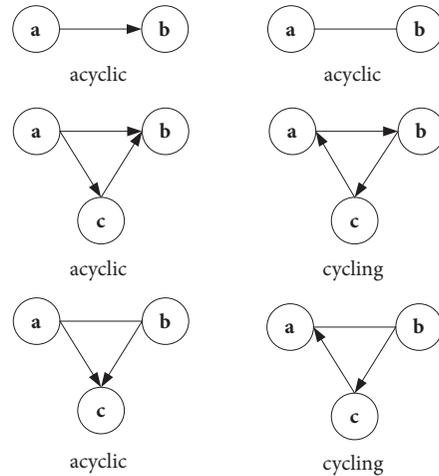


FIGURE 10: Rules to judge the feasibility of hybrid graphs.

- (iv) For transporting list of transporters, use dashed directed line.

Usually, a directed acyclic graph means the plan is feasible. Since this problem has cooperating tasks, a solution can be expressed as a hybrid graph. So to determine the feasibility of a solution, new judgment rules are presented in Figure 10. The acyclic graph means the solution is feasible, while the cycling graph means it is unfeasible.

Based on the rules, we can find that the solution in Figure 9 is feasible. To explain the rules, an example with 10

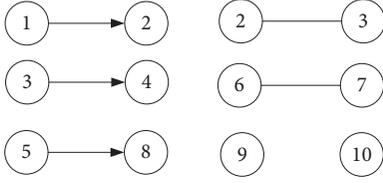


FIGURE 11: Topological diagram for the precedence and cooperating tasks ( $n=10$ ).

Task sequence 1 6 4 9 2 5 3 10 7 8

Transporter Assignment 1 2 4 3 2 1 3 2 4 1

FIGURE 12: A chromosome of block transportation scheduling ( $n=10$ ).

tasks and 4 transporters is given. The relationships among tasks are  $1 \rightarrow 2$ ,  $3 \rightarrow 4$ ,  $5 \rightarrow 8$ , and  $2 = 3$ ,  $6 = 7$ , which can be described in Figure 11.

Suppose one chromosome is given in Figure 12.

Based on this chromosome, the solution can be converted into a topological diagram (see Figure 13).

We can find that there is a cycle in Figure 13. The cycle is shown in Figure 14. So it can be judged that the chromosome is unfeasible.

**3.3.2. Chromosome Repair.** If unfeasibility is caused by transporter overweight, reassign the task to a capable transporter. The process is relatively simple, so we do not detail it anymore.

If unfeasibility is caused by the conflict to topological relationships that means there is a topological cycle, we should change the transporting list to repair it.

Take Figure 13 as an example to detail the repair steps. As shown in Figure 15, within the topological cycle, there are two the dashed arcs  $6 \rightarrow 2$  and  $4 \rightarrow 7$ . Choose one of them to break. In Figure 15(a), arc  $4 \rightarrow 7$  is broken. Then, Task 7 is temporarily unscheduled. So, assign Task 7 to another position and a new dashed arc is build, as shown in Figure 15(b). Then check the feasibility of new topological relationships again until there is no cycle, and that means the chromosome is feasible.

## 4. Computational Results

To verify the practicability and efficiency of the proposed algorithm for transporter scheduling problem considering precedence and cooperating constraints, different numerical simulations are tested and evaluated according to realistic situations.

The generating conditions are as follows.

- (i) Planning horizon  $T$  is 480 minutes
- (ii) Distances between two workshops/stock yards  $d = \text{Uniform}(200, 2000)$  in meters

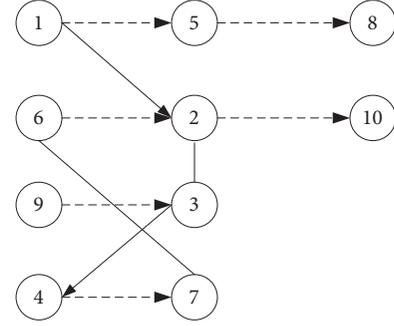


FIGURE 13: Topological diagram for the scheduling of block transportation tasks ( $n=10$ ).

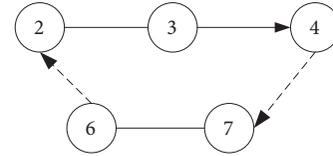


FIGURE 14: Topological cycle in the chromosome.

- (iii) Release time of tasks  $rt_i = \text{Uniform}(1, 360)$  in minutes
- (iv) Due time of tasks  $dt_i = \text{Uniform}(rt_i + 60, 480)$  in minutes
- (v) Weight of blocks are randomly generated from  $\text{Uniform}(100, 250)$  in tons
- (vi) 20% pair of tasks have precedence or cooperating relationships
- (vii) Loading and unloading time of blocks are randomly generated from  $\text{Uniform}(10, 20)$  in minutes
- (viii) Two flat transporters types are considered: 1# transporter: weight capacity is 150 tons; speed with load is 60m/min; speed without load is 100 m/min. 2# transporter: weight capacity is 200 tons; speed with load is 50m/min; speed without load is 80 m/min. The number of each type is  $m_1, m_2$

This is a multiobjective optimization problem. In Formula (1), weights  $w_1$ ,  $w_2$  and  $w_3$ , are set depending on actual demand. Representative results are summarized under  $w_1 = 0.7$   $w_2 = 0.2$   $w_3 = 0.1$ .

In Figure 16, the test is made with  $n = 30$  and  $m = 10$ . The dotted line denotes general genetic algorithm (GA), and the solid line denotes the proposed algorithm (GA-TS). After 500 iterations, the general GA did not converge, but GA-TS converged before 50 iterations. It is obvious that GA-TS converges much more speedily. Compared with the general genetic algorithm, the results show that the proposed algorithm has faster convergence speed.

In Figure 17, Tabu iterations are 0, 10, 20, 30, 40, and 50. After Tabu search, solution quality improves significantly. But when iterations exceed 10, the curve becomes stable. Thus, we select 10 iterations in the following tests.

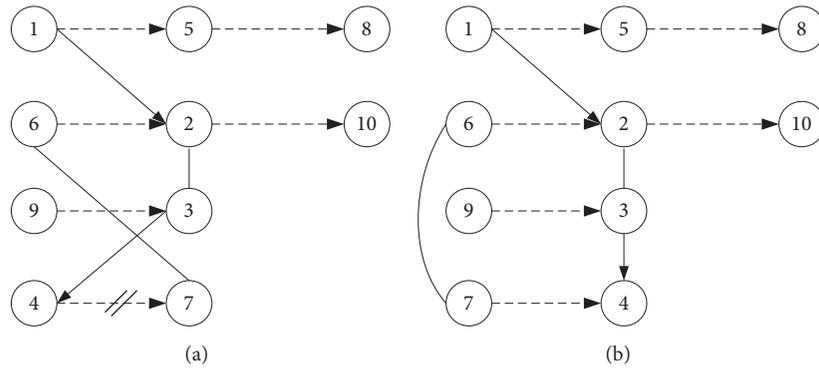


FIGURE 15: Chromosome repair.

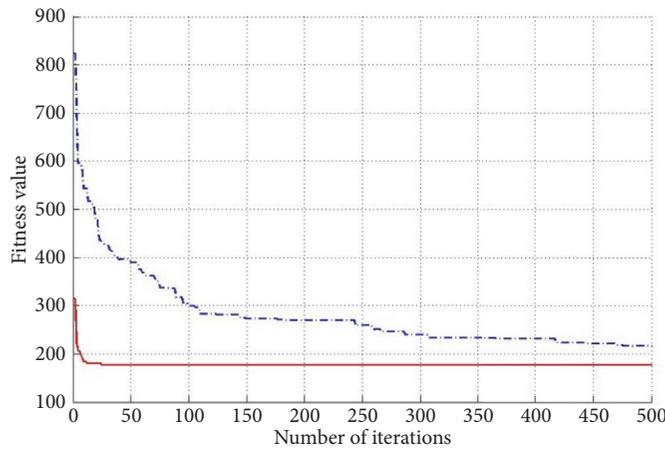


FIGURE 16: Algorithm convergence charts of GA and GA-TS.

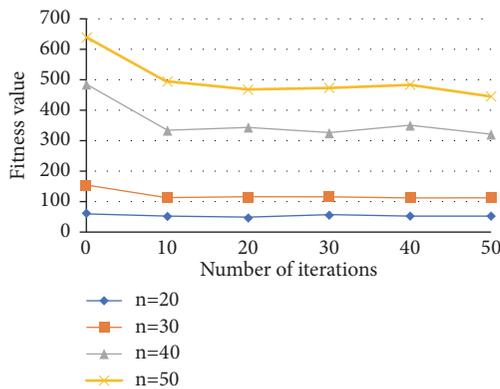


FIGURE 17: The objective values under different Tabu search iterations.

And with other pertaining to numerical tests, the algorithm parameters in this paper are determined as follows:

- (i) Population size: 20
- (ii) Number of generations: 100
- (iii) Assignment crossover probability: 0.8
- (iv) Assignment mutation probability: 0.1

- (v) Sequencing crossover probability: 0.6
- (vi) Sequencing mutation probability: 0.1
- (vii) number of iterations: 20
- (viii) number of iteration performed without improvement: 10
- (ix) length of Tabu table:  $(m1+m2)/2$

To evaluate the heuristic strategies of initial solutions, the problems are tested with three sets of block transportation tasks  $n$  as 20, 30, and 40, respectively, and the number of transporters is 4+4. Each test runs ten times to get an average result. “H1” is the heuristic strategy following the steps in Section 3.2.2 with Rule 1, while “H2” is the heuristic strategy following the steps in Section 3.2.2 with Rule 2. The result is presented in Table 3. We can find that H1 with Rule 1 gets less idle time of transporters, but H2 with Rule 2 gets better performance on time that related to production. Overall, we can choose a rule based on actual optimization requirements. In this paper, the scheduling scheme from these two heuristic rules is used as 60% of the initial population of the genetic algorithm to improve the efficiency of genetic algorithm.

For small-sized problems, we generate 5 instances shown in Table 4. The data of assembly blocks are randomly generated. The experiments are performed on a PC with Inter

TABLE 3: Comparison result of strategy rules (unit: min).

$n$	H1				H2			
	idle	delay	wait	$f$	idle	delay	wait	$f$
20	65.7	647.3	1054.8	511.2	222.9	209.5	417.3	405.9
30	99.9	1500.7	2089.7	591.4	313.1	324.9	569.3	426.6
40	168.1	3225.1	4262.8	741.4	466.3	801.1	1384.8	470.5

TABLE 4: Test results in small-sized instances.

$n$	$m$	CPLEX		GA-TS		GAP
		Opt (min)	Time (s)	$f$ (min)	Time (s)	
6(1,1)	2	22.5	9.83	22.5	12.78	0
8(1,1)	2	34.4	944.26	34.4	28.43	0
8(1,1)	3	20.7	970.76	21.7	31.27	5%
8(2,1)	3	20.7	1350.81	21.4	50.13	3%
10(1,1)	2	56.3	3600	60.2	45.75	6%
10(2,2)	3	32.6	3600	33.5	62.21	3%

TABLE 5: Test results in large-sized instances.

$n$	$m$	Objective value (min)			
		H1	H2	GA	GA-TS
20	8	338.1	262.9	337.8	161.2
30	8	398.8	290.2	513.3	195.7
40	8	549.9	566.7	620.3	288.6
50	8	2005.2	1266.8	2052.5	866.2
20	10	118.8	206.1	232.2	94.9
30	10	231.7	368.8	459.8	170.1
40	10	633.2	345.7	430.1	226.6
50	10	805	494.4	1143.8	322.6
30	12	362	480	433.9	211.6
40	12	583.5	476.2	538.6	301.8
50	12	1104.4	460.8	834	180

Pentium 4, 3.01GHz, and 2GB RAM. We use ILOG CPLEX for finding the optimal solutions with the mathematical programming presented in Section 2.2. The maximum calculate time of CPLEX is set to 3600s. The codes of proposed algorithm are written in Matlab.

For large-sized problems, we generate 11 instances shown in Table 5. We compare the proposed solution (GA-TS) with the two heuristics strategies (H1 and H2) and a GA algorithm for each test problem.

Table 4 gives the test results of small-sized instances. It shows the optimal solution and computation time of CPLEX and the relative performance of the proposed algorithm for each test instance. In the table,  $n=6(1,1)$  means the problem has six tasks. When the transporters are more than two, two types are involved. The first "1" within the parentheses means there is one pair of tasks that has precedence constraint, and the other "1" means one of the six tasks is overweight. GAP is used to evaluate the performance of the proposed algorithm, where  $GAP = ((f - opt)/f) \times 100\%$ . As the problem size is increasing, the computing time of CPLEX increases dramatically. For some instances with larger problem size, the CPLEX

cannot get the optimal solution within 3600s. Comparing the proposed algorithm with CPLEX, the proposed algorithm obtained optimal solutions in several problems, and when its solution is not optimal, the results are closer to the optimum one. Besides, the computing time did not increase rapidly. The number of precedence and cooperating constraints increases the difficulty; however, the proposed algorithm can get a satisfying solution while the computing time of the proposed algorithm does not increase a lot. The experimental results illustrate the efficiency of the proposed algorithm.

Table 5 gives the test results of large-sized instances. In these instances, the precedence and cooperating relationships are generated randomly. Within Table 5, H1 and H2 are two heuristic solutions gained by the experimental strategies which are obtained from experience. GA is designed as a general genetic algorithm flow. The results of GA are gained after 500 times iterations. We can see that results of GA are not stable even after 500 iterations; some result is worse than the solution gained by the experience strategies. While comparing the proposed algorithm (GA-TS) with the three algorithms, GA-TS obtained much better results. It is due to

using the strategies in initial solution and letting Tabu search process during each generation of GA. As the size of problem increases, the advantage of GA-TS becomes more and more obvious.

## 5. Conclusion

This paper researched a scheduling problem for block transportation with multitype transporters under precedence relationships and cooperating constraints. The problem includes three parts: assigning all of block transportation tasks to multiple transporters, sequencing the tasks for each transporter, and determining the start time to fulfill each task. The objective is to minimize logistics time, which includes empty travel time of transporters, waiting time, and delay time of block tasks. A MIP model was proposed for the problem. A hybrid topological graph was used to denote precedence and cooperating relationships of tasks, and a metaheuristic algorithm based on genetic algorithm (GA) and Tabu search (TS) was proposed. The performance of the algorithm was evaluated by comparing the algorithm to CPLEX in small-sized instances and two experience strategies and ordinary GA in large-sized instances. The results showed the efficiency and effectiveness of the proposed algorithm.

In this paper, it is assumed that the transportation tasks are executed smoothly on the road in shipyards. In reality, there must be other tasks executed on the same time which may cause the task to not be executed as plan and road traffic jam may happen. Therefore, it would be interesting to study the dynamic demands in the problem further.

## Data Availability

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

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The research is supported by National Natural Science Foundation of China under Projects (no. 71501125 and no. 71401098), Shanghai Pujiang Program (17PJJC051), and MIIT Research Project on High Technology of Ships named “Intelligent Shipyard Top-Level Architecture and Application of Production Logistics”.

## References

- [1] C. Park and J. Seo, “Comparing heuristic algorithms of the planar storage location assignment problem,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 46, no. 1, pp. 171–185, 2010.
- [2] Z. Zhang and J. Chen, “Solving the spatial scheduling problem: a two-stage approach,” *International Journal of Production Research*, vol. 50, no. 10, pp. 2732–2743, 2012.
- [3] C. Wang, Y.-S. Mao, B.-Q. Hu, Z.-J. Deng, and J. G. Shin, “Ship block transportation scheduling problem based on greedy algorithm,” *Journal of Engineering Science and Technology Review*, vol. 9, no. 2, pp. 93–98, 2016.
- [4] W. S. Lee, W. L. Lim, P. H. Koo, and C. M. Joo, “Transporter scheduling for block transportation in shipbuilding,” *Journal of the Korea Management Engineers Society*, vol. 11, no. 1, pp. 169–179, 2006 (Portuguese).
- [5] M.-I. Roh and J.-H. Cha, “A block transportation scheduling system considering a minimisation of travel distance without loading of and interference between multiple transporters,” *International Journal of Production Research*, vol. 49, no. 11, pp. 3231–3250, 2011.
- [6] B. S. Kim and C. M. Joo, “Ant colony optimisation with random selection for block transportation scheduling with heterogeneous transporters in a shipyard,” *International Journal of Production Research*, vol. 50, no. 24, pp. 7229–7241, 2012.
- [7] C. M. Joo and B. S. Kim, “Block transportation scheduling under delivery restriction in shipyard using meta-heuristic algorithms,” *Expert Systems with Applications*, vol. 41, no. 6, pp. 2851–2858, 2014.
- [8] H. Hu, Y. Wu, and T. Wang, “A Metaheuristic Method for the Task Assignment Problem in Continuous-Casting Production,” *Discrete Dynamics in Nature and Society*, vol. 2018, Article ID 8073648, 12 pages, 2018.
- [9] C. Park and J. Seo, “A GRASP approach to transporter scheduling and routing at a shipyard,” *Computers & Industrial Engineering*, vol. 63, no. 2, pp. 390–399, 2012.
- [10] Z.-Y. Zhang, J.-X. Xu, and F. Ji, “Shipbuilding yard scheduling approach based on genetic algorithm,” *Shanghai Jiaotong Daxue Xuebao/Journal of Shanghai Jiaotong University*, vol. 47, no. 7, pp. 1036–1042, 2013.
- [11] L. Zhen, E. P. Chew, and L. H. Lee, “An integrated model for berth template and yard template planning in transshipment hubs,” *Transportation Science*, vol. 45, no. 4, pp. 483–504, 2011.
- [12] G. Laporte and I. H. Osman, “Routing problems: a bibliography,” *Annals of Operations Research*, vol. 61, no. 1, pp. 227–262, 1995.
- [13] T. G. Crainic and G. Laporte, *Fleet Management And Logistics*, Kluwer Academic Publishers, Dordrecht, Mass, USA, 1998.
- [14] I.-M. Chao, “A tabu search method for the truck and trailer routing problem,” *Computers & Operations Research*, vol. 29, no. 1, pp. 33–51, 2002.
- [15] R. Y. Zhang, W. Y. Yun, and I. Moon, “A reactive tabu search algorithm for the multi-depot container truck transportation problem,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 45, no. 6, pp. 904–914, 2009.
- [16] S. Sterzik and H. Kopfer, “A tabu search heuristic for the inland container transportation problem,” *Computers & Operations Research*, vol. 40, no. 4, pp. 953–962, 2013.
- [17] R. Tavakkoli-Moghaddam, F. Taheri, M. Bazzazi, M. Izadi, and F. Sassani, “Design of a genetic algorithm for bi-objective unrelated parallel machines scheduling with sequence-dependent setup times and precedence constraints,” *Computers & Operations Research*, vol. 36, no. 12, pp. 3224–3230, 2009.
- [18] L. Zhen, D. Zhuge, and S. Zhu, “Production stage allocation problem in large corporations,” *Omega*, vol. 73, pp. 60–78, 2017.
- [19] Z.-H. Hu and C. Wei, “Synchronizing vehicles for multi-vehicle and one-cargo transportation,” *Computers & Industrial Engineering*, vol. 119, pp. 36–49, 2018.

- [20] A. Dohn, M. S. Rasmussen, and J. Larsen, “The vehicle routing problem with time windows and temporal dependencies,” *Networks*, vol. 58, no. 4, pp. 273–289, 2011.
- [21] M. Drexl, “Synchronization in vehicle routing—a survey of VRPs with multiple synchronization constraints,” *Transportation Science*, vol. 46, no. 3, pp. 297–316, 2012.
- [22] M. A. Salazar-Aguilar, A. Langevin, and G. Laporte, “Synchronized arc routing for snow plowing operations,” *Computers & Operations Research*, vol. 39, no. 7, pp. 1432–1440, 2012.
- [23] M. A. Salazar-Aguilar, A. Langevin, and G. Laporte, “The synchronized arc and node routing problem: Application to road marking,” *Computers & Operations Research*, vol. 40, no. 7, pp. 1708–1715, 2013.
- [24] L.-M. Rousseau, M. Gendreau, and G. Pesant, “The synchronized dynamic vehicle dispatching problem,” *INFOR. Information Systems and Operational Research*, vol. 51, no. 2, pp. 76–83, 2013.
- [25] U. Derigs, M. Pullmann, and U. Vogel, “Truck and trailer routing—problems, heuristics and computational experience,” *Computers & Operations Research*, vol. 40, no. 2, pp. 536–546, 2013.
- [26] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, USA, 1989.
- [27] H. Luo, G. Q. Huang, Y. Zhang, Q. Dai, and X. Chen, “Two-stage hybrid batching flowshop scheduling with blocking and machine availability constraints using genetic algorithm,” *Robotics and Computer-Integrated Manufacturing*, vol. 25, no. 6, pp. 962–971, 2009.

## Research Article

# Distance-Based Congestion Pricing with Day-to-Day Dynamic Traffic Flow Evolution Process

Qixiu Cheng,<sup>1</sup> Jiping Xing,<sup>1</sup> Wendy Yi,<sup>2</sup> Zhiyuan Liu ,<sup>1</sup> and Xiao Fu <sup>3</sup>

<sup>1</sup>Jiangsu Key Laboratory of Urban ITS, Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University, China

<sup>2</sup>School of Engineering and Advanced Technology, College of Sciences, Massey University, Auckland, New Zealand

<sup>3</sup>School of Transportation, Southeast University, China

Correspondence should be addressed to Zhiyuan Liu; zhiyuanl@seu.edu.cn

Received 19 September 2018; Accepted 19 December 2018; Published 6 January 2019

Academic Editor: Lu Zhen

Copyright © 2019 Qixiu Cheng 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 the distance-based congestion pricing in a network considering the day-to-day dynamic traffic flow evolution process. It is well known that, after an implementation or adjustment of a new congestion toll scheme, the network environment will change and traffic flows will be nonequilibrium in the following days; thus it is not suitable to take the equilibrium-based indexes as the objective of the congestion toll. In the context of nonequilibrium state, prior research proposed a mini-max regret model to solve the distance-based congestion pricing problem in a network considering day-to-day dynamics. However, it is computationally demanding due to the calculation of minimal total travel cost for each day among the whole planning horizon. Therefore, in order to overcome the expensive computational burden problem and make the robust toll scheme more practical, we propose a new robust optimization model in this paper. The essence of this model, which is an extension of our prior work, is to optimize the worst condition among the whole planning period and ameliorate severe traffic congestions in some bad days. Firstly, a piecewise linear function is adopted to formulate the nonlinear distance toll, which can be encapsulated to a day-to-day dynamics context. A very clear and concise model named logit-type Markov adaptive learning model is then proposed to depict commuters' day-to-day route choice behaviors. Finally, a robust optimization model which minimizes the maximum total travel cost among the whole planning horizon is formulated and a modified artificial bee colony algorithm is developed for the robust optimization model.

## 1. Introduction

Congestion toll is generally regarded as a potent economic instrument for transportation demand management (TDM) to alleviate the traffic congestion and improve the system performance in urban areas and also has received more and more attention both academically and practically. Since the successful implementation of congestion pricing in Singapore from 1975, many countries and cities (such as Norway, London, Stockholm and Milan) have implemented a road congestion pricing policy, which has achieved remarkable success in terms of easing urban traffic congestion [1]. However, all of the existing and implemented congestion toll schemes adopt a unique toll method, which makes the inequitable and ineffective problem. Therefore, in order to give full play to

congestion pricing in alleviating urban traffic congestion and improve the fairness and effectiveness of congestion pricing, it is necessary to consider the travel distance inside the charging cordon and establish the distance-based congestion toll scheme.

The distance-based toll scheme receives more and more attention recently in both the academic community and industrial circles. References [2–4] assumed that the distance-based toll is linearly proportional to the distance traveled in the charging area. However, it may be more efficient and effective in using the nonlinear toll function in the practical congestion toll scheme according to some recent studies. Lawphongpanich and Yin [5] assumed that the distance-based toll is a nonlinear form and used a piecewise linear function with two intervals to represent it. References [1, 6–9]

studied the distance-based toll using the piecewise linear toll function with multiple intervals to approximate the nonlinear toll function. It is worth noting that the next generation of road pricing systems in Singapore will adopt the distance-based toll scheme based on global navigation satellite system (GNSS) technology [10].

Generally, the total travel cost is regarded as the optimization objective of the congestion toll design problem. Most of the literatures studying the congestion pricing problem calculate the total travel cost (TTC) based on the equilibrium flows and make an evaluation based on the calculated TTC. However, when a toll pattern is implemented, the route flows will be totally different from day to day, because the implemented toll policy is an important component which will influence travelers' route choice behaviors. The system cannot reach an equilibrium state overnight. Therefore, for the optimal toll design problem during the whole planning period, the day-to-day dynamics models can better describe the network flow conditions, rather than the final equilibrium state. Besides, to avoid the complicated implementation of governments and confusions of travelers on the toll in practice, it is necessary to levy an unchanged toll in the whole period  $D$ ; for instance, Singapore's electronic road pricing (ERP) toll is adjusted every three months [4, 5] and kept unchanged in-between; thus  $D$  can be set as three months in this study.

During the whole period of  $D$ , the TTC will change from day to day because the traffic flows will change from day to day. Therefore, no toll pattern can give rise to a minimal TTC in all days of  $D$ . Liu et al. [7] proposed a mini-max regret model to solve the day-to-day dynamic congestion pricing (DCP) problem. However, this model is computationally demanding due to the calculation of minimal total travel cost for each day among the whole planning horizon. Therefore, in order to overcome the expensive computational burden problem and make the robust toll scheme more practical, we can use the mini-max total travel cost model to replace the mini-max regret model. The essence of this model, which is an extension of our previous work, is to optimize the worst condition among the whole planning period and ameliorate severe traffic congestions in some bad days.

The day-to-day dynamic flow evolution process is the foundation for the day-to-day DCP problem; lots of research work focus on the day-to-day dynamic flow evolution process [6–14]. As for the DCP, Wie and Tobin [15] considered the day-to-day DCP problem and used a convex control model to solve it. Friesz et al. [16] studied the day-to-day DCP aiming at maximizing the net present value of social welfare. More recently, Guo et al. [17] studied the dynamic tolls on each day and they are merely determined by the flows and tolls on the previous day. Ye et al. [18] studied the marginal-cost pricing scheme considering the day-to-day dynamics based on the trial-and-error method. Tan et al. Reference [19] studied the day-to-day DCP problem aiming at minimizing the total system cost and time. However, all of the aforementioned studies focus on deterministic day-to-day DCP problem. Recently, Rambha and Boyles [20] studied the stochastic day-to-day DCP problem. However, the objective

function in Rambha and Boyles [20] does not consider the network performances of each day. Cheng et al. [21] made a comprehensive review of urban dynamic congestion pricing and highlighted that there was an emerging research need to investigate the DCP problem.

The contributions of this paper are twofold. On the one hand, a finite learning process model named logit-type Markov adaptive learning model is proposed to depict commuters' day-to-day route choice behaviors. On the other hand, a mini-max total travel cost model, which can overcome the expensive computational burden problem in previous work and make the congestion toll scheme more practical, is proposed to solve the congestion toll problem considering nonequilibrium flow evolution processes. This paper is structured as follows. The next section first introduces the nonlinear distance toll which can be approximated with a piecewise linear toll function. A logit-type Markov adaptive learning model is then proposed in Section 3. Afterwards, a mini-max model for the optimal toll pattern that minimizes the maximum total travel cost among the whole planning horizon is introduced in Section 4, and a modified artificial bee colony (ABC) algorithm is developed for the robust optimization model in Section 5. Finally, conclusions are drawn in Section 6.

## 2. Problem Statement

A strongly connected network, denoted by  $G = (N, A)$ , is considered.  $N$  denotes the set of nodes and  $A$  denotes the set of directed links.  $W$  denotes the set of origin-destination (OD) pairs, and  $R^w$  denotes the set of routes connecting an OD pair  $w \in W$ .  $q^w$  is the traffic demand connecting the OD pair  $w \in W$ , and  $\mathbf{q} = (q^w, w \in W)^T$ .  $f_{wr}$  is the traffic flow on route  $r \in R^w$  connecting OD pair  $w \in W$ , and  $\mathbf{f} = (f_{wr}, r \in R^w, w \in W)^T$ .  $v_a$  is the traffic flow on link  $a \in A$ , and  $t_a(v_a)$  is the travel time (or link performance) function of link  $a \in A$ , which is assumed to be increasing, convex and continuously differentiable. The notations in this paper mostly follow that in Liu et al. [1, 7], which are summarized in Table 1.

A nonlinear-type distance-toll function is preferred according to Liu et al. [1, 7] and Meng et al. [8]; the general idea is to formulate the nonlinear distance-toll function  $\phi(\boldsymbol{\eta})$  as a piecewise linear function in terms of the travel distance  $\boldsymbol{\eta}$  in a cordon. It is clear to define the toll function  $\phi(\boldsymbol{\eta})$  on the range  $[\eta_0, \eta_K]$  with  $K$  equal intervals as shown in Figure 1. Note that  $\eta_0$  and  $\eta_K$  are the minimal and maximal route length in the charging cordon, respectively. Obviously, the piecewise linear toll function composes  $K$  straight line sections, and each line section is uniquely defined by the two ends points of each interval  $\mathbf{y} = (y_0, y_1, y_2, \dots, y_k, \dots, y_K)^T$  whose corresponding value of distance is  $\boldsymbol{\eta} = (\eta_0, \eta_1, \eta_2, \dots, \eta_k, \dots, \eta_K)^T$ . With this piecewise linear toll function discussed above, the continuous curve of the nonlinear distance-toll function can be characterized as a number of straight lines, which can be determined by  $K + 1$

TABLE 1: List of notations.

Notation	Explanation
$D$	The total planning period for one toll pattern.
$d$	The number of days after the toll implementation, $d = 1, 2, \dots, D$ .
$W$	The set of OD pairs.
$R^w$	The set of routes connecting an OD pair $w \in W$ .
$c_{wr}$	The travel cost on route $r \in R^w$ connecting OD pair $w \in W$ .
$\mathbf{f}$	The route flows over the entire network, $\mathbf{f} = (f_{wr}, r \in R^w, w \in W)^T$ .
$f_{wr}$	The traffic flow on route $r \in R^w$ connecting OD pair $w \in W$ .
$\mathbf{q}$	The travel demands, $\mathbf{q} = (q^w, w \in W)^T$ .
$q^w$	The travel demand connecting OD pair $w \in W$ .
$\mathbf{t}(\mathbf{v})$	The link travel time functions, $\mathbf{t}(\mathbf{v}) = (t_a(v_a), a \in A)^T$ .
$\mathbf{v}$	The link flows, $\mathbf{v} = (v_a, a \in A)^T$ .
$t_a(v_a)$	The travel time function of link $a \in A$ .
$v_a$	The traffic flow on link $a \in A$ .
$\delta_{ar}^w$	$\delta_{ar}^w = 1$ if route $r \in R^w$ contains link $a$ , and $\delta_{ar}^w = 0$ otherwise.
$\mathbf{y}$	The vertex values, $\mathbf{y} = (y_0, y_1, y_2, \dots, y_k, \dots, y_K)^T$ of the piecewise linear toll function.
$y_{\min}, y_{\max}$	The lower and upper bound of the distance-based toll.
$\phi(\boldsymbol{\eta})$	The toll charge function.
$K$	The total number of the intervals in the toll function $\phi(\boldsymbol{\eta})$ .
$\boldsymbol{\eta}$	Column vector of the travel distance in a cordon, $\boldsymbol{\eta} = (\eta_{wr}, r \in R^w, w \in W)^T$ .
$\boldsymbol{\tau}$	Column vector of the distance-based toll $\boldsymbol{\tau} = (\tau_{wr}, r \in R^w, w \in W)^T$ .
$N_c, N_e, N_o$	Number of the colony size, the employed bees, and the onlookers.
$\gamma_0, \gamma_1, \gamma_2$	Parameters used in the day-to-day dynamics model.

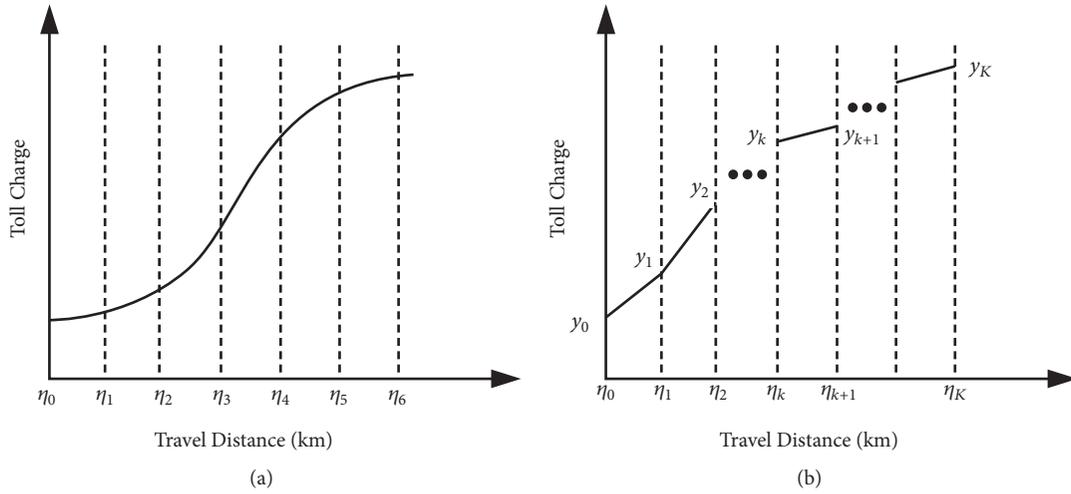


FIGURE 1: Piecewise linear toll.

points. For example, assume that  $y_{\min} = 2.0$ ,  $y_{\max} = 5.0$ , and  $\Delta_y = 0.5$ , and then  $K = (5.0 - 2.0)/0.5 = 6$ , which means the piecewise linear toll function can be determined with 6 intervals, and these 6 intervals can be uniquely defined by 7 vertices. By the way, the proposed methodology here can be easily generalized for the case of unequal intervals between the range  $[\eta_0, \eta_K]$ .

It is practical to define a nondecreasing distance-toll function in real life; thus we should have

$$y_{\min} = y_0 \leq y_1 \leq y_2 \leq \dots \leq y_k \leq \dots \leq y_K \leq y_{\max} \quad (1)$$

In a real application, we only need to know a particular route length, then we can calculate the corresponding toll value in terms of the piecewise linear toll function. Let  $\eta_{wr}$  denote the

length portion of route  $r \in R^w$  in the cordon. Suppose  $\eta_{wr}$  locates in the  $k$ th distance interval of  $\phi(\boldsymbol{\eta})$  shown in Figure 1, then the distance toll of route  $r \in R^w$  is

$$\tau_{wr} = \phi(\eta_{wr}) = y_{k-1} + \frac{\eta_{wr} - \eta_{k-1}}{\eta_k - \eta_{k-1}} (y_k - y_{k-1}) \quad (2)$$

The total/generalized travel cost on route  $r \in R^w$  connecting OD pair  $w \in W$

$$c_{wr} = \sum_a t_a \delta_{ar}^w + \frac{\tau_{wr}}{\kappa} \quad (3)$$

where  $\kappa$  is the travelers' value-of-time.

From (2) we can see that each toll pattern  $\tau$  is uniquely determined by the vertex value  $\mathbf{y} = (y_0, y_1, y_2, \dots, y_k, \dots, y_K)^T$ . Let  $\Omega_y$  be the set of all the feasible  $\mathbf{y}$ . Then, the optimal congestion pricing becomes a problem to obtain the optimal  $\mathbf{y}^* \in \Omega_y$ . Before introducing the model for the optimal  $\mathbf{y}^*$ , the day-to-day dynamics model is first discussed in the next section.

### 3. Day-To-Day Dynamics Model

A reasonable day-to-day dynamics model should well reflect the realistic route adjustment process and learning behavior of commuters [15, 16]. As for the travel behavior of an individual commuter, his/her route choice on current day is dependent on his/her route choice decision as well as what others did in the previous day. Essentially, this reconsideration process of the route choice from day to day is an experienced weighted learning model [22]. Compared to the reinforcement learning model in market entry games [23], the modification of this model is that commuter's route choice in that past may influence his/her own current route choice, and this route choice may also influence other commuters' route choice behaviors [24]. As for an ordinary commuter, he/she would concern about the utilization for every route (no matter whether it was his/her currently chosen route) from day to day to make his/her route choice in the next day more rational.

Erev et al. [25] showed that subjects in the market entry games were mainly influenced by what happened in the most recent times, and this effect would be more obvious in the day-to-day dynamics process [24]. For instance, commuters' route choice behaviors are highly affected by the most recently unexpected incidents, such as unexpected network disruptions and adverse weather conditions. Thus, in the proposed day-to-day dynamics model of this paper, we assume that a commuter's route choice on the current day is only dependent on his/her and other commuters' route choice decisions in the previous day, and this is a finite learning process in essence. The most obvious characteristic is that the route flow on day  $d + 1$  decreases as its actual travel cost of that route on day  $d$  increases and vice versa. It is obvious that the route choice decision of the current day is the baseline of the route choice decision of the next day. For the sake of presentation but without loss of generality, all of the commuters are assumed at the initial state of day 1, and then traffic flows of the whole network evolve from day to day.

A very clear and concise finite learning process named logit-type Markov adaptive learning model is proposed to depict commuters' day-to-day route choice behaviors. Specifically, the baseline probability of choosing route  $r$  on day  $d + 1$ , which is the same route as day  $d$ , can be expressed through the following multinomial logit-type function:

$$\Pr^{baseline}(R_{d+1} = r) = \frac{\exp(A_{wr}^{baseline}(\mathbf{y}, d))}{\sum_k \exp(A_{wk}^{baseline}(\mathbf{y}, d))} \quad (4)$$

where  $R_{d+1}$  is the route choice decision on day  $d + 1$ ,  $A_{wr}^{baseline}(\mathbf{y}, d) = -\gamma_1 C_{wr}(\mathbf{y}, d)$  is the *attraction of* or the *propensity towards* choosing route  $r \in R^w$  with the toll pattern  $\mathbf{y}$  on day  $d$ , and  $\gamma_1$  is a positive response sensitivity parameter.

Then, the flow of route  $r$  can be calculated by

$$f_{wr}^{baseline}(\mathbf{y}, d + 1) = q^w \cdot \Pr^{baseline}(R_{d+1} = r) \quad (5)$$

From this baseline model, we can see clearly that the routes with high travel costs on day  $d$  hold weaker attraction to commuters compared to those routes with low travel costs on the same day, and this is coincided with the nature of day-to-day dynamics process.

According to the baseline model, the following model is proposed to estimate the actual attraction of route  $r$  on day  $d$ :

$$A_{wr}(\mathbf{y}, d) = \begin{cases} \gamma_1 (C_{wi}(\mathbf{y}, d) - C_{wr}(\mathbf{y}, d)) & \text{if } r \neq i \text{ and } C_{wi}(\mathbf{y}, d) < C_{wr}(\mathbf{y}, d) \\ \gamma_0 & \text{if } r = i \\ \gamma_2 (C_{wi}(\mathbf{y}, d) - C_{wr}(\mathbf{y}, d)) & \text{if } r \neq i \text{ and } C_{wi}(\mathbf{y}, d) \geq C_{wr}(\mathbf{y}, d) \end{cases} \quad (6)$$

where  $\gamma_0$  and  $\gamma_2$  are positive parameters. The probability of choosing route  $r$  on day  $d + 1$  and the route flow can

be expressed through the following multinomial logit-type functions:

$$\Pr(R_{d+1} = r) = \frac{\exp(A_{wr}(\mathbf{y}, d))}{\sum_k \exp(A_{wk}(\mathbf{y}, d))} \quad (7)$$

$$f_{wr}(\mathbf{y}, d+1) = q^w \cdot \Pr(R_{d+1} = r) \quad (8)$$

It is worth noting that  $\gamma_0$  and  $\gamma_2$  are inertia and regret effect in this model, respectively. Compared with other day-to-day dynamics models, e.g., Guo et al. [26], He et al. [27], and Cantarella and Watling [28], the proposed model in this paper is more concise, and there is no explicit route adjustment parameter, which is actually reflected by the inertia and regret effect of  $\gamma_0$  and  $\gamma_2$ , respectively.

#### 4. Robust Optimization Model

As discussed in Introduction, the whole network environment will be changed after a period of days, and the days  $D = 90$  in this paper. Therefore, after 90 days, a new toll scheme will be performed to optimize the whole network with the already changed network environment, making a new process of day-to-day dynamics. At same time,  $d$  should be reset to 1 when the new toll scheme is implemented. Hence, the study period is from  $d = 1$  to  $d = D$ .

When a toll pattern  $\mathbf{y}$  is implemented, the route flows will be totally different from day to day, because the implemented toll scheme is an important component which will influence travelers' route choice behaviors. Let  $\mathbf{f}(\mathbf{y}, d)$  be the column vector of route flows on day  $d$  in terms of a toll pattern  $\mathbf{y}$ , and  $\mathbf{f}(\mathbf{y}, d)$  is clearly determined by (4)-(8). The authorities' objective is to minimize the total travel costs for each day rather than merely for the final equilibrium state. Since the commuters' route choice behavior follows the logit-type Markov adaptive learning model introduced in Section 3, the system's optimal performance is reflected by the one with minimal total travel cost. On day  $d$ , the total travel cost can be calculated by

$$TTC(\mathbf{y}, d) = \mathbf{f}(\mathbf{y}, d)^T \cdot \mathbf{c}(\mathbf{y}, \mathbf{f}, d) \quad (9)$$

It may be impossible for just one particular toll pattern to optimize the system's performance for all the days/scenarios (from day 1 to day  $D$ ). If a particular toll scheme  $\mathbf{y}$  is implemented on  $d = 1$ , then there would be a new day-to-day route flow evolution process. In the context of day-to-day dynamics, network flows are varying each day, so any toll pattern that can give rise to minimum total travel cost in a particular day may lead to bad traffic conditions in some other days. From the viewpoints of policy-maker, the deterioration of some worst days is more harmful than the loss of efficiency on the good cases, both temporally and spatially. Thus, it is a better strategy in practice that: compromising the efficiency on the optimal end of some good days so that the severe traffic congestions in some bad days could be ameliorated. Following this logic, this paper then takes the mini-max total travel cost as the objective; thus

$$\mathbf{y}(d) \in \arg \min_{\mathbf{y} \in \Omega_y} \max TTC \quad (10)$$

subject to the day-to-day route flows introduced in Section 3.

It is clear that model (10) is a robust optimization model, which can also be deemed as a bilevel model, where the upper level is a mini-max total travel cost model, and the lower level reflects the day-to-day dynamic flows, which is discussed in Section 3. The optimal solution of model (10) is a robust toll scheme that considers the system's performance on each day of the study horizon.

#### 5. Solution Algorithms

All of the existing solution methods (such as sensitivity analysis method, system optimal relaxation method, and gap function method) are not suitable to solve the proposed bilevel robust model due to the complexity of the flow evolution process  $\mathbf{f}(\mathbf{y}, d)$ . More specifically, the day-to-day dynamics model of  $\mathbf{f}(\mathbf{y}, d)$  has no closed form when  $d > 2$ , in spite of an initial flow pattern  $\mathbf{f}(\mathbf{y}, 1)$  is given. Therefore, a heuristic algorithm named artificial bee colony algorithm is adopted in this paper.

The ABC algorithm was originally proposed by Karaboga [29] for solving unimodal and multimodal numerical optimization problems. Recently, this algorithm is used to solve transportation problems [3, 22–25, 27]. We also use the ABC algorithm to figure out the mini-max TTC problem here due to its advantages of good local search mechanism which can enhance the solution quality compared to other evolutionary algorithms such as genetic algorithm [30]. In this paper, we proposed a modified ABC algorithm to solve the robust programming model for the optimal toll design. According to an initial food source, we can obtain the TTC for each day and the maximum TTC among the whole planning period, then the mini-max TTC can be calculated through the modified ABC algorithm. Finally, the robust optimal toll pattern can be output. The procedures of this algorithm are summarized in Figure 2. For a detailed description of the ABC algorithm, the readers are referred to some other references, such as [3, 25, 27].

#### 6. Conclusions

This paper studies the nonlinear distance-based toll with day-to-day dynamic traffic flow evolution. After an implementation or adjustment of a new congestion toll scheme, the network environment will change and traffic flows will be nonequilibrium in the following days, which makes it not suitable to take the equilibrium-based indexes as the objective of the congestion toll. A mini-max total travel cost model, which can overcome the expensive computational burden problem in previous work and make the congestion toll scheme more practical, is then proposed to solve the congestion toll problem considering nonequilibrium flow evolution processes. The essence of the mini-max total travel cost model is to optimize the worst condition among the whole planning period and ameliorate severe traffic congestions in some bad days, and this takes into consideration the network performance on each day of the study horizon rather than the final equilibrium state.

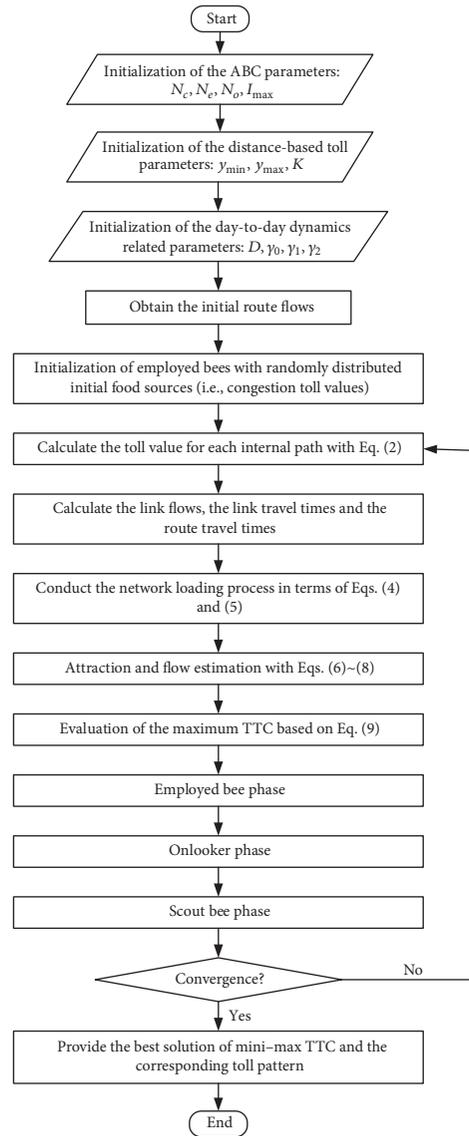


FIGURE 2: Flowchart of the solution algorithm.

## Data Availability

No external data were used to support this study.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study is supported by the National Natural Science Foundation of China (nos. 71501038 and 71601045), the Key Project of National Natural Science Foundation of China (no. 51638004), and the Scientific Research Foundation of Graduate School of Southeast University (no. YBPY1885).

## References

- [1] Z. Liu, S. Wang, and Q. Meng, "Optimal joint distance and time toll for cordon-based congestion pricing," *Transportation Research Part B: Methodological*, vol. 69, pp. 81–97, 2014.
- [2] A. May and D. Milne, "Effects of alternative road pricing systems on network performance," *Transportation Research Part A: Policy and Practice*, vol. 34, no. 6, pp. 407–436, 2000.
- [3] G. Mitchell, A. Namdeo, and D. Milne, "The air quality impact of cordon and distance based road user charging: An empirical study of Leeds, UK," *Atmospheric Environment*, vol. 39, no. 33, pp. 6231–6242, 2005.
- [4] A. Namdeo and G. Mitchell, "An empirical study of estimating vehicle emissions under cordon and distance based road user charging in Leeds, UK," *Environmental Modeling & Assessment*, vol. 136, no. 1-3, pp. 45–51, 2007.

- [5] S. Lawphongpanich and Y. Yin, "Nonlinear pricing on transportation networks," *Transportation Research Part C: Emerging Technologies*, vol. 20, no. 1, pp. 218–235, 2012.
- [6] Q. Cheng, Z. Liu, and A. Seloune, "Nonlinear Distance-Based Dynamic Pricing Considering Congestion-Level Correction," *International Journal of Transportation*, vol. 6, no. 2, pp. 1–8, 2018.
- [7] Z. Liu, S. Wang, B. Zhou, and Q. Cheng, "Robust optimization of distance-based tolls in a network considering stochastic day to day dynamics," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 58–72, 2017.
- [8] Q. Meng, Z. Liu, and S. Wang, "Optimal distance tolls under congestion pricing and continuously distributed value of time," *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 5, pp. 937–957, 2012.
- [9] X. Sun, D. Huang, and Q. Cheng, "Optimal Nonlinear Distance Toll for Cordon-Based Congestion Pricing Considering Equity Issue," *Journal of Harbin Institute of Technology (New Series)*, vol. 23, no. 6, pp. 73–79, 2016.
- [10] Land Transport Authority of Singapore, "Land Transport Master Plan 2013," Singapore Land Transport Authority, 2013.
- [11] P. Olszewski and L. Xie, "Modelling the effects of road pricing on traffic in Singapore," *Transportation Research Part A: Policy and Practice*, vol. 39, no. 7-9, pp. 755–772, 2005.
- [12] Z. Liu, Q. Meng, and S. Wang, "Speed-based toll design for cordon-based congestion pricing scheme," *Transportation Research Part C: Emerging Technologies*, vol. 31, pp. 83–98, 2013.
- [13] X. Guo and H. X. Liu, "Day-to-Day Dynamic Model in Discrete-Continuum Transportation Networks," *Transportation Research Record*, vol. 2263, no. 1, pp. 66–72, 2018.
- [14] X. He and H. X. Liu, "Modeling the day-to-day traffic evolution process after an unexpected network disruption," *Transportation Research Part B: Methodological*, vol. 46, no. 1, pp. 50–71, 2012.
- [15] B. Wie and R. L. Tobin, "Dynamic congestion pricing models for general traffic networks," *Transportation Research Part B: Methodological*, vol. 32, no. 5, pp. 313–327, 1998.
- [16] T. L. Friesz, D. Bernstein, and N. Kydes, "Dynamic Congestion Pricing in Disequilibrium," *Networks and Spatial Economics*, vol. 4, no. 2, pp. 181–202, 2004.
- [17] R. Guo, H. Yang, H. Huang, and Z. Tan, "Day-to-Day Flow Dynamics and Congestion Control," *Transportation Science*, vol. 50, no. 3, pp. 982–997, 2016.
- [18] H. Ye, H. Yang, and Z. Tan, "Learning marginal-cost pricing via a trial-and-error procedure with day-to-day flow dynamics," *Transportation Research Part B: Methodological*, vol. 81, pp. 794–807, 2015.
- [19] Z. Tan, H. Yang, and R. Guo, "Dynamic congestion pricing with day-to-day flow evolution and user heterogeneity," *Transportation Research Part C: Emerging Technologies*, vol. 61, pp. 87–105, 2015.
- [20] T. Rambha and S. D. Boyles, "Dynamic pricing in discrete time stochastic day-to-day route choice models," *Transportation Research Part B: Methodological*, vol. 92, pp. 104–118, 2016.
- [21] Q. Cheng, Z. Liu, F. Liu, and R. Jia, "Urban dynamic congestion pricing: an overview and emerging research needs," *International Journal of Urban Sciences*, vol. 21, no. sup1, pp. 3–18, 2017.
- [22] T. H. Ho, C. F. Camerer, and J.-K. Chong, "Self-tuning experience weighted attraction learning in games," *Journal of Economic Theory*, vol. 133, no. 1, pp. 177–198, 2007.
- [23] I. Erev and A. E. Roth, "Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria," *American Economic Review*, vol. 88, no. 4, pp. 848–881, 1998.
- [24] V. Mak, E. J. Gisches, and A. Rapoport, "Route vs. Segment: An Experiment on Real-Time Travel Information in Congestible Networks," *Production Engineering Research and Development*, vol. 24, no. 6, pp. 947–960, 2015.
- [25] I. Erev, E. Ert, and A. E. Roth, "A choice prediction competition for market entry games: an introduction," *Games*, vol. 1, no. 2, pp. 117–136, 2010.
- [26] R.-Y. Guo, H. Yang, H.-J. Huang, and Z. Tan, "Link-based day-to-day network traffic dynamics and equilibria," *Transportation Research Part B: Methodological*, vol. 71, pp. 248–260, 2015.
- [27] X. He, X. Guo, and H. X. Liu, "A link-based day-to-day traffic assignment model," *Transportation Research Part B: Methodological*, vol. 44, no. 4, pp. 597–608, 2010.
- [28] G. E. Cantarella and D. P. Watling, "A general stochastic process for day-to-day dynamic traffic assignment: Formulation, asymptotic behaviour, and stability analysis," *Transportation Research Part B: Methodological*, vol. 92, pp. 3–21, 2016.
- [29] D. Karaboga, *An Idea Based on Honey Bee Swarm for Numerical Optimization*, 2005.
- [30] J. Chen, Z. Liu, S. Zhu, and W. Wang, "Design of limited-stop bus service with capacity constraint and stochastic travel time," *Transportation Research Part E: Logistics and Transportation Review*, vol. 83, pp. 1–15, 2015.

## Research Article

# Applying Big Data Analytics to Monitor Tourist Flow for the Scenic Area Operation Management

Siyang Qin <sup>1</sup>, Jie Man <sup>2</sup>, Xuzhao Wang,<sup>2</sup> Can Li,<sup>2</sup> Honghui Dong,<sup>2</sup> and Xinquan Ge<sup>1,3</sup>

<sup>1</sup>School of Economics and Management, Beijing Jiaotong University, 3 ShangyuanCun, Haidian District, Beijing 100044, China

<sup>2</sup>School of Traffic and Transportation, Beijing Jiaotong University, 3 ShangyuanCun, Haidian District, Beijing 100044, China

<sup>3</sup>School of Economics and Management, Beijing Information Science and Technology University, 12 Qinghe Xiaoying East Road, Haidian District, Beijing 100192, China

Correspondence should be addressed to Siyang Qin; syqin@bjtu.edu.cn and Jie Man; 16120770@bjtu.edu.cn

Received 7 September 2018; Revised 11 November 2018; Accepted 18 November 2018; Published 1 January 2019

Academic Editor: Lu Zhen

Copyright © 2019 Siyang Qin 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.

Considering the rapid development of the tourist leisure industry and the surge of tourist quantity, insufficient information regarding tourists has placed tremendous pressure on traffic in scenic areas. In this paper, the author uses the Big Data technology and Call Detail Record (CDR) data with the mobile phone real-time location information to monitor the tourist flow and analyse the travel behaviour of tourists in scenic areas. By collecting CDR data and implementing a modelling analysis of the data to simultaneously reflect the distribution of tourist hot spots in Beijing, tourist locations, tourist origins, tourist movements, resident information, and other data, the results provide big data support for alleviating traffic pressure at tourist attractions and tourist routes in the city and rationally allocating traffic resources. The analysis shows that the big data analysis method based on the CDR data of mobile phones can provide real-time information about tourist behaviours in a timely and effective manner. This information can be applied for the operation management of scenic areas and can provide real-time big data support for “smart tourism”.

## 1. Introduction

With the rapid development of China's economy, the large and medium-sized cities have entered the Leisure Era. The quality of leisure has become an important evaluation criterion of the living quality of urban residents [1] and an essential part of public life. Witnessing the continuous growth of tourist flow in the scenic spots, it is vital to understand the accurate and real-time travel behaviour information of different types of tourists [2–10]. In recent years, the statistical monitoring methods for the tourist flow mainly focus on video surveillance, entrance gates, and other means [11]. The development of network and Internet technology also has enabled technical means of monitoring based on the WiFi provided by scenic areas and mobile phone application terminals. These flow monitoring methods require cameras, gateways, WiFi base stations, and other equipment to support them. It is difficult to install and implement in all scenic areas. Most mobile applications (such as WeChat) are oriented towards young and middle-aged users and cannot monitor tourists of all types.

The conventional estimation of macroscopic travel demands mostly relies on the empirical judgment on historic data by transportation practitioners. It is costly and the resolution is limited. Therefore, it is practical to develop new and better tools to automate the identification of tourist flows from large-scale mobility data sets. Some research is conducted with smart card fare data [12, 13] and mobile phone data [14].

The quantity of mobile phone users in China is constantly increasing. According to data released by China's Ministry of Industry and Information Technology (MIIT), as of January 2017, the number of mobile phone users in China reached 1.32 billion, and the penetration rate of mobile phones reached 96.2%. The number of mobile phone users in Beijing reached 3.869 million, and the penetration rate of mobile phones reached as high as 178.3%. Nearly everyone who travels carries their cell phone. Based on the continuous development of positioning technology in mobile communication, the tourist flow analysis with Call Detail Record (CDR) data can provide real-time valid data for scenic flow control, tourist diversion, traffic dispersion, safety management, and so on. And these

can support to improve the operation management of the scenic area.

Through the comprehensive analysis of scenic spots and tourists' travel behaviour, researchers have explored tourist flow and tourist behaviours in the corresponding scenic spots. Ferrari [15], through the processing and analysis of the Call Detail Record (CDR) data of mobile phones, analysed the temporal and spatial regulations of personal behaviour and the nature and law of events occurring in the city to provide a theoretical basis for the management of events [15]. Ahas and Silm studied the temporal and spatial distribution of Estonian tourists, offering big data support for tourism planning [16–19]. Based on the CDR data of mobile phones, Dong [20, 21] analysed the spatial and temporal situation of population movement within the Sixth Ring Road area in Beijing at both the regional and road network levels. Etison [22], utilizing the CDR data, established a resident traffic flow monitoring and management system for real-time data monitoring and statistics for mobile phone users in specific areas.

The existing research results have no specific target population for the analysis of mobile phone users and do not cover the research of users' individual behaviour. Therefore, this paper intends to analyse the main tourist attractions in Beijing by employing handset CDR data, focusing on the Forbidden City as the main research object. The author aims to make full use of the advantages of big data to effectively analyse the residence time, the spatial and temporal distribution of tourists, and the behaviour of tourists. The experiment results show that this big data analysis technology can become technical support for the management of urban tourist areas and effectively improve the accuracy of the current tourist flow monitoring.

## 2. Data Description

**2.1. Call Detail Record (CDR) Data.** Positioning with the mobile communication network is a technology that acquires information through mobile Call Detail Records (CDR) in the communication network.

**2.1.1. Location Mode of the Cellular Network.** Each zone covered by the mobile communication network is generally assumed as a regular hexagon, where each mobile user can act as a mobile station (MS), as shown in Figure 1. Regardless of the user's state, whether mobile or stationary, as long as the user calls or accesses the Internet, the user of the mobile phone will exchange data with the nearest mobile phone base station (BS). Therefore, the data containing the base station ID is recorded. Depending on the location of the base station, the current location of the user can be estimated. The specific positioning is shown in Figure 1.

When the user moves from Cell-1 to Cell-7, a total of three location areas (LACs) and seven cells are passed through in sequential order from Cell-1, Cell-2, Cell-3, Cell-4, Cell-5, and Cell-6 to Cell-7. If the users switch on and off, interact with the network, or receive the data regarding user steps, the user's location information is updated since LAC<sub>1</sub>, LAC<sub>2</sub>, and LAC<sub>3</sub> belong to different receiving areas. Taking Cell-2 and Cell-3 of LAC<sub>1</sub> and Cell-4 of LAC<sub>2</sub> as examples, if

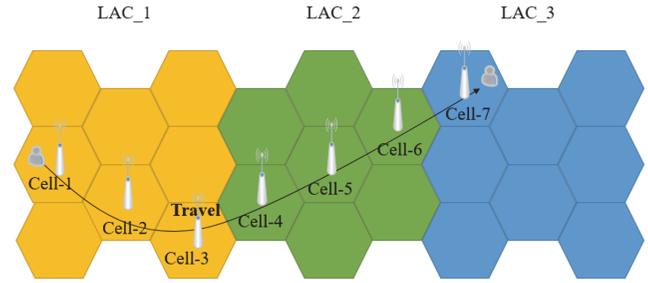


FIGURE 1: Location mode of mobile communication network.

there is no interaction with the network in the process of their moving, such as calling, texting, or surfing the Internet, the user's location will not be updated, since Cell-2 and Cell-3 are from the same location area (LAC<sub>1</sub>). In other words, when two cells belong to the same location area, the two cells have the same location area code. Although the cell area is changed, the location area remains intact, and thus the change in location will not be recorded in this situation. Since Cell-3 and Cell-4 belong to different location areas, when the user moves from Cell-3 to Cell-4, which is equivalent to the user moving from LAC<sub>1</sub> to LAC<sub>2</sub>, the centre will record the user's location changes regardless of whether the user has data interaction with the network during the movement.

This cell phone positioning method is called the Cell-ID positioning method, which is utilized as a web-based cell phone positioning technology without installing additional equipment, acquiring maintenance costs and upgrading the existing mobile network. Although this method cannot obtain the location information accurately, the macroscopic tourist flow analysis used in this paper has already met its accuracy requirement.

**2.1.2. Base Station Data of the Cellular Network.** In general, each base station (BS) has its own fixed location area (LAC). When a cell phone is registered in the communication network, the network will page the LAC location of the cell phone and obtain the corresponding Cell-ID. In this paper, the LAC is intercepted from the location area identification code (LAI) to identify the location area in the GSM network. The MSC area (hereafter called the MSC) is composed of all the location areas under its control. In addition to administering the subordinate location area, the MSC is also responsible for the reception of facsimile data. The geographical location of the base station (BS) mainly consists of information such as its administrative region, latitude, and longitude. The network attributes of the base station (BS) mainly cover information such as the BSC number used to control the communication, the running status, the antenna height, and the base station type.

The Cell-ID and LAC of the selected base station will be taken as the basic characteristic attributes of the base station. The recorded information includes the user's mobile phone ID, the user's phone status, the location area (LAC) of the user, the Cell-ID, the type of event triggered, and the recording position. The record table structure and data samples are shown in Table 1.

TABLE 1: Samples of the mobile phone base station.

ID	Status	BTS	Cell-ID	Community	SITENO	LAC
1	Working	139	45723	Tongzhou DistrictG1	45721	4280
2	Working	138	45722	Tongzhou DistrictG2	45721	4280
...	.....	...	.....	.....	.....	...
ID	Longitude	Latitude	Antenna Height	Antenna Azimuth	Coverage Condition	Station Location
1	116.887742	39.809074	33	120	Residential Area	Exercise Plaza in Yueshang Village, Tongzhou District
2	116.887742	39.809074	33	240	Residential Area	Exercise Plaza in Yueshang Village, Tongzhou District
...	.....	...	.....	.....	.....	...

TABLE 2: Example of CDR.

IMSI	Cell-ID	LAC	CTIME	TIMECHAR
96cbf1288b3ae673f1711c3falea9ac6	4259	11033	1422720000	20150201010000
54895a8b525d586e5de85417e0763eb3	4341	19507	1422720000	20150201010000
.....	.....	.....	.....	.....

Mobile communication networks cover a large amount of the population activity area, especially in urban areas. As shown in Figure 2, more than 70% of the base stations in Beijing are located in the urban area within the Sixth Ring Road area, with a total of 52,759 base stations. Meanwhile, according to statistics from the Ministry of Industry and Information Technology, the number of phone users in Beijing reached 20.62 million in 2015. With the mobile phone base station functioning as a fixed traffic detector and the user mobile phone as a mobile detector, the urban travel holographic information acquisition system is practicable. Therefore, the CDR data can be used in the analysis of extremely complicated travel behaviour.

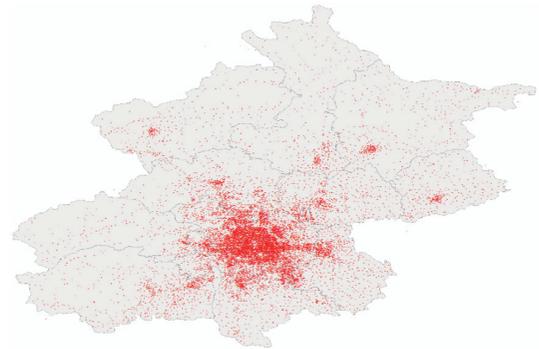


FIGURE 2: Mobile phone base station distribution in Beijing.

2.1.3. *User Data of the Cellular Network.* As with the base station attributes and types, when a user moves, updating the user location data acquired by the GSM network also must be defined by the user attributes. Each location update will generate a user location datum correspondingly. The GSM network encodes the user information not only to monitor the change of its location but also to confirm that users are properly connected to each other when using the network. Hence, it is necessary to correctly address this issue through encoding. Each user has a cell phone number during the use procedure. To protect the privacy of the user, the system will desensitize the user code hiding the user information, in the process of network storage. In this paper, based on the existing data, the author selected IMSI, Cell-ID, LAC, CTIME, and TIMECHAR as the location-updating attributes of the users under the condition of ensuring data quality. The Cell-ID and LAC meaning is consistent with the base station attribute property. The final selected mobile network user locations are updated in the format shown in Table 2.

The user data of the mobile communication adopted in this paper are recorded for one year. The data are collected every two seconds, with a data size of 52,759 pieces. Since there are 86,400 seconds in a day, this collection will amount

to 450 million pieces per day with an average daily data size of 40G. To address large amounts of data quickly, preprocessing the data, cleaning the noise, and converting the format are necessary steps.

2.2. *Processing Platform.* The CDR data are a kind of typically big data. Taking Beijing as an example, the CDR data for one day exceed 400 million pieces, which requires a large amount of calculation for processing. However, the processing capacity and I/O performance of a single machine cannot support such a large data calculation. At the same time, traditional relational databases, such as oracle, can build clusters, but when the amount of data reaches a certain limit, the query processing speed will become very slow, and the performance of the machine is very high. Thus, the Spark big data platform is considered to handle the CDR data in this paper.

The concept of Resilient Distributed Dataset (RDD) is adopted in the Spark framework. Considering that Map Reduce cannot complete effective data sharing at all stages of the parallel computing, RDD in the Spark framework makes up for this defect. Using this efficient data sharing and

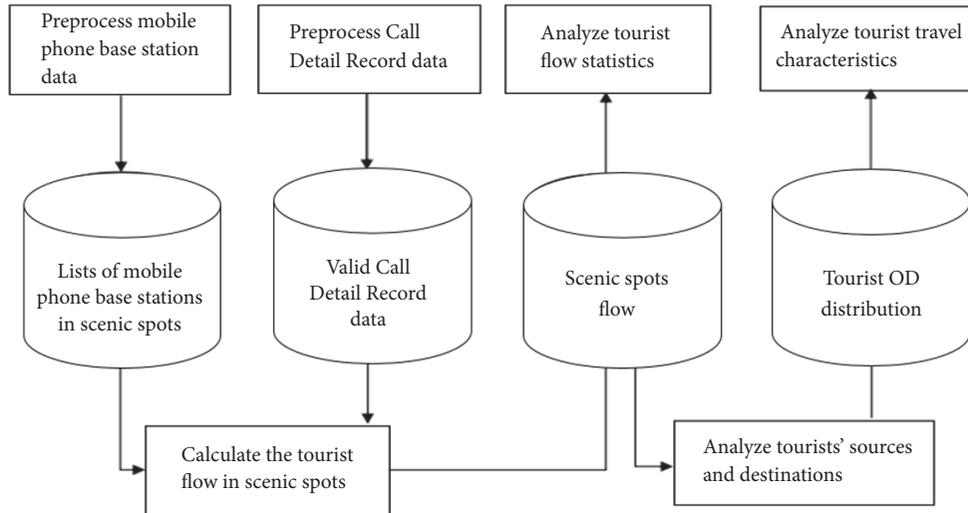


FIGURE 3: The overall structure.

Map Reduce-like operating interfaces, various proprietary types of calculations can be effectively expressed in the Spark framework, and similar performance can be achieved. According to the popular classification of the application area, big data processing can be divided into complex batch data processing, interactive data query based on historical data, and streaming data processing based on real-time data streaming. Because of the abundant expression capability of RDD, the unified large data processing platform capable of simultaneously dealing with the above three situations is derived on the basis of the Spark core. The goal of the Spark ecosystem is to integrate batch processing, interactive processing and streaming processing into the same software stack. In this paper, the Spark SQL interface is used, which provides a distributed SQL engine with a query speed 10 ~ 100 times higher than hive.

### 3. Tourist Flow Statistics and Characteristic Analysis Method

As shown in Figure 3, the overall flow chart depicting the scenic spot flow calculation and tourist travel characteristics analysis, the preprocessing of mobile phone base station data, and the CDR data are completed first to obtain the lists of all the base stations of each surveyed scenic spot and active phone users. Next, through matching the CDR data of the base stations in the scenic spots, the flow of each scenic spot will be obtained, and the tourist flow characteristics will be analysed. Based on the origin and the destination of tourists in the scenic spots, the tourist OD spatial distribution and the travel characteristics of tourists in the relative scenic spots can be obtained.

#### 3.1. Data Preprocessing

**3.1.1. Preprocessing of Mobile Phone Base Station Data.** Because the base station data are continuously improved under the operation of a communication company, there are

many basic data without value. The Spark SQL has processed 52,759 base station raw data into 43,022 valid pieces of data and extracted Cell-ID, LAC, longitude and latitude, base station type, coverage area type, and base station location as new attribute values. The specific details are shown in Table 3.

This paper analyses the main scenic spots in Beijing, including the Forbidden City, the Summer Palace, and the Olympic Forest Park. Working with data from more than 40,000 base stations in Beijing, Python language is adopted to write scripts on the ArcGIS platform, handling 20 scenic spots in a batch. A buffer zone of 100 metres is designed, and then the base stations are screened out in the scenic area by matching with the location of the scenic area, as shown in Figure 4. The Cell-ID and LAC attributes of the base station are selected to store in the lists of base stations in the scenic spots.

**3.1.2. Processing of the CDR Data.** The IMSI number, the only code in the CDR data for identifying the phone user, belongs to the STRING type, which is inconvenient for subsequent operations. Therefore, the IMSI number will be transformed into the LONG type through the HASH code, and its uniqueness and nonnegativity will be verified to ensure that one IMSI number can still determine one phone user. At the same time, due to the ping-pong switching phenomenon in the CDR data, “silent users” and ping-pong switching data have been filtered out by setting a threshold of frequency (PCF), leaving 67% of the data as the source for traffic information collection [9]. After completion of the above two steps, 40G of data per day can be reduced to approximately 16G, which significantly reduces the running time of the subsequent processing. The new CDR data table is shown in Table 4.

**3.2. Tourist Flow Algorithm Based on the CDR Data.** In this paper, the tourist flow statistics of the scenic spots are divided into five parts, namely, the total flow, influx, outflow, stagnant flow, and net increment and variation of the scenic spot

TABLE 3: Processed base station data.

Cell-ID	LAC	Longitude	Latitude	Type	Coverage	Address
435	4310	115.4275	39.9705	1	Residential area	Linchang Road side, Xiaolongmen Village, Qingshui Town, Mentougou District
6495	4148	115.4344	39.96334	1	Road traffic	Xiaolongmen Village, Qingshui Town, Mentougou District, Beijing
.....	.....	.....	.....	.....	.....	.....

TABLE 4: Example of processed CDR data.

HASH_ID	CTIME	LAC	CELL_ID
3304557668498423790	1423328400	4125	30519
5415176000769466540	1423328400	4352	13802
.....	.....	.....	.....



FIGURE 4: Mobile phone base stations in the scenic area. (a) Spatial distribution of typical scenic spots in Beijing. (b) Generating a list of mobile phone base stations in the scenic area.

during a certain period of time. The calculation method is as follows.

The specified study period is  $[a, b]$ , and the time interval from time  $a$  to time  $b$  is  $t$  hours; at the same time, a time period of  $[c, a]$  is set whose time interval is also  $t$  hour(s). Let the phone users of the first time period be collection  $B$ , in which the total number of users is recorded as  $C_b^t$ , and let the phone users of the second time period be collection  $C$ , in which the total number of users is recorded as  $C$ .

(1) The total flow of the scenic spot during a period of time is the total tourist number of the scenic spot in the time period  $[a, b]$ , with the collection expressed as  $B = C_b^t$ .

(2) The stagnant flow of the scenic spot during a certain period of time is the number of tourists in the scenic spot both within the time period  $[c, a]$  and within the next time period  $[a, b]$ , with the collection expressed as  $U_{ab} = |C \cap B| = |C_c^t \cap C_b^t|$ .

(3) The influx of the scenic spot during a certain period of time is the increased tourist number at the scenic spot during a certain period of time, that is, the number of tourists who were not in the scenic spot within the time period  $[c, a]$  but who appeared within the time period  $[a, b]$ , with the collection expressed as  $I_{ab} = B - U_{ab} = C_b^t - |C_c^t \cap C_b^t|$ .

(4) The outflow of the scenic spot during a certain period of time is the number of tourists who appeared within the time period  $[c, a]$  but did not appear within  $[a, b]$ , with the collection expressed as  $O_{ab} = C - U_{ab} = C_c^t - |C_c^t \cap C_b^t|$ .

(5) The net increment of the scenic spot during a certain period of time is the net increase in tourist number at the scenic spot during the time period  $[a, b]$ , calculated by the total influx minus the total outflow over this time period; this item can also be understood as the increased tourist number (which can be negative) of the time period  $[a, b]$  compared with the time period  $[c, a]$ . The collection is expressed as  $R_{ab} = I_{ab} - O_{ab} = B - C = C_b^t - C_c^t$ .

(6) The variation of the scenic spot during a certain period of time is the net increase in tourist number during the time period  $[0, b]$ , that is, the user number calculated by the net increase in tourist number within the time period  $[a, b]$  plus the total outflow within  $[0, a]$ . The collection is expressed as  $S_{0b} = R_{ab} + S_{0a}$ .

The calculation process is shown in Figure 5.

3.3. Tourist OD Analysis. The origin and destination analysis of tourists will reflect the important characteristics of tourist travel. Therefore, this paper considers using O (origin) and D (destination) of the trip to analyse the origin and destination of tourists and travel characteristics. The existing traffic zone division has been referred to the conduct tourist OD analysis.

3.3.1. Traffic Zone. This paper adopts the traffic zone division in the literature [20]. According to the processed mobile phone base station data and the CDR data (see Table 4 for the data format), the mobile phone base station is first defined in

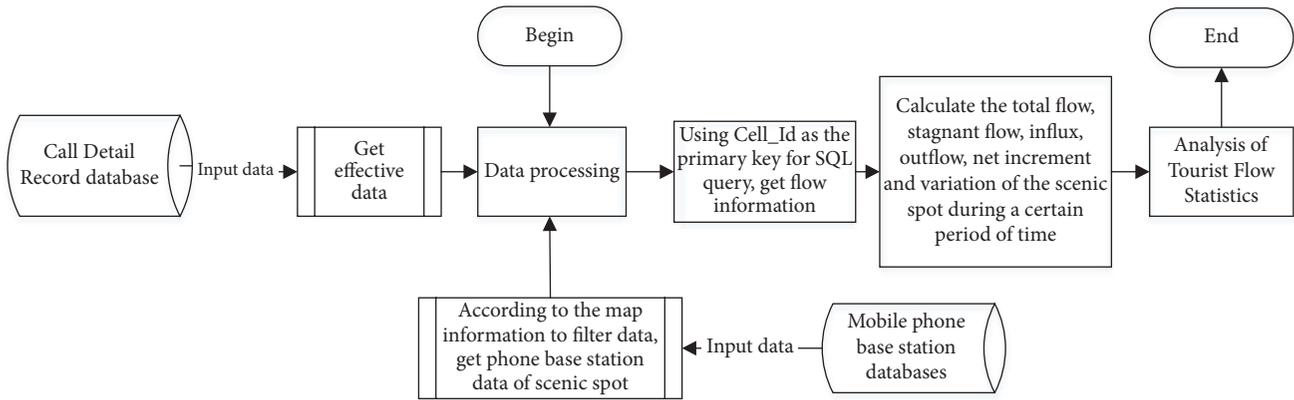


FIGURE 5: Tourist flow algorithm based on the CDR data.

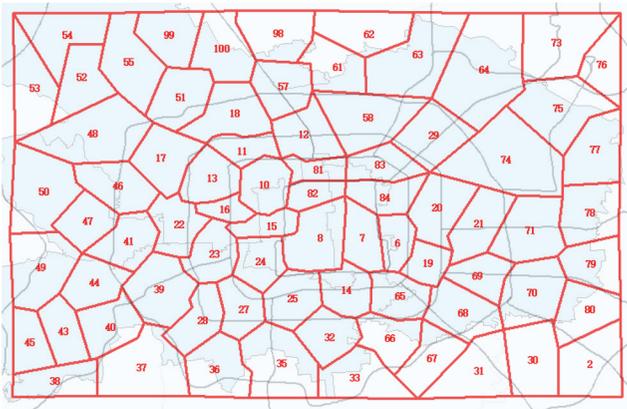


FIGURE 6: Traffic zone division of the area within the sixth ring in Beijing.

terms of traffic semantics as the residential area, work area, and road traffic. Then, it matches the geographic information system and divides these mobile phone base stations into “traffic zones” based on the mobility characteristics of the population directly related to traffic. Figure 6 depicts the division results for traffic zones in the urban area within the sixth ring in Beijing according to the above method.

According to the above method, each divided traffic zone contains a large number of base stations. First, the base station trajectory of each user needs to be extracted by a specific algorithm, through which the continuous location switching of the user at base stations can be obtained. The data in this step provide the database for acquiring the dynamic urban traffic OD, because the data are obtained by using the traffic zone as a unit and converting the user’s base station location information into the traffic zone location information to obtain the user’s zone switching trajectory.

**3.3.2. Tourist OD Analysis Algorithm Based on the CDR Data.** This paper primarily examines the OD flow of tourists in key scenic spots. Since February is a low season for tourism, the operating hours of most scenic spots are from 8:30 am to 16:30 pm. This paper takes the Summer Palace as an example to study the method for the tourist OD analysis.

The algorithm procedure is as follows:

(1) *Select Tourists Visiting the Scenic Spot.* The scenic spot opens at 8:30 am every day. Although it will take approximately 2 to 3 hours to visit the entire area, data should first be screened and should meet the following two conditions: (1) select users who remain in the scenic spot without going to other areas from 8:30 am to 11 am; (2) exclude those who remain in the scenic spot from 8:30 am to 4:30 pm (workers). The results provide user IDs for the tourists in the scenic spot.

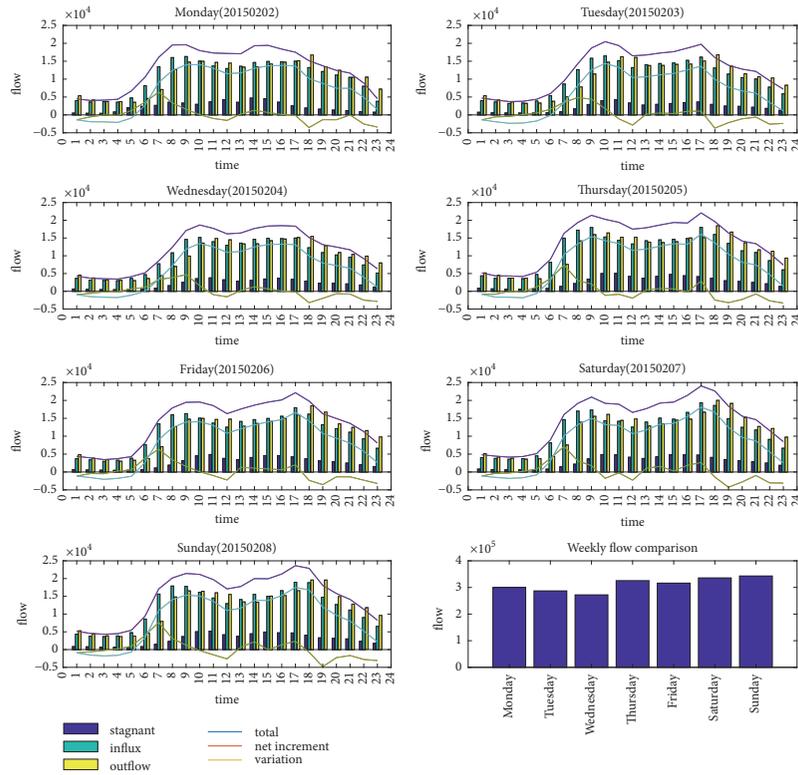
(2) *Obtain Tourists’ OD Information.* For tourists whose IDs have been selected, information on which traffic zone they originate from and where they finally arrive should be determined. This information is obtained by reverse querying, that is, using a user’s ID number to identify where she or he passed from 5 am to 8 am and from 11 am to 2 pm. Therefore, the user’s movement route from 5 am to 2 pm is determined.

(3) *Convert the Trajectory Data into an OD Matrix.* The trajectories obtained for each user may contain a significant amount of location data because the user is constantly moving. However, researchers only need the origin and destination of the users. To improve the fault tolerance, the result position data calculated by the median of the position data in this period multiplied by 0.7 plus the average of the position data in this period multiplied by 0.3. Traffic zones are determined based on the location range of the zones. The number of people in each traffic zone is then counted, and the OD matrix is finally produced.

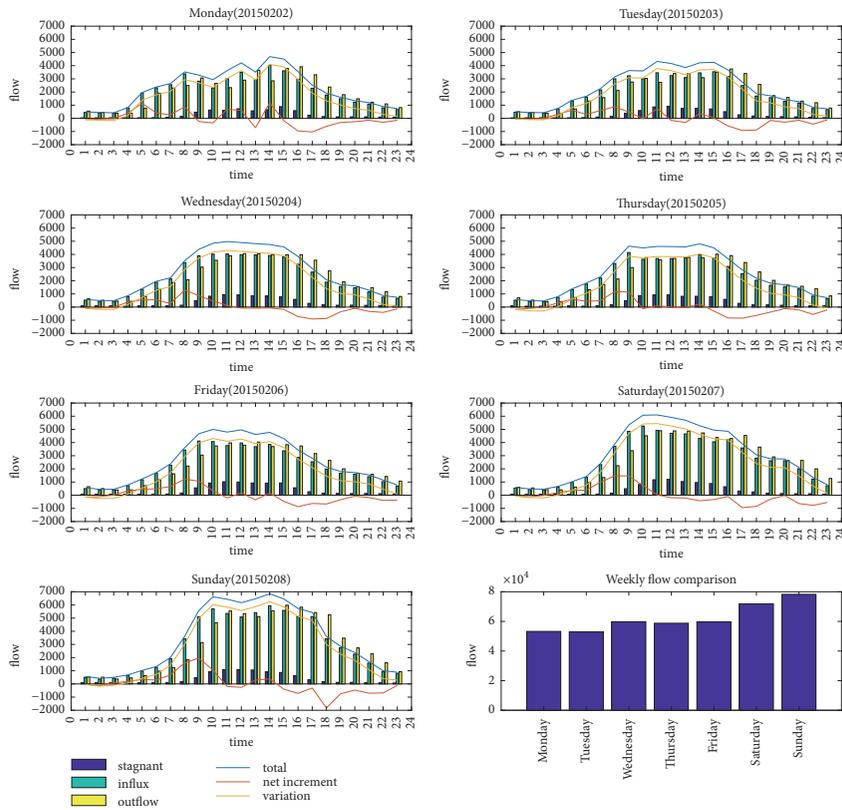
## 4. Experiment and Result Analysis

**4.1. Tourist Flow Statistics.** According to the above algorithm, we collect the statistics of tourist flow in 20 scenic spots and select 2 typical scenic spots including the Olympic Forest Park, Badaling Great Wall, obtaining the flow chart of the scenic spots as shown in Figure 7.

From the flow chart in Figure 7, we observe that the trend in the daily changes over time is similar in the same scenic spot. In general, most tourists begin their visit at 9:00~10:00 am and leave by 17:00~18:00 pm. At the same time, this trend



(a)



(b)

FIGURE 7: Weekly tourist flow chart of the scenic spots. (a) Olympic Forest Park. (b) Badaling Great Wall.

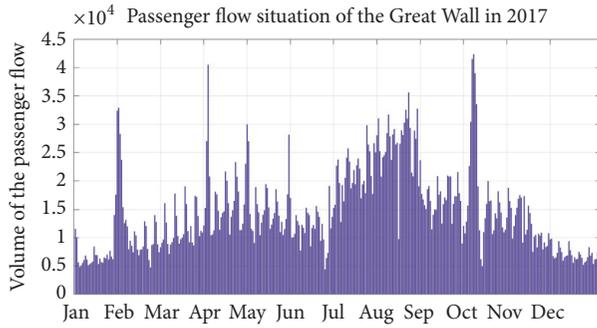


FIGURE 8: Tourist flow of the Badaling Great Wall for one year.

is not the same for traditional scenic areas and general parks; traditional scenic spots such as the Great Wall are closed after 5:30 pm. As a result, there are far more tourists remaining in the afternoon than the tourists who come in, and there are no visitors after 18:00 pm. For a general leisure park such as the Olympic Forest Park, another entering peak occurs after work from 17:00 pm to 18:00 pm, and a departure peak appears from 21:00 pm to 22:00 pm, consistent with daily behaviour.

With regard to another aspect, the number of tourists on the weekends and holidays is much higher than that on working days as shown in Figure 8. Examining the number of tourists in the scenic spots, through time dimensions we find that the number of tourists on weekends far exceeds the number of tourists on working days in the traditional scenic spots. The tourist number in the summer holidays is apparently higher than other seasons. And the tourist flow reached the peaks in holidays such as the Spring Festival, the Qingming Festival, the Dragon Boat Festival and the National day. The reason is that the resident population in Beijing has more time to go out a tour on weekends and holidays.

Comparing the total number of tourists in different scenic spots through spatial dimensions, we obtain the tourist flow order in Beijing shown in Figure 9. Some scenic spots are the hot spots for the tourists.

**4.2. OD Analysis of the Scenic Spot.** To vividly describe the origin and destination of tourists, taking the Summer Palace as an example, we conduct an analysis of tourists (total 1,733 tourists) who visited the palace from 8 am to 11 am on February 8, 2015. Combined with the tourist OD analysis algorithm above, the tourist origin and destination information of the Summer Palace scenic area are obtained together with the GIS platform. Figure 10(a) shows the number of tourists visiting the Summer Palace, whereas Figure 11(b) depicts the number of tourists who move from the Summer Palace to other places.

To better distinguish the different tourist flows, we utilize the thickness and colour of arrows and lines. As the flow rate increases, the width of the arrows and lines increases gradually, and the colour changes from green to red. From the spatial distribution map of the tourists, we can determine whether the origin of tourists (tourist attractions in the

Summer Palace area) or the whereabouts of tourists (Summer Palace area tourists) show “wave” spread, layer by layer, in line with regular traffic rules. Although tourists generally follow the law of “the distance increases and the amount decreases” after one tour, due to the particularity of the scenic spots, tourists choose to visit not only nearby scenic spots such as Summer Palace and Tsinghua, Peking University, but also farther scenic spots such as the Forbidden City and the Beijing Zoo.

To fully and comprehensively analyse the origin and destination of tourists in the Summer Palace, the data not only should be described qualitatively and visually from the space level but also should be described accurately and quantitatively from the statistics level to clarify the relationship between travel and distance.

First, the latitude and longitude of the centre point in the Summer Palace are extracted directly from the GIS platform to calculate the distance between the origin point (the ascent point) of each tourist and the centre of the scenic spot, that is, the geographical distance (spherical distance) between two points on the map. Then, we calculate the number of the above distances that are the same. Third, the arrival and departure probabilities of the tourists in the Summer Palace are calculated. Finally, we fit the distance distribution of the tourists by taking the travel distance as the abscissa and the scenic attraction quantity (occurrence quantity) probability as the ordinate. It is found that the composite exponential function is the best fit. Among the calculations, the fitting formula of the distance distribution of tourist origin is shown as follows:

$$P_O(d) = 0.5065 * e^{(-0.4146d)} + 0.1078 * e^{(-0.05757d)} \quad (1)$$

$$P_D(d) = 0.7927 * e^{(-0.9213d)} + 0.06561 * e^{(-0.1281d)} \quad (2)$$

In the formula,  $d$  is the distance between the scenic spot and the visitor’s location;

$P_O(d)$ , in the case of  $d$ , is the attraction probability of the scenic spots;

$P_D(d)$ , in the case of  $d$ , is the occurrence probability of the scenic spots.

From the formula and Figure 11, it is observed that the origin and destination of tourists are mainly distributed in the close range. That is to say, the majority of tourists still prefer to visit the closer scenic spots.

## 5. Conclusions

This paper presents a method based on the CDR data to analyse the tourist flow of scenic spots, including the collection and processing of the CDR data, tourist flow, travel OD, and other statistical analysis, which is all the helpful information for the operation management of the scenic spots. The conclusion is as follows:

(1) Through an analysis with the CDR data in the scenic spots in Beijing, the results show that the method can effectively analyse the tourist flows and other behaviour information, which can provide big data support to alleviate



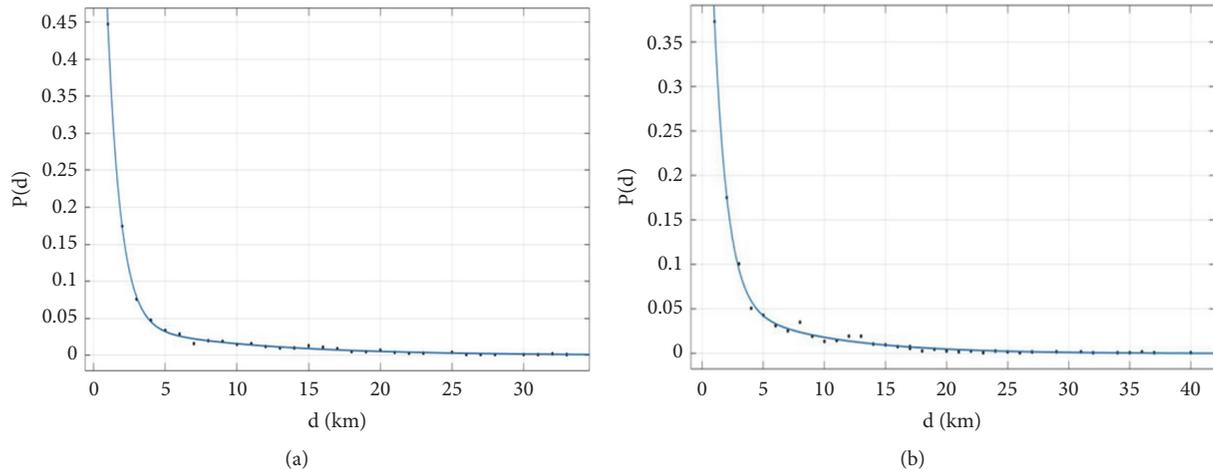


FIGURE 11: Distance distribution of visitors in the Summer Palace. (a) Origin distance distribution of tourists. (b) Destination distance distribution of tourists.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported in part by the National Science and Technology Pillar Program of China (Grant 2014BAG01B02) and Graduate Education Fund (Grant 145260522) and the 2014 National Advanced Training Course (Grant KIL14018530).

## References

- [1] J. Mair and M. Whitford, "An exploration of events research: event topics, themes and emerging trends," *International Journal of Event and Festival Management*, vol. 4, no. 1, pp. 6–30, 2013.
- [2] G. Yao, "A study on the construction framework of intelligent tourism," *Journal of Nanjing University of Posts & Telecommunications*, vol. 12, no. 2, pp. 13–16, 2012.
- [3] Y. B. Yan and H. E. Wen-Juan, "Analysis on the time-space evolution of the domestic tourism flow to China," *Economic Geography*, 2013.
- [4] N. Xianzhong et al., "On the future and the characteristics about domestic tourist flow to jiuzhaigou," *Journal of Mountain Research*, 1998.
- [5] R. Chen, C.-Y. Liang, W.-C. Hong, and D.-X. Gu, "Forecasting holiday daily tourist flow based on seasonal support vector regression with adaptive genetic algorithm," *Applied Soft Computing*, vol. 26, pp. 435–443, 2015.
- [6] A. Worobiec, L. Samek, P. Karaszkiwicz et al., "A seasonal study of atmospheric conditions influenced by the intensive tourist flow in the Royal Museum of Wawel Castle in Cracow, Poland," *Microchemical Journal*, vol. 90, no. 2, pp. 99–106, 2008.
- [7] X. Wang, H. Dong, Y. Zhou, K. Liu, L. Jia, and Y. Qin, "Travel distance characteristics analysis using call detail record data," in *Proceedings of the 29th Chinese Control and Decision Conference (CCDC '17)*, pp. 3485–3489, May 2017.
- [8] E. C. Higham, "Tourist flow reasoning: the spatial similarities of tourist movements," 1996.
- [9] D. G. Wang, "The influence of Beijing-Shanghai high-speed railway on tourist flow and time-space distribution," *Tourism Tribune*, vol. 29, no. 1, pp. 75–82, 2014.
- [10] S. E. Zhong et al., "Spatial patterns of tourist flow: problems and prospects," *Human Geography*, 2010.
- [11] H. Dong, X. Wang, C. Zhang, R. He, L. Jia, and Y. Qin, "Improved Robust Vehicle Detection and Identification Based on Single Magnetic Sensor," *IEEE Access*, vol. 6, pp. 5247–5255, 2018.
- [12] A. Soltani, M. Tanko, M. I. Burke, and R. Farid, "Travel patterns of urban linear ferry passengers: analysis of smart card fare data for brisbane, queensland, Australia," in *Transportation Research Record Journal of the Transportation Research Board*, no. 2535, pp. 79–87, Transportation Research Board of the National Academies, Washington, DC, USA, 2015.
- [13] S. Rahman, J. Wong, and C. Brakewood, "Use of mobile ticketing data to estimate an origin-destination matrix for new york city ferry service," in *Transportation Research Record Journal of the Transportation Research Board*, no. 2544, pp. 1–9, Transportation Research Board of the National Academies, Washington, DC, USA, 2008.
- [14] F. Calabrese, M. Diao, G. Di Lorenzo, J. Ferreira, and C. Ratti, "Understanding individual mobility patterns from urban sensing data: a mobile phone trace example," *Transportation Research Part C: Emerging Technologies*, vol. 26, pp. 301–313, 2013.
- [15] L. Ferrari, M. Mamei, and M. Colonna, "Discovering events in the city via mobile network analysis," *Journal of Ambient Intelligence and Humanized Computing*, vol. 5, no. 3, pp. 265–277, 2014.
- [16] R. Ahas et al., "Mobile Positioning in Space& Time Behaviour Studies: Social Positioning Method Experiments in Estonia," *American Cartographer*, vol. 34, no. 4, pp. 259–273, 2007.
- [17] R. Ahas, A. Aasa, Ü. Mark, T. Pae, and A. Kull, "Seasonal tourism spaces in Estonia: Case study with mobile positioning data," *Tourism Management*, vol. 28, no. 3, pp. 898–910, 2007.

- [18] R. Ahas, A. Aasa, A. Roose, Ü. Mark, and S. Silm, "Evaluating passive mobile positioning data for tourism surveys: An Estonian case study," *Tourism Management*, vol. 29, no. 3, pp. 469–486, 2008.
- [19] R. Ahas, A. Aasa, S. Silm, and M. Tiru, "Daily rhythms of suburban commuters' movements in the Tallinn metropolitan area: case study with mobile positioning data," *Transportation Research Part C: Emerging Technologies*, vol. 18, no. 1, pp. 45–54, 2010.
- [20] H. Dong, M. Wu, X. Ding et al., "Traffic zone division based on big data from mobile phone base stations," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 278–291, 2015.
- [21] H. Dong, M. Wu, Q. Shan et al., "Urban residents travel analysis based on mobile communication data," in *Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems: Intelligent Transportation Systems for All Modes (ITSC '13)*, pp. 1487–1492, October 2013.
- [22] Y. Etison et al., "System and method for real-time monitoring of a contact center using a mobile computer," US20150098561, 2015.

## Research Article

# An Improved Genetic Algorithm Based Robust Approach for Stochastic Dynamic Facility Layout Problem

Yunfang Peng, Tian Zeng, Lingzhi Fan, Yajuan Han, and Beixin Xia 

School of Management, Shanghai University, Shanghai 200444, China

Correspondence should be addressed to Beixin Xia; [bxia@shu.edu.cn](mailto:bxia@shu.edu.cn)

Received 31 August 2018; Accepted 25 November 2018; Published 5 December 2018

Guest Editor: Xinchang Wang

Copyright © 2018 Yunfang Peng 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 deals with stochastic dynamic facility layout problem under demand uncertainty in terms of material flow between facilities. A robust approach suggests a robust layout in each period as the most frequent one falling within a prespecified percentage of the optimal solution for multiple scenarios. Monte Carlo simulation method is used to randomly generate different scenarios. A mathematical model is established to describe the dynamic facility layout problem with the consideration of transport device assignment. As a solution procedure for the proposed model, an improved adaptive genetic algorithm with population initialization strategy is developed to reduce the search space and improve the solving efficiency. Different sized instances are compared with Particle Swarm Optimization (PSO) algorithm to verify the effectiveness of the proposed genetic algorithm. The experiments calculating the cost deviation ratio under different fluctuation level show the good performance of the robust layout compared to the expected layout.

## 1. Introduction

Nowadays, the increasingly fierce market competition and the growing variable demands of customers have gradually made the production mode shift from high volume and repetitiveness to high mix and low volume. How to meet the variant production requirements is thought to be an important objective for new intelligent manufacturing system, which is an advanced manufacturing mode of next generation. Designing agile facilities makes much sense to satisfy these requirements [1]. Dynamic changes, such as fluctuations in product quantity, varieties in product mix, introduction of new products, and discontinuation of existing products, are frequently taking place in high-mix and low-volume production environment. As a result of these changes, the previous layout becomes less efficient, which makes material handling cost increased. The uncertainty of demand has brought great challenges to design a suitable facility layout. Therefore, it is of great theoretical and practical significance to study the facility layout problem under dynamic environment [2].

The research on facility layout problem (FLP) started in the 1950s. Koopmans and Beckmann [3] initially defined the

facility layout problem as the assignment of facilities to discrete locations with the objective of minimizing the material handling cost. Static facility layout is obtained according to the deterministic material flow when the demand is constant. But when the demand changes frequently with time, static layout becomes no more suitable for various periods in the planning horizon. Dynamic facility layout problem (DFLP) divides the planning horizon into several discrete time periods, with different product demand. Designing flexible layout and designing robust layout are two approaches to deal with DFLP [4]. Flexible layout optimization is to design an optimal layout for each period in the multiperiod planning horizon so that the total material handling and rearrangement cost is minimized. The facility layout is changed according to the production demand in different periods. On the other hand, the robust approach is to design a fixed robust layout to minimize the total material handling cost over the entire time planning horizon [5]. Although the robust layout is not an optimal layout for a particular time period, its performance is good in each period. In a traditional DFLP, the demand in each period is determined by demand forecasting. But, in reality, the product demand in one period is difficult to forecast accurately. Therefore, it would be more valuable to

study DFLP, assuming the product demand is stochastic in each period. Considering uncertainty of the product demand in each period leads to stochastic dynamic facility layout problem (SDFLP) [6].

## 2. Literature Review

The approaches for the DFLP can be classified into four categories: exact methods, heuristics, metaheuristics, and hybrid approaches [1, 4]. Rosenblatt [7] who is the pioneer in DFLP presented a dynamic programming (DP) formulation. Based on this formulation, both optimal and heuristic procedures are developed. Lacksonen and Ensore [8] extended five algorithms to solve the DFLP, which is modeled as a modified Quadratic Assignment Problem (QAP) formulation. And the cutting plane algorithm was illustrated to be the best for all the test problems. To avoid the computational complexity of the DP and QAP formulations, Urban [9] proposed a new heuristic algorithm based on the steepest-descent pairwise-interchange procedure, which performs well in most of situations. Recently, more specifically metaheuristic and hybrid approaches, such as genetic algorithm (GA), tabu search (TS), and simulated annealing (SA), have been widely applied for DFLP. Kaku and Mazzola [10] defined a TS heuristic utilizing a dynamic tabu list for dynamic plant layout problem. Madhusudanan et al. [5] applied a robust layout to minimize the total material handling cost over all periods. They proposed a mathematical model for robust approach and designed a SA algorithm to solve the model. Tayal and Singh [11] presented mathematical formulation for multiobjective stochastic dynamic facility layout problem and solved it by SA and chaotic simulated annealing (CSA) metaheuristics. The experiment results observed that CSA performs better than SA. GA has been proven to be effective to generate suboptimal solutions for large-scale dynamic facility layout problems. Fazlelahi et al. [12] devised a customized permutation-based robust genetic algorithm in dynamic manufacturing environments, which is expected to be generating a unique robust layout for all the manufacturing periods.

Kulturel-Konak surveyed recent developments in designing robust and flexible facilities layout under uncertainty [1]. Webster and Tyberghein [13] measured the flexibility of a layout as the ability to react to disturbances caused by future change. They analyzed the annual material handling costs to measure the flexibility. Gupta [14] solved the FLP by Monte Carlo simulation to randomly generate the flow between all pairs of departments. Chan and Malmberg [15] also used Monte Carlo simulation to empirically search for robust solutions for dynamic line layout problem. Rosenblatt and Lee [16] solved the single period plant layout problem under stochastic demand by a robustness approach. They defined the robustness of a layout as the frequency it falls within a prespecified percentage of the optimal solution for various sets of scenarios. The robustness approach searches for a reliable layout for all scenarios but not the optimal layout for any given scenario. Besides the scenario-based robust optimization to deal with uncertainty, some other methods such as stochastic programming or fuzzy programming are widely used. Stochastic programming employs probabilistic

models and describes the uncertainty by probability distributions. Moslemipour and Lee [6] considered the randomly changing product demands as independent normally distributed random variables with known probability density function. SA metaheuristic algorithm was utilized to solve the mathematical model. Fuzzy programming models uncertain parameters with fuzzy numbers and establishes constraints using fuzzy sets and membership functions. Considering the uncertainty of material flows, Cheng et al. [17] introduced fuzzy numbers to represent the material flows between department pairs. Then, GA was applied to solve this hard fuzzy combinatorial problem. Kaveh et al. [18] modeled the DFLP as fuzzy programming and solved the models by a hybrid intelligent algorithm including GA, simulated annealing, and fuzzy simulation.

Although previous studies have significantly improved FLP with uncertainty, most of articles assumed that the demand is in exact probability distribution or is defined by fuzzy numbers. In fact, the information about the uncertainty is sometimes lacking and its behavior is difficult to predict. Therefore, a scenario-based method is applied in this paper to describe the demand uncertainty. Designing flexible layout and designing robust layout are two approaches to cope with dynamic layout problem. In this article, these two approaches are combined. In each period, a robust layout inspired from Rosenblatt and Lee [16] for DFLP considering the assignment of transport devices under uncertain demands is present. The robust layout is the most frequent layout falling within a prespecified percentage of the optimal solution for different sets of scenarios generated by Monte Carlo simulation. To improve the search speed of finding the robust layout, an improved adaptive genetic algorithm with population initialization strategy is proposed to reduce the search space and improve the efficiency of solving the model.

The rest of this paper is organized as follows. In the next section, the dynamic facility layout problem that considers the assignment of transport devices is modelled. Then, a robust approach based on Monte Carlo simulation is proposed to deal with material flow uncertainty in Section 3. After that, an improved genetic algorithm is developed to solve the mathematical model. Some numerical results are compared and the advantages of the robust layout are illustrated in Section 5. Finally, the conclusions are given.

## 3. Dynamic Facility Layout Problem

DFLP considers material flow over multiple time periods. The material flow between facilities changes over time. But traditional dynamic layout optimization is studied under the condition that the demand in each period is constant; it cannot effectively solve the problem with demand fluctuation. Therefore, on the basis of proposing an improved adaptive genetic algorithm, Monte Carlo simulation method is used to describe the effect of demand fluctuation on the material flow. In our research, the assignment of transport devices (such as conveyor, AGV, and tow train) is an important decision because of different unit material handling cost for each transport device. Although we consider the fluctuation of material flow in each period, the problem becomes a

TABLE 1: Notations.

Indices:	
$i, j$	Indices for facilities where $i, j=1, \dots, I$ , $I$ = number of facilities in the layout
$t$	Index for periods where $t=1, \dots, T$ , $T$ =number of periods in the planning horizon
$k, h$	Index for locations of facility where $k, h=1, \dots, L$ , $L$ =number of locations in the layout
$z$	Indices for transport device types where $z=1, \dots, G$ , $G$ =number of transport device types
Parameters:	
$D_{kh}$	The distance between location $k$ and location $h$
$A_{kh}$	The cost of rearrangement from location $k$ to location $h$
$P_{khz}$	The cost of transporting unit material from position $k$ to position $h$ using device $z$
$Q_{tij}$	The random variable of the total material flow from facility $i$ to facility $j$ in period $t$
Variables:	
$x_{tik}$	Binary variable:1, if facility $i$ is placed at location $k$ in period $t$ ; 0 otherwise
$y_{tzij}$	Binary variable:1, if using device $z$ to transport material between facility $i$ and facility $j$ during period $t$ ; 0 otherwise

determined dynamic facility problem when one scenario is generated.

The mathematical model for DFLP is discussed as follows and Table 1 gives the notations used.

The objective (3) is to minimize the total material handling cost (see (1)) and the rearrangement cost (see (2)). Constraints (4) ensure that each facility should be placed in exactly one location in each period. Constraints (5) indicate that a location can place at most one facility in each period. Constraints (6) ensure that only one transport device can be used between two facilities in each period.

$$MHC = \sum_{t=1}^T \sum_{z=1}^Z \sum_{i=1}^I \sum_{j=1}^I \sum_{k=1}^L \sum_{h=1}^L x_{tik} \cdot x_{tjh} \cdot y_{tzij} \cdot Q_{tij} \quad (1)$$

$$\cdot P_{khz} \cdot D_{kh}$$

$$AC = \sum_{t=1}^{T-1} \sum_{i=1}^I \sum_{k=1}^L \sum_{h=1}^L x_{tik} \cdot x_{(t+1)ih} \cdot A_{kh} \quad (2)$$

$$\min TC = MHC + AC \quad (3)$$

$$\sum_{k=1}^L x_{tik} = 1, \quad \forall t, i \quad (4)$$

$$\sum_{i=1}^I x_{tik} \leq 1, \quad \forall t, k \quad (5)$$

$$\sum_{z=1}^G y_{tzij} = 1, \quad \forall t, i, j \quad (6)$$

#### 4. Robust Approach Based on Monte Carlo Simulation

Considering demand uncertainty in terms of  $Q_{tij}$ , Monte Carlo simulation method is used to generate  $N$  different scenarios. In each period, we find a robust layout, which is the

most frequent layout falling within a prespecified percentage of the optimal solution for different sets of production scenarios. Robust layout constraints (7) and (8) are defined to ensure that the gap of cost between robust layout and optimal layout in each scenario falls within a prespecified percentage.

$$RCC_n = \frac{TC_n - TC_n^{opt}}{TC_n^{opt}} \quad (7)$$

$$RCC_n \leq M \quad (8)$$

In these constraints,  $n$  is the index for scenarios where  $n=1, \dots, N$ .  $RCC_n$  is the robust control coefficient of the  $n$ th scenario.  $TC_n^{opt}$  is optimal cost of the  $n$ th scenario, while  $TC_n$  denotes the total cost of the robust layout of the  $n$ th scenario.  $M$  is the prespecified percentage that is suggested to be less than 15% [19]. The detailed process to find the robust layout is as follows (Figure 1).

*Step 1.* Initially set  $n=1$  and candidate solution set  $S_n = \Phi$ .

*Step 2.* Generate the expected scenario with the material flow  $Q_{tij}$  setting to the expected value.

*Step 3.* Obtain the optimal layout  $Layout(n)$  and minimum cost  $TC_n^{opt}$  by genetic algorithm.

*Step 4.* Generate the initial population and improve the solution through evolution process until the solution satisfies the robust layout constraints. Add current layout into the candidate solution set  $S_n$ .

*Step 5.* If the number of elements in  $S_n$  is less than  $N_{max}$ , go back to the fourth step; otherwise, output set  $S_n$  and go to next step.

*Step 6.* If the number of simulations is less than  $N$ ,  $n=n+1$ . Generate a new scenario corresponding to the random numbers  $Q_{tij}$  in a certain range and go back to Step 2. Else, go to the next step.

*Step 7.* Output the most frequent layout (robust layout) contained in  $S_n$  where  $n=1..N$ . If more than two layouts are satisfied, search for the robust layout  $Layout'$  with minimal cost deviation from the optimal cost.

## 5. An Improved Adaptive Genetic Algorithm Design

Genetic algorithm (GA) is an excellent heuristic algorithm, which is a random optimization algorithm for the simulation of biological evolution in nature. Because of its strong practicability and robustness, it is widely used in the field of facility layout optimization, scheduling, and transportation. But genetic algorithm has the shortcomings of easy falling into local optimal and premature convergence. Therefore, in view of the above shortcomings, combined with the characteristics of the facility layout problems, an improved genetic algorithm is developed.

*5.1. Population Initialization.* Because the decision variables of the model are all 0-1 variables, binary coding is adopted. Each cell contains  $T$  chromosomes. The first  $I \times L$  position in a chromosome defines the layout of the facility, and the latter  $G \times I \times I$  position denotes the assignment of transport devices. In order to improve the efficiency of the algorithm and combine the constraint conditions of the model, a new population initialization strategy is proposed on the premise of ensuring diversity, which is as follows.

A sample of two facilities located in three alternative locations with three transport devices in two periods is shown in Figure 2. In the initialization of chromosome, the front-end part, due to  $I < L$ , generates a matrix  $A$  with  $L \times L$  dimensions, which has only one element with value 1 in each row and each column, such as  $A=[0,1,0; 1,0,0; 0,0,1]$ . After that, we randomly select  $I$  rows to compose matrix  $B$ ;  $B=[0,1,0; 0,0,1]$ , thus generating the layout gene position  $Chrom1x=[0,1,0,1,0,0]$ . The back-end part needs to be initialized according to the transport devices in each period. If  $Q_{tij}$  is 0, then directly generate  $1 \times G$  dimension zero vector. If  $Q_{tij}$  is not 0, generate the  $1 \times G$  dimension row vector with only one element of 1 and the rest is 0. Then it is converted to  $1 \times (G \times I \times I)$  gene position  $Chrom1y=[0,0,0,1,0,0,0,0,1,0,0,0,0]$ . Finally, it is combined with  $Chrom1x$ , so an initialization chromosome  $Chrom1=[0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0]$  is formed. According to the total number of time periods  $T$ , a cell is composed of  $T$  chromosomes.

*5.2. Selection Operation.* Selection operations are the process of eliminating individuals with low adaptation from an old population and selecting excellent individuals with high fitness. Compared to the common genetic algorithm, the selection process is divided into two steps, namely, cell selection and chromosome selection, due to the particularity of this problem. We use traditional roulette wheel selection method to select the cell and chromosome.

*5.3. Mutation and Crossover Operations.* Crossover is a random selection of individuals' genes and changes them to

produce evolutionary operations for new individuals. Mutation operation is a crucial step in genetic algorithm, which ensures the diversity of population and avoids premature convergence. However, for different optimization problems, it is necessary perform tests repeatedly to determine cross probability  $P_c$  and mutation probability  $P_m$ , which is a very tedious task, and it is difficult to find the best value for each problem. So the idea of adaptive genetic algorithm is applied to determine the probability of cross and mutation.

$$P = \frac{\max(\text{popfitness}) - \min(\text{popfitness})}{\max(\text{popfitness})} \quad (9)$$

Because binary coding is adopted in this problem, the mutation and cross operations are combined together. During cross and mutation operations, it is necessary to fragment the chromosomes so as to find the corresponding facilities, locations, and types of transport devices. For the front-end part  $Chrom1x=[0,1,0; 0,0,1]$ , as shown in Figure 3, assuming the mutation position is 2, the corresponding value 1 indicates that facility 1 is placed on location 2. Then value changes from 1 to 0, and the remaining two locations are randomly selected to place facility 1. If it is placed on location 1, there is no need for adjustment. If it is placed on position 3, we found that location 3 has already been placed by facility 2. Then, facility 2 is adjusted to the original variant location (location 2). If the mutation position is 3 corresponding to value 0, which means facility 1 is not placed on location 3, then change the value from 0 to 1, and meanwhile change the value 1 on position 2 to 0. After that change, we find that facilities 1 and 2 are both placed on location 3; then we randomly place facility 2 on the remaining locations. As a result, in condition 1, the values in positions 2 and 3 are crossed, and the values in positions 5 and 6 are exchanged. Meanwhile, in condition 2, the values in positions 1 and 2 are exchanged.

For the variation of the back-end part, as shown in Figure 4, if the mutation position is 2 corresponding to value 0, because this position represents no transport devices between facility 1 and facility 1, there is no need to do any operation. If the mutation position is 4 corresponding to value 0, we change value 0 to 1 and change the remaining position to 0. If the variation position is 5 corresponding to value 1, we change value 1 to 0. The residual position is randomly selected as 1. As a result, in conditions 2 and 3, the values on positions 4 and 5 are exchanged.

*5.4. Termination Criteria.* The algorithm is terminated when the generation reached the maximal number  $M$ , or the best value did not improve for  $0.05 \times M$  generations.

## 6. Computational Analysis

Because the transport device is not mentioned in most of articles, there are no standard benchmark instances. 20 determined instances with different size as article [20] are randomly generated. The size of instance is denoted by  $TJJL_G$ , where  $T$  is the number of periods,  $I$  is the number of facilities,  $L$  is the number of locations, and  $G$  is the number of transport device types. We compare the improved

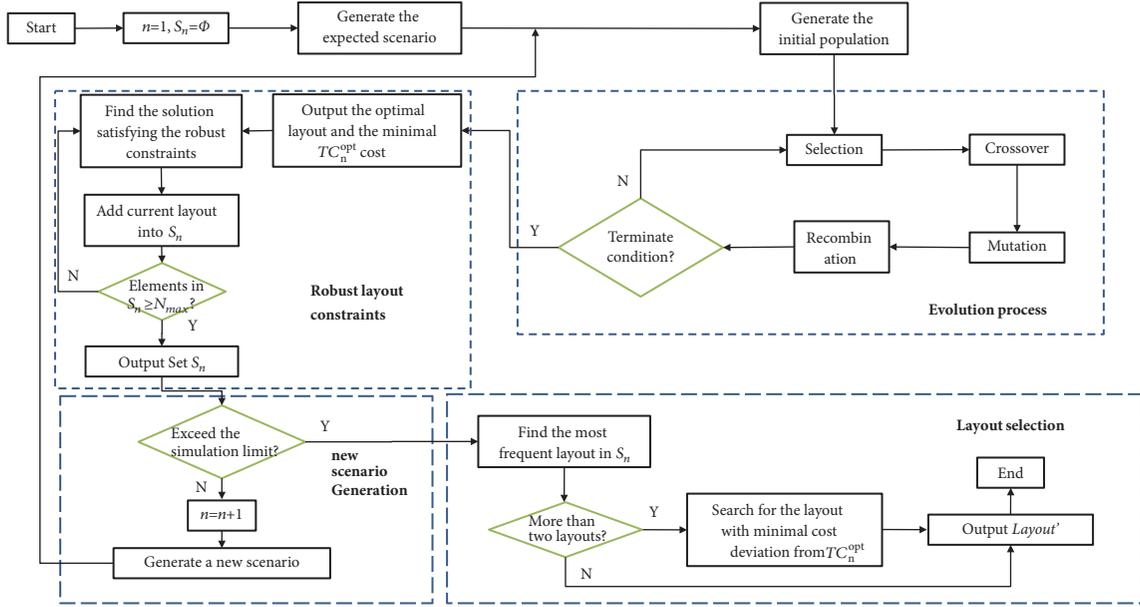


FIGURE 1: The robust layout generation process based on MCS.

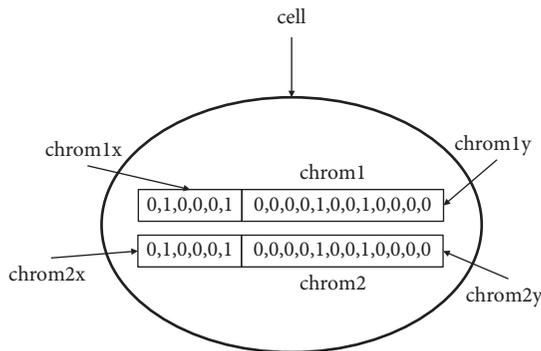


FIGURE 2: An example of a cell.

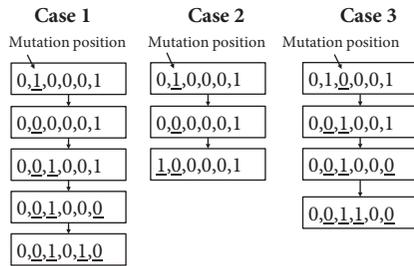


FIGURE 3: An example of mutation and cross operation of front-end part.

genetic algorithm with the Particle Swarm Optimization (PSO) algorithm proposed in [21]. The parameters are set as follows: the number of populations is 20 and the maximal number of generations is 100. In order to eliminate the influence of the random factors, the average values of the 20 simulation results are compared. A personal computer with

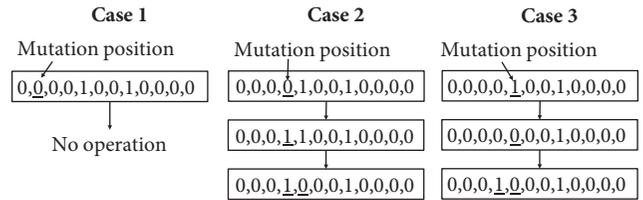


FIGURE 4: An example of mutation and cross operation of back-end part.

2.7GHz Intel Core i5, 8G RAM, is used to run the algorithms that are coded by matlab2014b. The results are shown in Table 2. It is found that the improved genetic algorithm is superior to the PSO in both solving speed and solution quality in almost all the instances except for instances 8\_24\_30\_10 and 14\_32\_36\_18. Especially for large-scaled instances, the improved genetic algorithm presented in this paper has more advantages.

In order to illustrate the effectiveness of the robust layout proposed in this paper, we compare it with the traditional expected layout (the optimal layout in each period with expected material flow). Cost departure ratio denoted by  $H$  is defined to describe the deviation ratio between the total cost of the evaluated layout and the optimal cost in each scenario.  $H_r$  represents the robust layout cost deviation ratio, while  $H_d$  represents expected layout cost deviation ratio.

$$H = \frac{\sum_{n=1}^N TC_n - \sum_{n=1}^N TC_n^{opt}}{\sum_{n=1}^N TC_n^{opt}} \quad (10)$$

For the stochastic variable material flow  $Q_{tij}$ , assume that the expected average value is  $a$ , and the fluctuation level

TABLE 2: Comparison of algorithms.

Instance scale $T\_J\_L\_G$	Improved genetic algorithm		PSO	
	Time(s)	Total cost	Time(s)	Total cost
4_8_12_4	40.4572	140	47.4022	140
4_10_16_6	58.7843	178	68.4300	178
6_14_20_6	94.8301	312	142.0803	314
8_20_26_8	180.2430	478	277.0293	480
8_24_30_10	210.4790	648	312.0019	642
10_24_30_8	260.4341	750	377.0341	750
10_30_36_14	380.4339	1078	500.0391	1090
14_32_36_18	1009.2215	1977	1228.0012	1975
14_38_42_22	1900.0034	2130	2277.1014	2134
20_48_52_32	7708.9988	4701	8792.1044	4710
28_60_68_48	15366.4222	7999	18722.0924	8073
35_70_80_60	23177.7788	9876	26718.8819	9915
40_80_90_70	32421.6644	13077	37877.0701	13203
45_90_100_80	40500.7721	18093	45722.3307	18210
50_100_110_90	56783.4451	24088	60421.7701	24191
55_105_115_95	69803.1111	34330	81445.7769	34700
60_110_120_100	77099.2299	40099	89001.7721	40280
60_115_125_105	84123.7798	46788	93100.1109	46989
65_120_130_110	93401.1194	52103	101012.1981	52708
65_125_135_115	99088.7711	58021	107662.0061	58810

is  $b$ . It means that the material flow is in the rage of  $[(1-b)*a, (1+b)*a]$ . We randomly generate  $N=200$  scenarios, and the prespecified percentage  $M$  in robust constraint is 15%. The expected layout is obtained in the scenario when  $Q_{tij}$  is equal to  $a$ . The cost deviation ratios  $H_r$  and  $H_d$  are calculated under different fluctuation level in three instances. The results are shown in Table 3.

From Table 3, we can see that, with the increase of fluctuation level, the cost deviation ratio will increase for both robust layout and expected layout. We can conclude that the greater the fluctuation is, the bigger difference between robust/expected layout cost and minimum cost is. Moreover, for all these three different sized instances, the cost deviation ratio of robust layout is much less than expected layout. As the problem size increases, the cost deviation ratio of expected layout raises, while the cost deviation ratio of robust layout reduces. It is obviously concluded that the robust approach proposed in this paper is effective to deal with dynamic layout problem under demand uncertainty, especially for large sized problem.

## 7. Conclusions

In this paper, an improved genetic algorithm is designed to solve dynamic facility problem considering the assignment of transport devices. The developed algorithm has given near-optimal solutions to different sized case studies randomly generated. In addition, a robust layout is suggested with

uncertain material flow in each period. The robust layout is the most frequent one falling within a prespecified percentage of the optimal solution for multiple scenarios generated by Mont Carlo simulation. Although robust layout may not be optimal in any scenario, the deviation from the best cost is small. Its performance is much better than the expected layout through sensitivity analysis.

More uncertain factors except for material flow can be investigated in our future research, such as the unit material handling cost and the flexible routing. Moreover, unequal-sized facilities can be considered and more efficient metaheuristic algorithm needs to be developed in the future.

## 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 work was supported by the National Natural Science Foundation of China (Grants no. 51405283, no. 71401098, and no. 71572104).

TABLE 3: Comparison of cost deviation ratio.

Fluctuation level	6_14_20_6		20_48_52_32		40_80_90_70	
	$H_r$	$H_d$	$H_r$	$H_d$	$H_r$	$H_d$
1.00%	4.11%	10.75%	2.71%	13.82%	1.72%	15.90%
1.45%	4.08%	10.88%	2.90%	14.50%	1.80%	16.98%
1.90%	4.22%	11.03%	3.11%	15.70%	1.83%	17.88%
2.35%	4.27%	11.77%	3.23%	16.10%	1.88%	19.09%
2.80%	4.29%	12.09%	3.29%	16.77%	1.91%	20.12%
3.25%	4.66%	12.80%	3.30%	17.23%	1.93%	21.98%
3.70%	4.88%	12.91%	3.45%	17.46%	1.94%	22.10%
4.15%	4.91%	13.70%	3.56%	17.56%	1.96%	22.70%
4.60%	4.97%	13.92%	3.71%	18.10%	1.99%	22.90%
5.05%	4.99%	14.10%	3.77%	18.22%	2.01%	23.11%
5.50%	5.01%	14.45%	3.89%	18.28%	2.09%	23.34%
5.95%	5.09%	14.67%	3.91%	18.45%	2.21%	23.56%
6.40%	5.03%	14.72%	3.94%	18.66%	2.32%	23.88%
6.85%	5.21%	14.92%	3.99%	18.88%	2.26%	23.99%
7.30%	5.27%	14.98%	4.01%	19.10%	2.45%	24.67%
7.75%	5.29%	15.21%	4.03%	19.23%	2.48%	24.99%
8.20%	5.33%	15.35%	4.07%	19.45%	2.29%	25.16%
8.65%	5.35%	15.62%	4.11%	19.57%	2.38%	25.37%
9.10%	5.66%	16.11%	4.13%	19.82%	2.44%	25.99%
9.55%	5.70%	16.32%	4.22%	20.03%	2.51%	26.81%
10.00%	5.72%	16.78%	4.25%	20.70%	2.68%	27.32%

## References

- [1] S. Kulturel-Konak, "Approaches to uncertainties in facility layout problems: Perspectives at the beginning of the 21st Century," *Journal of Intelligent Manufacturing*, vol. 18, no. 2, pp. 273–284, 2007.
- [2] S. Emami and A. S. Nookabadi, "Managing a new multi-objective model for the dynamic facility layout problem," *The International Journal of Advanced Manufacturing Technology*, vol. 68, no. 9-12, pp. 2215–2228, 2013.
- [3] T. C. Koopmans and M. Beckmann, "Assignment problems and the location of economic activities," *Econometrica*, vol. 25, pp. 53–76, 1957.
- [4] G. Moslemipour, T. S. Lee, and D. Rilling, "A review of intelligent approaches for designing dynamic and robust layouts in flexible manufacturing systems," *International Journal of Advanced Manufacturing Technology*, vol. 60, pp. 11–27, 2012.
- [5] V. Madhusudanan Pillai, I. B. Hunagund, and K. K. Krishnan, "Design of robust layout for dynamic plant layout problems," *Computers & Industrial Engineering*, vol. 61, no. 3, pp. 813–823, 2011.
- [6] G. Moslemipour and T. S. Lee, "Intelligent design of a dynamic machine layout in uncertain environment of flexible manufacturing systems," *Journal of Intelligent Manufacturing*, vol. 23, no. 5, pp. 1849–1860, 2012.
- [7] M. J. Rosenblatt, "Dynamics of Plant Layout," *Management Science*, vol. 32, no. 1, pp. 76–86, 1986.
- [8] T. A. Lacksonen and E. E. Enscore, "Quadratic assignment algorithms for the dynamic layout problem," *International Journal of Production Research*, vol. 31, no. 3, pp. 503–517, 1993.
- [9] T. L. Urban, "A Heuristic for the dynamic facility layout problem," *Institute of Industrial Engineers (IIE). IIE Transactions*, vol. 25, no. 4, pp. 57–63, 1993.
- [10] B. K. Kaku and J. B. Mazzola, "A tabu-search heuristic for the dynamic plant layout problem," *INFORMS Journal on Computing*, vol. 9, no. 4, pp. 374–384, 1997.
- [11] A. Tayal and S. P. Singh, "Formulating multi-objective stochastic dynamic facility layout problem for disaster relief," *Annals of Operations Research*, 2017.
- [12] F. Zarea Fazlelahi, M. Pournader, M. Gharakhani, and S. J. Sadjadi, "A robust approach to design a single facility layout plan in dynamic manufacturing environments using a permutation-based genetic algorithm," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 230, no. 12, pp. 2264–2274, 2016.
- [13] D. B. Webster and M. B. Tyberghein, "Measuring flexibility of job-shop layouts," *International Journal of Production Research*, vol. 18, no. 1, pp. 21–29, 1980.
- [14] R. M. Gupta, "Flexibility in layouts: A simulation approach," *Material Flow*, vol. 3, pp. 243–250, 1986.
- [15] W. K. Chan and C. J. Malmborg, "A Monte Carlo simulation based heuristic procedure for solving dynamic line layout problems for facilities using conventional material handling devices," *International Journal of Production Research*, vol. 48, no. 10, pp. 2937–2956, 2010.
- [16] M. J. Rosenblatt and H. L. Lee, "A robustness approach to facilities design," *International Journal of Production Research*, vol. 25, no. 4, pp. 479–486, 1987.
- [17] R. Cheng, M. Gen, and T. Tozawa, "Genetic search for facility layout design under interflows uncertainty," in *Proceedings of the 1995 IEEE International Conference on Evolutionary Computation*, pp. 400–405, December 1995.

- [18] M. Kaveh, V. M. Dalfard, and S. Amiri, "A new intelligent algorithm for dynamic facility layout problem in state of fuzzy constraints," *Neural Computing and Applications*, vol. 24, no. 5, pp. 1179–1190, 2014.
- [19] P. Kouvelis, A. A. Kurawarwala, and G. J. Gutiérrez, "Algorithms for robust single and multiple period layout planning for manufacturing systems," *European Journal of Operational Research*, vol. 63, no. 2, pp. 287–303, 1992.
- [20] A. Kheirkhah, H. Navidi, and M. Messi Bidgoli, "Dynamic facility layout problem: a new bilevel formulation and some metaheuristic solution methods," *IEEE Transactions on Engineering Management*, vol. 62, no. 3, pp. 396–410, 2015.
- [21] F. Jolai, R. Tavakkoli-Moghaddam, and M. Taghipour, "A multi-objective particle swarm optimisation algorithm for unequal sized dynamic facility layout problem with pickup/drop-off locations," *International Journal of Production Research*, vol. 50, no. 15, pp. 4279–4293, 2012.

## Research Article

# Optimizing the Composite Cost Involved in Road Motor-Transporting Trucks by Taking into Account Traffic Condition

Hongying Fei  and Chengyi Zhang

*School of Management, Shanghai University, China*

Correspondence should be addressed to Hongying Fei; [feihy@shu.edu.cn](mailto:feihy@shu.edu.cn)

Received 7 September 2018; Revised 28 October 2018; Accepted 15 November 2018; Published 2 December 2018

Guest Editor: Xinchang Wang

Copyright © 2018 Hongying Fei and Chengyi 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.

With the rapid development of modern automotive logistics industry, vehicle logistics has drawn more and more attention. Since the vehicle transporters mainly are the severe-polluting heavy-duty vehicles and their exhaust emissions vary under different traffic conditions, it is necessary to improve the planning of road motor-transporting services by taking into account road traffic condition, especially for urban areas. This study aims at minimizing the composite cost, including both the economic cost related to the driver cost and fuel consumption, and the social cost related to the vehicle emissions. The dynamic road traffic condition is imitated dynamically with a discretization technique. A metaheuristic is applied with data collected from a dense district in a huge city. Experimental results show that the proposed approach can always converge quickly to the best solution and the solution with minimal composite cost can always dominate the other solutions with classic route optimization goals.

## 1. Introduction

Besides the labor cost of truck drivers, vehicle fuel consumption and emissions are also critical aspects in the transportation planning process, especially for companies using heavy-duty trucks in recent years.

Having been the largest automobile market and factory, thousands of motors are being transported from their assembly factories to the 4S dealers by motor-transporting trucks, which are normally heavy-duty diesel trucks. It has been reported that the transportation cost takes up about 80% of the total logistics cost for the whole process of vehicle delivery in china, which is even higher than doubled cost in Europe. Therefore, it is quite important for the managers to arrange the vehicle delivery process in an efficient and effective manner.

Early vehicle routing problems are mainly focused on the minimization of economic transportation cost, travel time, or/and empty loading rate; however, it is no more the case in recent years. Since more and more people realize that the transportation plays a significant role in

air pollution, the corporations, especially logistics companies, are forced to take their responsibilities of reducing vehicles emissions while performing their delivery services.

Urban roads are limited resources, while the number of vehicles has been increasing dramatically in recent years. In consequence, traffic jams can be observed everywhere. Once involved in the traffic jam, the vehicles consume much more fuel, resulting in a high level of pollutant emissions. Thus the traffic condition, which dynamically changes, has an important impact on vehicles emissions.

How can we optimize the delivery routes of the motor-transporting trucks so as to minimize both economic cost and pollutant emissions? A non-traffic-jam shortest path may be ideal solution. Unfortunately, it is almost impossible to attain such situation in the real world, especially in metropolitan areas. On the contrary, we should make a concession to the existence of traffic congestions and take actions to improve the transportation plan based on a thorough investigation of the traffic condition in the targeted area, as is the objective of this study.

The paper is organized as follows: the literature review on time-dependent vehicle routing problems, especially the studies taking into account traffic conditions, is presented in Section 2. Section 3 is dedicated to the problem description and some preliminaries. The key points of the proposed metaheuristic are described in Section 4, and numerical experiments are presented in Section 5 and this paper ends up with an overall conclusion made in Section 6.

## 2. Literature Review

According to the literature, numerous researchers were interested in dealing with time-dependent vehicle routing problems (VRP) since this kind of problems arises naturally in a variety of applications [1]. Since the travel time impacts greatly the vehicle speed, which is an important factor of car emissions, the green vehicle routing problems, that is, the minimizing-emission VRPs, are normally considered as variants of time-dependent VRPs strongly related to road traffic conditions. In this section, a brief literature review of studies focused on time-dependent VRPs, especially traffic-condition-related VRPs, in recent years is given.

Time-dependent routing problems may be classified with respect to a number of criteria, such as the topology of the network, the objectives, and the constraints. Although some authors, such as Franceschetti et al. [2], provided an analytical characterization of the optimal solution for a single-arc version, most of the studies dealt with network version. As for the paths connecting two nodes, most of the work supposed that only one arc is defined between two nodes, though some others allowed multiple selection on arcs between two nodes [3, 4].

As for the objectives, it was observed that most of the early studies were considered to minimize the travel time [5–13], and social costs are considered widely in recent years, such as the congestion charge [14] and emissions [2, 15–19]. It reveals the importance the society attaches to the environmental protection.

With regard to the methodology, most of the studies were based on heuristics (e.g., [4, 7, 9, 12, 14, 20]) and metaheuristics, such as Tabu search [3, 5, 6, 13, 16], Variable Neighborhood Search [8], Ant Colony System [10], and particle swarm optimization [17, 21].

To sum up, vast publications can be observed in the literature related to time-dependent vehicle routing problem, and more and more scholars focused their studies on emission-minimizing problems by taking into account the traffic condition. Except some that provided analytic model for single-arc problem, most of the studies are based on heuristics or metaheuristics, though few obtain great achievements on dealing with the real-world traffic congestions, especially for the heavy-duty motor-transporting trucks. In this study, the dynamic change of traffic congestion is considered to obtain grounded evaluation of the velocity of motor-transporting trucks, and a metaheuristic will also be applied to solve this time-dependent vehicle routing problem with an objective to minimize both economical cost (driver cost and fuel consumption) and social cost (emission of  $\text{CO}_2$ , CO, and  $\text{NO}_x$ ).

## 3. Problem Description

*3.1. Assumptions.* As described in the previous section, in general, motor-transport trucks start from a given depot and deliver cars to 4S dealers' parks, with each 4S dealer being served by one truck, according to predefined orders, and finally return back to the same depot to complete the route. The aim of this paper is to define the best path for each motor-transport truck by taking into account the uncertainty of road traffic condition, especially the congestion, so as to minimize composite cost, composed of not only the economic cost, fuel cost, and maneuver cost but also the social cost represented as the emission of those trucks. It is worth mentioning that since the vehicle emission depends on their running conditions which are closely correlated with the congestion of the route, the distribution of road congestion has been discretized in this study to imitate the real situation in an efficient way.

Some important assumptions are as follows:

- (1) Distribution of the road traffic congestion situation is investigated in advance.
- (2) There is one depot. The locations of the depot and 4S dealers are given.
- (3) The capacity of a truck is enough to serve any 4S dealer, and the requirements of a 4S dealer cannot be shared among different motor-transport trucks.
- (4) The route of each truck starts from and eventually ends at the storage depot.
- (5) The storage in the depot could meet all customers' needs.
- (6) Not only will the motors be transported but also the motor-transport trucks are uniform.
- (7) The composite cost is composed of three parts: fuel consumption, drivers' salaries, and the exhaust emissions of motor-transport trucks.
- (8) Tree traffic conditions are considered, as usually shown in various navigation maps using green, yellow, and red. The trucks run with at the maximum authorized speed under "green" condition, the limited speed under "yellow" condition and the extremely low speed, defined as 10 in this study, under "red" condition. The distributions of various road traffic conditions during the whole day have been discretized into 24 intervals, each corresponding to one hour, and evaluated after we conducted an analysis of the real situation of each route with the data collected from navigation maps during certain observation period.
- (9) One and only one driver is responsible for a motor-transport truck.
- (10) No stop accounts for road junctions.

*3.2. Notations.* The vehicle routing problem considered in this study can be described in an undirected graph  $G = \{N, A\}$ .  $N$  represents the nodes, including the depot  $i_0$ , 4S

dealers, and road junctions.  $A$  represents the set of arcs linking pairs of nodes according to the traffic road map.

#### Notation

$H$ : Total number of motor-transporting trucks
$\Omega$ : Global set of nodes, including the depot, 4S dealers, and road junctions
$\Omega_L$ : Set of 4S dealers
$\Omega_h$ : Set of nodes passed by the truck $h \in \{1, \dots, H\}$
$P_d$ : Driver's hourly salary
$P_e$ : Estimated emission-treatment fee per km per vehicle
$P_f$ : Fuel consumed per km per vehicle
$C_d$ : Total driver cost
$C_f$ : Total fuel-consumption cost
$C_s$ : Total social cost considering the treatment of car emissions
$t_i^h$ : Departure time of truck $h$ at node $i$
$t_F^h$ : Arrival time of truck $h$ at the depot, node 0

$\Delta t$ : Length of unit time interval used in this study to discretize the distribution of traffic condition

$d_{ij}$ : Travel distance between nodes  $i$  and  $j$

$v_{ij}^G$ : Maximum speed, under "Green" traffic condition, along the route connecting nodes  $i$  and  $j$

$V_{ij}^Y$ : Speed of the truck under "Yellow" traffic condition along the route connecting nodes  $i$  and  $j$

$V_{ij}^R$ : Speed of the truck under "Red" condition along the route connecting nodes  $i$  and  $j$

$\bar{V}_{ij}(t)$ : Expected speed of the truck on the route connecting nodes  $i$  and  $j$  at time  $t$

$P_{ij}^G(t)$ : Probability of "Green" condition for the route  $(i, j)$  at time  $t$

$P_{ij}^Y(t)$ : Probability of "Yellow" condition for the route  $(i, j)$  at time  $t$

$P_{ij}^R(t)$ : Probability of "Red" condition for the route  $(i, j)$  at time  $t$

$T_0$ : Earliest possible leaving time from the depot

$T_f$ : Latest possible arrival time at the depot

#### Decision Variables

$$x_{ij}^h(t) = \begin{cases} 1, & \text{truck } h \text{ takes the route connecting nodes } i \text{ and } j \text{ at time } t \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

**3.3. Preliminaries.** As mentioned in the previous section, this study aims at minimizing the composite cost including the labor cost  $C_d$  related to drivers' salaries, the fuel cost  $C_f$  related to the consumption of fuel along the trajectories, and the social cost  $C_s$  considering the treatment of car emissions. In this part, the calculations of these costs are detailed, respectively.

**3.3.1. Evaluation of Drivers' Labor Cost.** The drivers' labor cost,  $C_d$ , is mainly determined by the total time used by truck drivers to complete their missions. In this study, the total travel time of trucks is applied to evaluate the drivers' labor cost.

Since the drivers must be along with their trucks from their departure from the depot until their arrival at the same depot, in one word, drivers' labor costs can be defined as the drivers' hourly pay,  $P_s$ , multiplied by their traveling time:

$$C_d = P_s \cdot \sum_{h=1}^H (t_0^h - t_F^h) \quad (2)$$

where  $t_0^h$  and  $t_F^h$ ,  $h \in \{1, \dots, H\}$  denote the departure time of truck  $h$  from the depot and the arrival time of that truck at the depot, respectively. Obviously, all  $t_0^h$ ,  $h \in \Omega_h$  are normally

given in advance according to the delivery schedule, while  $t_F^h$  depends on not only the departure of truck  $h$  from the depot but also the traffic condition of the road which is uncertain. In this study, the distribution of the road traffic condition for each segment of the road is predefined according to the survey from a popular navigation map during a certain observation period and then discretized into several parts, each corresponding to one time unit, so that the arrival time of the truck at successive nodes can be calculated dynamically with a recursive formula. The specific solving process for a certain segment of the route is shown in Figure 1 as an example.

According to the dynamic threshold shown in a popular application of navigation map, one of the most popular navigation maps in China, dynamic traffic conditions are recorded and generally visualized with three colors. If one route's road condition is quite clear, the corresponding line on the map is marked green and the vehicle can run at the maximum authorized speed. If the vehicles on one route should slow down to a predefined threshold (e.g., 30 km/h in Baidu Map), the corresponding line is marked yellow and the vehicles are supposed to run at that threshold. Red lines in the navigation maps indicate that those lines are facing congestion and no cars can run at a speed over 5 km/h; that is, all the cars must experience stop-and-goes.

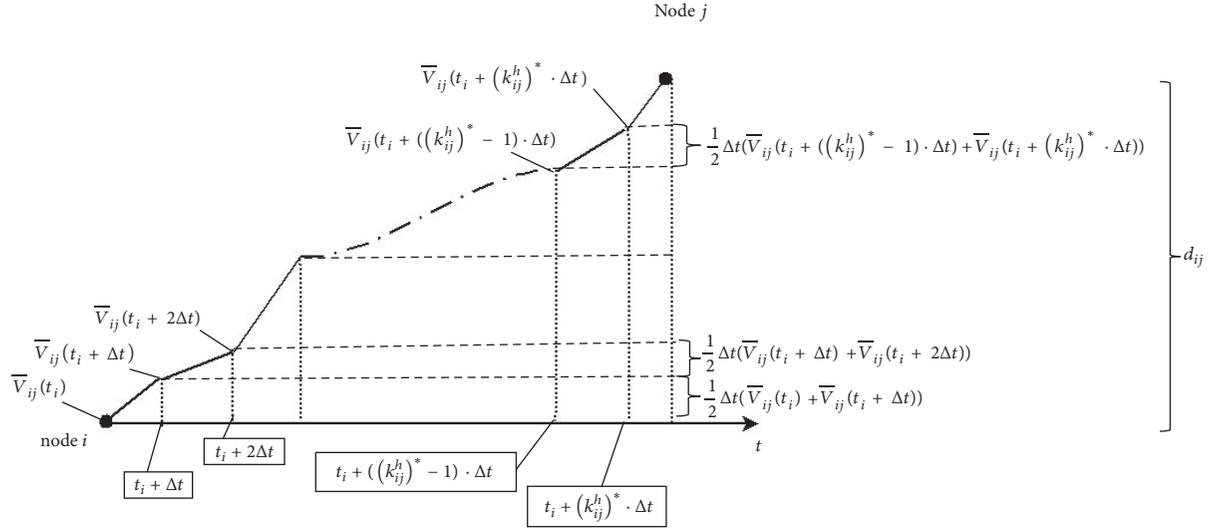


FIGURE 1: Description of how to define the arrival time of the truck at the successive node in condition that the departure time at the precedent node is given.

Suppose that a truck passed by road junction  $i$  at time  $t_i^h$  towards road junction  $j$ , where  $h \in \{1, \dots, H\}$ ,  $i \in \Omega$ , and  $j \in \Omega$ . The route between these two road junctions (nodes) can be divided into  $(k_{ij}^h)^* + 1$  segments, where any of the first  $(k_{ij}^h)^*$  segments corresponds to the distance, covered by the truck, which is calculated according to the current traffic condition within one unit time interval  $\Delta t$ , and the truck's travel duration for the last segment may be part of the time unit  $\Delta t$ , as shown in Figure 1.

Since a predefined distribution of traffic condition is given for each segment at any time, the expected speed of any vehicle passing through the route that connects nodes  $i$  and  $j$  at time  $t$  can be calculated with the following formula:

$$\bar{V}_{ij}(t) = V_{ij}^G(t) \cdot P_{ij}^G(t) + V_{ij}^Y(t) \cdot P_{ij}^Y(t) + V_{ij}^R(t) \cdot P_{ij}^R(t) \quad (3)$$

In consequence, if the truck  $h$  departed from node  $i$  at time  $t_i^h$ , its arrival time at the successive node  $j$  can be defined as follows:

$$(k_{ij}^h)^* = \operatorname{argmin} \left\{ \Delta t \cdot \left[ \frac{1}{2} \bar{V}_{ij}(t_i^h) + \sum_{k'=1}^{k-1} \bar{V}_{ij}(t_i^h + k' \cdot \Delta t) + \frac{1}{2} \bar{V}_{ij}(t_i^h + k \cdot \Delta t) \right] \mid \Delta t \right. \quad (4)$$

$$\cdot \left[ \frac{1}{2} \bar{V}_{ij}(t_i^h) + \sum_{k'=1}^{k-1} \bar{V}_{ij}(t_i^h + k' \cdot \Delta t) + \frac{1}{2} \bar{V}_{ij}(t_i^h + k \cdot \Delta t) \right] > d_{ij} \left. \right\}$$

$$t_j^h = t_i^h + (k_{ij}^h)^* \cdot \Delta t + \varepsilon_{ij}(t_i^h) \quad (5)$$

where

$$\varepsilon_{ij}(t_i^h) = \frac{d_{ij} - \Delta t \left[ (1/2) \bar{V}_{ij}(t_i^h) + \sum_{k'=1}^{(k_{ij}^h)^*-1} \bar{V}_{ij}(t_i^h + k' \cdot \Delta t) + (1/2) \bar{V}_{ij}(t_i^h + (k_{ij}^h)^* \cdot \Delta t) \right]}{\bar{V}_{ij}(t_i^h + (k_{ij}^h)^* \cdot \Delta t)} \quad (6)$$

Constraint (4)-(6) indicates the calculation of the passing time of the truck  $h$  at each node by discretizing the route that connects the pair of nodes with segments covered by the truck  $h$  within unit time interval  $\Delta t$ .

3.3.2. *Evaluation of Social Cost That Is Related to the Treatment of Car Emissions.* The social cost  $C_s$  mainly depends on

the amount of exhaust emissions. Since Li [22] has created the formula of vehicle emission related to the three most important factors CO, HC, and  $NO_x$  after taking real-world experiments, the formulas are applied in this study to evaluate the emission of motor-transport trucks, which are normally heavy diesel vehicles (HDV). The corresponding fitting models are shown in Table 1.

TABLE 1: Fitting models based on comprehensive emission factor of HC, CO, and NO<sub>x</sub> for HDVs.

Pollutant gases	Seasons	Emission factor fitting models (unit: gram/unit.km)
HC	Summer	$E_{HC}(v) = 15.2935 - 0.4129v + 0.0050v^2 - 2 \cdot 10^{-5}v^3$
	Winter	$E_{HC}(v) = 15.2935 - 0.4129v + 0.0050v^2 - 2 \cdot 10^{-5}v^3$
CO	Summer	$E_{CO}(v) = 64.5194 - 2.2960v + 0.0319v^2 - 0.0001v^3$
	Winter	$E_{CO}(v) = 64.5194154 - 2.2960v + 0.0319v^2 - 0.0001v^3$
NO <sub>x</sub>	Summer	$E_{NO_x}(v) = 77.3436 - 1.6314v + 0.0179v^2 - 3 \cdot 10^{-5} \cdot v^3$
	Winter	$E_{NO_x}(v) = 77.3436 - 1.6314v + 0.0179v^2 - 3 \cdot 10^{-5} \cdot v^3$

It is worth mentioning that since HC was excluded from the list of pollutant emissions in “Environmental Air Quality Standards,” only the costs for CO and NO<sub>x</sub> will be discussed in this study. Furthermore, as shown in Table 1, no significant difference is observed between different seasons; in consequence, two universal fitting models, presented in (7) and (8), are applied in this study to evaluate the emission of CO and NO<sub>x</sub>, respectively.

$$E_{CO}(v) = 64.5194 - 2.2960v + 0.0319v^2 - 0.0001v^3 \quad (7)$$

$$E_{NO_x}(v) = 77.3436 - 1.6314v + 0.0179v^2 - 3 \cdot 10^{-5} \cdot v^3 \quad (8)$$

It is obvious that the vehicles' speed is strongly related to the road traffic conditions and the road restrictions; considering that certain regularity can be observed in the road traffic conditions in practice, in this study, the road traffic conditions are imitated with the data collected during an observation period for estimating the expected speed for the routes that connect the pair of nodes in the network.

Afterwards, formulas (3)-(6) are applied to estimate the emission of trucks on those routes by using the corresponding expected speed at the passing time.

Furthermore, since “Technical Guidelines for Environmental Impact Assessment-Atmospheric Environment (HJ2.2-2008)” assessed that the conversion formula for the hourly average and daily average of NO<sub>2</sub> and NO<sub>x</sub> is NO<sub>2</sub> = 0.9 · NO<sub>x</sub>, NO<sub>x</sub> can be replaced by NO<sub>2</sub> for further analysis without significant influence on the experimental results.

As for the treatment fee of the exhaust gases, the environmental degradation costs for CO and NO<sub>2</sub> are considered, where the environmental degradation costs are the corresponding depreciation costs for the environmental degradation because the nature absorbed the polluting gases of the trucks during the distribution process.

According to the regulation “sewage charges standard management approaches” promulgated by the Chinese government, the standard polluting discharge fee is set as 0.6 RMB per pollutant equivalent. The specific pollutant equivalent value is 16.7kg and 0.95kg for CO and NO<sub>2</sub>, respectively. In consequence, unit control cost (UCC) for different pollutant gases can be calculated with the following formulas:

$$\text{Pollutant Equivalents} = \frac{\text{Quality of Emissions Unit (kg)}}{\text{Corresponding Equivalent Value (kg)}} \quad (9)$$

$$\text{Unit Control Cost (UCC)} = \frac{\text{Quality of the pollutant per Emission Unit (kg)}}{\text{Corresponding Equivalent Value (kg)} * \text{Adjustment coefficient}} \quad (10)$$

In consequence, the social cost for one motor-transporting truck running with speed  $v$  (km/h) can be defined with formula (11). It is worth noting that the parameter  $10^{-3}$  is used in formula (11) for the purpose of adjustment because the unit “gram/km” is used in formulas (7) and (8) for calculating the emission, while the unit “kg” is applied in formula (9) for calculating the unit control cost.

$$\begin{aligned} P_e(v) &= UCC(NO_2) \cdot E_{NO_2}(v) + UCC(CO) \cdot E_{CO}(v) \\ &= \frac{0.6 \cdot 1}{0.95 \cdot 1} \cdot E_{NO_2}(v) \cdot 10^{-3} + \frac{0.6 \cdot 1}{16.7 \cdot 1} \cdot E_{CO}(v) \quad (11) \\ &\quad \cdot 10^{-3} \end{aligned}$$

Although the emission is not in linear relationship with the vehicle's speed, it can be calculated as the sum of the

emissions for each segment in which the vehicle traveled during unit time interval and the quality of the results depends on how the unit time interval is defined. In the extreme case, if the unit time interval is small enough, no significant difference will be observed between the sum of those segmental emission values and the exact emission. In consequence, the calculation of emission is simplified by using the segmental processing method with the average expected speed for each segment. The total social cost could be calculated with the following formula:

$$C_s = \sum_{h=1}^H \sum_{i \in \Omega} \sum_{j \in \Omega} d_{ij} \cdot \left[ P_e \left( \frac{1}{2} \bar{v}_{ij} (t_i^h) \right) \right]$$

$$\begin{aligned}
& + \sum_{k=1}^{(k_{ij}^h)^*} (P_e (\bar{V}_{ij} (t_i^h + k \cdot \Delta t))) \\
& + P_e \left( \frac{1}{2} \bar{V}_{ij} (t_i^h + (k_{ij}^h)^* \cdot \Delta t + \varepsilon_{ij} (t_i^h)) \right) \Big] \quad (12)
\end{aligned}$$

3.3.3. *Evaluation of the Cost Related to the Consumption of Fuel*  $C_f$ . The fuel cost  $C_f$  could be obtained by the accumulation of the fuel cost per unit distance, and the calculation process applied by Wen et al. [12] for a Diesel LGV Euro II type vehicle, formula (13), is adopted in this study.

$$\begin{aligned}
EF(v) = & 77.43 + 0.009v - 0.015v^2 + 0.00015v^3 \\
& + 519v^{-1} - 70v^{-2} \quad (13)
\end{aligned}$$

where  $v$  represents the vehicle's travel speed (unit: km/h);  $EF(v)$  represents the fuel consumption (unit: gram/ km). It is worth noting that one liter of diesel weights about 840 grams. Although the price of diesel is strongly related to that of petroleum which keeps changing, this parameter cannot be regarded as sensitive in this model, so a predefined constant works. Therefore, the price of diesel is set as 5.4 RMB/L, the average price in the targeted region during the observation period, and the cost related to the consumption of fuel for one truck running with an average speed  $\bar{V}$  per kilometer (unit: RMB/(vehicle·Km)) can be evaluated with formula (14), where the average speed is calculated by using the segmental processing method based on the average speed for segments in which the vehicle traveled within unit time interval  $\Delta t$ . In consequence, the total fuel cost can be calculated with formula (15), which is quite similar to formula (12), except that  $P_f$  is used instead of  $P_e$ .

$$\begin{aligned}
P_f(\bar{V}) = & 5.4 \cdot \left( 77.43 + 0.009\bar{V} - 0.015\bar{V}^2 \right. \\
& \left. + 0.00015\bar{V}^3 + 519\bar{V}^{-1} - 70\bar{V}^{-2} \right) \div 840 \quad (14)
\end{aligned}$$

$$\begin{aligned}
C_f = & \sum_{h=1}^H \sum_{i \in \Omega} \sum_{j \in \Omega} d_{ij} \cdot \left[ P_f \left( \frac{1}{2} \bar{V}_{ij} (t_i^h) \right) \right. \\
& + \sum_{k=1}^{(k_{ij}^h)^*} (P_f (\bar{V}_{ij} (t_i^h + k \cdot \Delta t))) \\
& \left. + P_f \left( \frac{1}{2} \bar{V}_{ij} (t_i^h + (k_{ij}^h)^* \cdot \Delta t + \varepsilon_{ij} (t_i^h)) \right) \right] \quad (15)
\end{aligned}$$

### 3.4. Mathematical Model

$$\min C_{total} = C_d + C_f + C_s \quad (16)$$

Subject to  
(2), (12), and (15),

$$\sum_{i=1}^H \sum_{i \in \Omega_h} \sum_{j \in \Omega_h} x_{ij}^h(t) \leq H, \quad T_0 \leq t \leq T_f \quad (17)$$

$$\int_{T_0}^{T_f} \sum_{j \in \Omega_0} x_{0j}^h(t) = \int_{T_0}^{T_f} \sum_{i \in \Omega_0} x_{io}^h(t) \quad h \in \{1, \dots, H\} \quad (18)$$

$$\int_{T_0}^{T_f} \sum_{h=1}^H \sum_{i \in \Omega \setminus \{j\}} x_{ij}^h(t) = 1, \quad j \in \Omega_L \quad (19)$$

$$\int_{T_0}^{T_f} \sum_{i \in \Omega_L} \sum_{h=1}^H x_{ij}^h(t) = 1, \quad h \in \{1, \dots, H\} \quad (20)$$

$$x_{ij}^h(t) \in \{0, 1\}, \quad T_0 \leq t \leq T_f \quad (21)$$

The objective is to minimize both transportation cost, including drivers' labor cost and fuel consumption, and treatment cost of pollutant emissions. Constraint (17) indicates that the number of the trucks which leave from the storage depot could not exceed the total number of the trucks owned by storage depot. Constraint (18) denotes that all trucks must start from and end at the only depot. Constraints (19) and (20) indicate that one 4S dealer could only be served by one truck.

## 4. Methodology

Since VRP problems are normally NP-hard, it is quite hard to find their exact solutions efficiently and it is important to make a compromise between the quality of the solution and the efficiency of the approach. Considering that genetic algorithm (GA) has been proven to have great parallelism, robustness, and a strong search capability, it is applied to this study as well.

*4.1. Encoding and Decoding.* In this study, the natural number coding is applied, and each natural number corresponds to a position of the node. In total, three kinds of node are defined: storage depot (starting point/end point), vehicle storage points, and the road junction node the trucks may pass by.

Suppose that there are in total  $R$  trucks in the depot and  $k$  4S dealer shop demanded for service, and  $N$  road junctions are involved in the road map where the storage depot and 4S dealers are connected. Since one truck can be just responsible for one 4S shop,  $k$  trucks should be scheduled.

By using number 0 to separate the route of different trucks, chromosome is arranged as follows:

$$\begin{aligned}
& (0+1, i_1^1+1, i_2^1+1, i_3^1+1, \dots, i_{g_1}^1+1, j_1+1, i_{g_1+1}^1+1, \dots, i_{g_1+h_1}^1+1, \\
& 0+1, 0, 0+1, i_1^2+1, i_2^2+1, i_3^2+1, \dots, i_{g_2}^2+1, j_2+1, i_{g_2+1}^2+1, \dots, i_{g_2+h_2}^2+1, \\
& 0+1, 0, \dots, 0+1, i_1^k+1, i_2^k+1, i_3^k+1, \dots, i_{g_k}^k+1, j_k+1, i_{g_k+1}^k+1, \dots, \\
& i_{g_k+h_k}^k+1, 0+1, 0, t_0^1, t_0^2, t_0^3, t_0^4, \dots, t_0^k)
\end{aligned}$$

Storage depot is noted as "0";  $i_k, k \in \{1, \dots, N\}$  represents the road junctions and 4S dealers passed by the corresponding truck;  $j_l, l \in \{1, \dots, L\}$  represents the 4S dealer

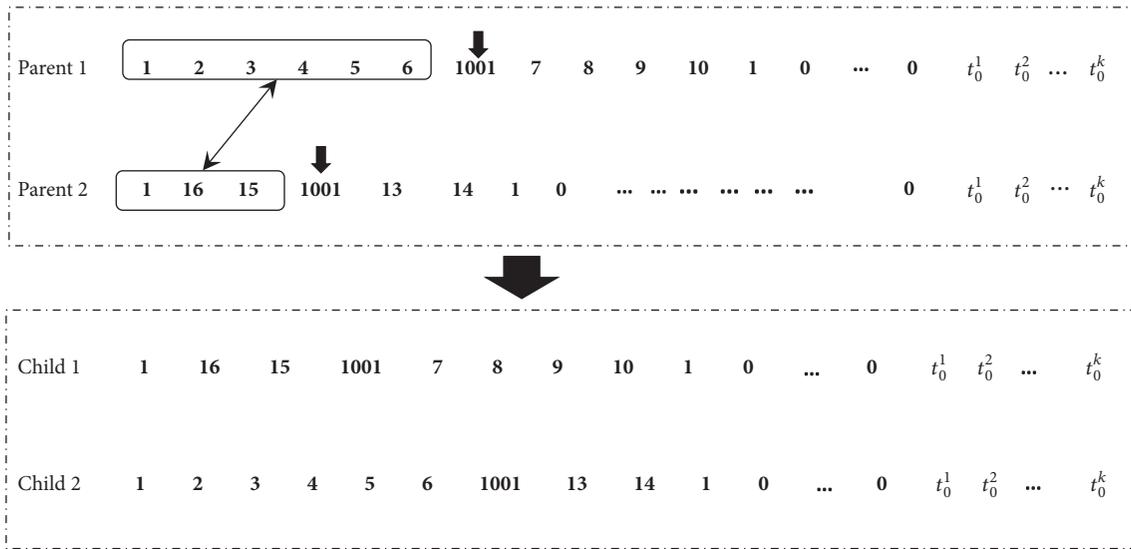


FIGURE 2: Crossover procedure.

served by this truck. In the process of encoding, all path nodes need to plus 1 as the programming language. In fact, this chromosome is composed of  $k+1$  parts: the first  $k$  parts correspond to  $k$  motor-transporting trucks' visiting orders, which are separated by "0", and the last component represents trucks' departure time from the depot.

For example, a possible solution for the problem with 3 motor-transporting trucks serving 3 4S dealers can be as follows: truck 1 is the first one leaving from the storage depot, numbered as node 0, and passes through five route junctions, numbered as node 1 to node 5, to serve the first 4S dealer, numbered as node 1000. Afterwards, this truck will go back to the depot, node 0, passing by four road junctions, nodes 6 to 9. The gene representing the route of this truck can be encoded as  $(0+1, 1+1, 2+1, 3+1, 4+1, 5+1, 1000+1, 6+1, 7+1, 8+1, 9+1, 0+1)=(1, 2, 3, 4, 5, 6, 1001, 7, 8, 9, 10, 1)$ ; truck 2 starts from the storage depot, node 0, and passes through nodes 15 and 17 before serving the 4S dealer, node 1003, and then returns back to the depot by passing through nodes 18 and 16. The gene corresponding to truck 2 can be encoded as  $(0+1, 15+1, 17+1, 1003+1, 18+1, 16+1, 0+1) = (1, 16, 18, 1004, 19, 17, 1)$ . Truck 3 starts also from the depot, serves 4S dealer 1004 after passing through three nodes, 19, 14, and 12, and returns back to the depot by passing through node 13. Similarly, the gene representing the path of truck 3 can be  $(0+1, 19+1, 14+1, 12+1, 1004+1, 13+1, 0+1)=(1, 20, 15, 13, 1005, 14, 1)$ . When the departure time of these trucks is given, the last part of the gene can be defined as  $(t_0^1, t_0^2, t_0^3)$ . Combine all those genes together by inserting three "0"s as boundaries; a chromosome representing the mentioned arrangement is finally defined:  $(1, 2, 3, 4, 5, 6, 1001, 7, 8, 9, 10, 1, 0, 1, 16, 18, 1004, 19, 17, 1, 0, 1, 16, 18, 1004, 19, 17, 1, 0, 1, 20, 15, 13, 1005, 14, 1, 0, t_0^1, t_0^2, t_0^3)$ .

**4.2. Generation of Initial Population.** Consider the fact that if all the paths were created randomly, then the generated routes may not be feasible probably because of the existence of loop

paths. Floyd algorithm is applied to generate a set of shortest paths as the basis of each initial solution so as to improve the quality of the initial population.

Afterwards, randomly generate a number  $k$  between 0 and 1. If  $k$  is smaller than a fixed value set previously, then 2 nodes need to be inserted to the current shortest path, ensuring that the distances between the pair of inserted nodes as well as the distance between the inserted nodes and current nodes should be the shortest ones.

The selected node may be inserted either before or after a 4S dealer. The judgement should be decided on each specific matter. It is worth noting that the random generation procedure may result in infeasible solutions, so a preventive feasibility check is necessary. If the solution is infeasible, then fine-tuning will be performed to make the solution feasible, that is, to make all the routes meet with each other and remove the circles except those that will return back to the depot.

**4.3. Fitness Function and Selection.** In this study, the inverse of the objective value is used as the fitness function, as shown in formula (22), and roulette wheel is applied as the selection scheme.

$$F(s) = \frac{1}{f_s} \tag{22}$$

where  $F(s)$  represents the fitness of individual  $s$  and  $f_s$  is the objective value corresponding to this solution.

**4.4. Crossover and Mutation.** Partially matched crossover applied in this study is to exchange the paths before a selected 4S dealer according to the crossover probability. An example is shown in Figure 2 describing how two children are generated after the paths before 4S dealer 1001 are exchanged.

Mutation takes place according to the predefined mutation probability, and the principle of mutation process is

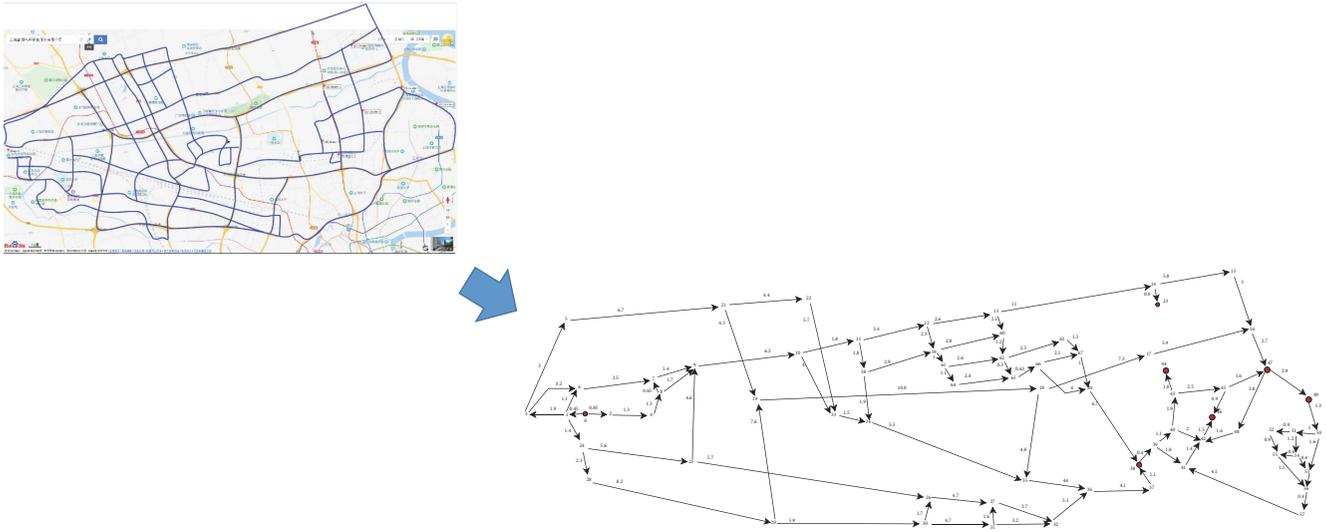


FIGURE 3: Creation of the simplified road map.

to generate a new path to replace a gene of a selected chromosome.

## 5. Case Study

For the purpose of validation of the proposed scheme, the proposed method has been applied with the data collected from district Baoshan at Shanghai, one of the biggest cities in China, and imitates daily delivery of motors from the storage depot of a big vehicle assembly factory located at the northwest of Shanghai to its 4S dealers in the targeted region.

**5.1. Data Preparation.** As shown in Figure 3, a simplified version of the road map of the targeted district is designed as follows: firstly, extract the possible paths, that is, the paths in which the heavy trucks are authorized, on the road map; then, number not only the depot but also the selected 4S dealers and all the involved road junctions, and link the pair of nodes with straight lines for all possible paths. As shown in Figure 4, node 0 represents the storage depot of the targeted vehicle assembly factory, nodes 23, 38, 39, 44, 46, 47, and 49 correspond to seven 4S dealers, which are observed and selected in the targeted region with specific address according to the service catalogue of the targeted vehicle brand, respectively, and the rest of the nodes represent road junctions. The numbers noted on the line between two nodes are the actual distance (unit: km).

The scheduled departure time of trucks is randomly selected between [9, 18], indicating the regular working time of the truck drivers, and we suppose that the trucks set off on the hour.

As for the algorithm, the size of the population is set as 60, so only the best 30 individuals will be retained in each iteration, and the rest of inferior individuals will be replaced by the individuals generated by crossover or mutation operations.

The algorithm is coded with Matlab2013 on a computer with 2.3GHZ processor and 4.0GB memory with Windows 10. The procedure will terminate after 100 iterations. Crossover probability is set as 0.85, and mutation probability is set as 0.1.

### 5.2. Discussion on Results

**5.2.1. Discussion Algorithm Convergence.** In order to test the stability of the proposed GA with 100 iterations, a set of monoobjective variants together with the original problems, the problem aiming at minimizing the composite cost, have been tested.

As shown in Figure 4, the best solution is obtained at around the 50th iteration for minimizing the composite cost; the solution with the minimal fuel cost was found at around the 40th iteration; the solution with the minimal emission cost is obtained at around the 60th iteration and the solution with minimal travel time is already attained around the 20th iteration. To sum up, the proposed GA can be always converged to the best solution within 100 iterations, so the result is quite encouraging.

**5.2.2. Discussion on Solution Quality.** After we compare the solution for minimizing the composite cost (total cost) with other monoobjective solutions, it is observed that all the solutions pick the same distance (distance = 501.10 km), though the values for the other indicators vary a lot, as shown in Table 2. With further analyses, it can be observed that

- (1) the amount of emission is highly correlated with the consumption of fuel, though the “green” mode is not always the solution with minimal fuel consumption,
- (2) The “green” drive mode, that is, the mode where minimizing the emissions needs much more travel time than the composite solution (more than 8%),

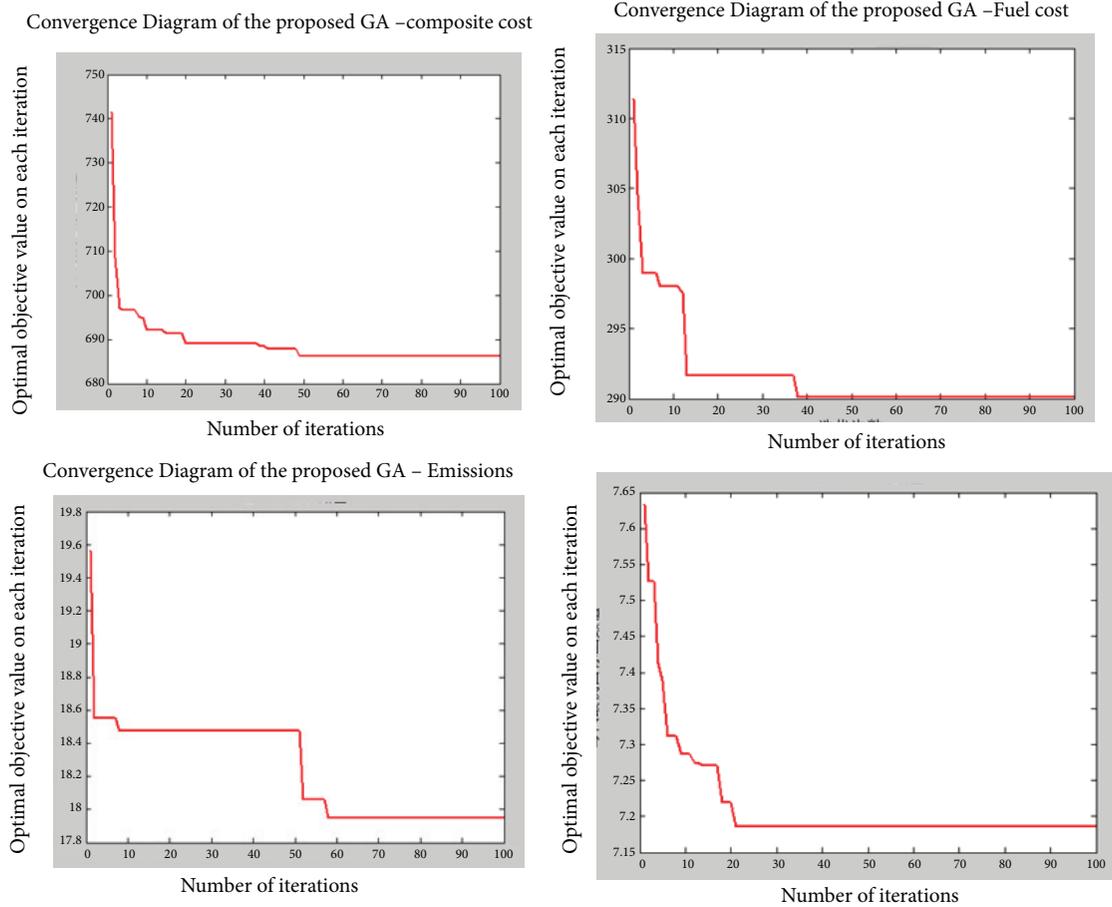


FIGURE 4: Test on algorithm convergence with different objectives.

TABLE 2: Comparison among solutions with different optimization objectives.

Objective	Social cost (RMB)	Fuel cost (RMB)	Drivers' cost (RMB)	distance (km)	Travel time (hour)	Total cost (RMB)
Min total cost	19.49	304.94	361.75	501.10	7.23	686.17
Min travel time	20.25	311.73	359.27	501.10	7.19	691.25
Min $C_f$	18.00	290.12	394.86	501.10	7.90	702.98
Min $C_s$	17.95	291.64	392.21	501.10	7.84	701.79

- (3) The solution with minimal composite cost can dominate the other solution in total cost while not greatly degrading its quality on other objectives.

A further study was organized by comparing the solution obtained by the proposed algorithm with the shortest-path solutions obtained by Floyd algorithm within different time slots. Since road traffic condition is not taken into consideration by Floyd algorithm, the starting time for all trucks remains the same, that is, the trucks set off on hour from 9:00 to 17:00 to evaluate the cost within different time slots, as shown in Table 3.

Obviously, all costs vary during different time slots, indicating that the traffic condition can greatly influence

the emissions, consumption of fuels, and travel time, which shows the importance of our work.

While comparing those solutions with the solution obtained by the proposed method to minimize the composite cost, it can be observed that our solution dominates all of the shortest-path solutions!

## 6. Conclusions

With the rapid development of the logistics industry, the business managers have realized that effective reduction of logistics cost has direct relationship with the enterprises' economic interests and the improvement of competitive advantage by taking their social responsibilities. This study aims at optimizing the performance of motor-transporting

TABLE 3: Shortest-path solutions obtained by Floyd algorithms within various time slots.

Departure time	Optimal distance (km)	Total cost (RMB)	Social cost (RMB)	Drivers' cost (RMB)	Fuel cost (RMB)	Total time (hour)
9:00	501.1	756.01	21.94	390.74	343.33	7.81
10:00	501.1	745.43	21.95	384.75	338.73	7.69
11:00	501.1	795.80	23.51	401.21	371.08	8.02
12:00	501.1	691.96	20.20	360.28	311.48	7.20
13:00	501.1	693.55	19.71	366.61	307.24	7.33
14:00	501.1	752.59	21.52	391.41	339.66	7.82
15:00	501.1	792.67	22.27	412.20	358.20	8.24
16:00	501.1	799.39	21.87	423.85	353.67	8.47
17:00	501.1	791.44	20.11	434.68	336.64	8.69
Average value	501.1	757.65	21.45	396.19	340.00	7.92

services by taking into account dynamic traffic conditions. Based on a discretization of the distribution of the road traffic condition, the influence of the road congestion on vehicle speeds has been taken into account so as to create optimal delivery routes for motor-transporting trucks with a composite objective of minimizing both economic cost (driver cost and fuel cost) and social cost (vehicle emissions). A genetic algorithm, one of the most popular metaheuristics, has been proposed to solve this problem.

With the data collected from the Baoshan district at Shanghai, one of the biggest cities in China, numerical experiments showed that the proposed algorithm can always converge to the best solution within 100 iterations. Considering the objective values, the solution with composite objective can dominate the others in general, while no significant deterioration is observed for other objectives. The results are quite encouraging.

## 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.

## References

- [1] M. Gendreau, G. Ghiani, and E. Guerriero, "Time-dependent routing problems: A review," *Computers & Operations Research*, vol. 64, pp. 189–197, 2015.
- [2] A. Franceschetti, D. Honhon, T. Van Woensel, T. Bektaş, and G. Laporte, "The time-dependent pollution-routing problem," *Transportation Research Part B: Methodological*, vol. 56, pp. 265–293, 2013.
- [3] M. Setak, M. Habibi, H. Karimi, and M. Abedzadeh, "A time-dependent vehicle routing problem in multigraph with FIFO property," *Journal of Manufacturing Systems*, vol. 35, pp. 37–45, 2015.
- [4] Y. Huang, L. Zhao, T. Van Woensel, and J.-P. Gross, "Time-dependent vehicle routing problem with path flexibility," *Transportation Research Part B: Methodological*, vol. 95, pp. 169–195, 2017.
- [5] S. Zheng, C. Jiandong, L. Xiaomin, and L. Keqiang, "Urban pickup and delivery problem considering time-dependent fuzzy velocity," *Computers & Industrial Engineering*, vol. 60, no. 4, pp. 821–829, 2011.
- [6] T.-Y. Liao and T.-Y. Hu, "An object-oriented evaluation framework for dynamic vehicle routing problems under real-time information," *Expert Systems with Applications*, vol. 38, no. 10, pp. 12548–12558, 2011.
- [7] A. L. Kok, E. W. Hans, and J. M. J. Schutten, "Vehicle routing under time-dependent travel times: the impact of congestion avoidance," *Computers & Operations Research*, vol. 39, no. 5, pp. 910–918, 2012.
- [8] S. Kritzing, K. F. Doerner, R. F. Hartl, G. Kiechle, H. Stadler, and S. S. Manohar, "Using Traffic Information for Time-Dependent Vehicle Routing," *Procedia - Social and Behavioral Sciences*, vol. 39, pp. 217–229, 2012.
- [9] J. F. Ehmke, A. Steinert, and D. C. Mattfeld, "Advanced routing for city logistics service providers based on time-dependent travel times," *Journal of Computational Science*, vol. 3, no. 4, pp. 193–205, 2012.
- [10] C. Verbeeck, K. Sörensen, E.-H. Aghezzaf, and P. Vansteenwegen, "A fast solution method for the time-dependent orienteering problem," *European Journal of Operational Research*, vol. 236, no. 2, pp. 419–432, 2014.
- [11] S. Mancini, "Time dependent travel speed vehicle routing and scheduling on a real road network: The case of torino," *Transportation Research Procedia*, vol. 3, pp. 433–441, 2014.
- [12] L. Wen, B. Çatay, and R. Eglese, "Finding a minimum cost path between a pair of nodes in a time-varying road network with a congestion charge," *European Journal of Operational Research*, vol. 236, no. 3, pp. 915–923, 2014.
- [13] D. Taş, N. Dellaert, T. van Woensel, and T. de Kok, "The time-dependent vehicle routing problem with soft time windows and stochastic travel times," *Transportation Research Part C: Emerging Technologies*, vol. 48, pp. 66–83, 2014.
- [14] L. Wen and R. Eglese, "Minimum cost VRP with time-dependent speed data and congestion charge," *Computers & Operations Research*, vol. 56, pp. 41–50, 2015.

- [15] A. Franceschetti, E. Demir, D. Honhon, T. Van Woensel, G. Laporte, and M. Stobbe, "A metaheuristic for the time-dependent pollution-routing problem," *European Journal of Operational Research*, vol. 259, no. 3, pp. 972–991, 2017.
- [16] J. F. Ehmke, A. M. Campbell, and B. W. Thomas, "Vehicle routing to minimize time-dependent emissions in urban areas," *European Journal of Operational Research*, vol. 251, no. 2, pp. 478–494, 2016.
- [17] M. Naderipour and M. Alinaghian, "Measurement, evaluation and minimization of CO<sub>2</sub>, NO<sub>x</sub>, and CO emissions in the open time dependent vehicle routing problem," *Measurement*, vol. 90, pp. 443–452, 2016.
- [18] M. Behnke, T. Kirschstein, and C. Bierwirth, "An emission-minimizing vehicle routing problem with heterogeneous vehicles and pathway selection," in *Operations Research Proceedings 2016*, Oper. Res. Proc., pp. 285–291, Springer, Cham, 2018.
- [19] F. Köster, M. W. Ulmer, D. C. Mattfeld, and G. Hasle, "Anticipating emission-sensitive traffic management strategies for dynamic delivery routing," *Transportation Research Part D: Transport and Environment*, vol. 62, pp. 345–361, 2018.
- [20] M. Soysal, J. M. Bloemhof-Ruwaard, and T. Bektaş, "The time-dependent two-echelon capacitated vehicle routing problem with environmental considerations," *International Journal of Production Economics*, vol. 164, pp. 366–378, 2015.
- [21] N. Norouzi, M. Sadegh-Amalnick, and R. Tavakkoli-Moghaddam, "Modified particle swarm optimization in a time-dependent vehicle routing problem: minimizing fuel consumption," *Optimization Letters*, vol. 11, no. 1, pp. 121–134, 2017.
- [22] T. Z. Li, *Study on evaluation and forecast technology of urban traffic atmospheric environment influence [PhD. thesis]*, Southeast University, Nanjing, China, 2001.

## Research Article

# Optimal Scheme for Process Quality and Cost Control by Integrating a Continuous Sampling Plan and the Process Yield Index

Chunzhi Li, Shurong Tong , and Keqin Wang

School of Management, Northwestern Polytechnical University, No. 1, Dongxiang Road, Chang'an District, Xi'an 710129, China

Correspondence should be addressed to Shurong Tong; stong@nwpu.edu.cn

Received 28 April 2018; Revised 2 July 2018; Accepted 30 October 2018; Published 19 November 2018

Academic Editor: Lu Zhen

Copyright © 2018 Chunzhi 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.

The single level continuous sampling plan (CSP-1) is an in-line process control tool that has been commonly adopted in various manufacturing industries. However, CSP-1 is designed for only satisfying the quality constraint. At the same time, CSP-1 has disadvantages with the high probabilities of both Type I and Type II errors due to its inherent deficiency coming from the operating procedure. In this work, an optimal scheme for process quality and cost control is proposed to monitor the process cost and improve the process quality. The CSP-1 and the process yield index ( $S_{pk}$ ) are integrated in the present scheme, which work independently and complementarily. The four parameters (clearance number, inspecting fraction, sample size, and critical value) are designed in the proposed scheme under simultaneously considering the quality and cost constraints. The sole feasible inspection scheme in CSP-1 under the two constraints is found and used for controlling the process quality. The probabilities of Type I and Type II errors are concurrently controlled at the stipulated level with the risk control scheme, which is constructed with two nonlinear inequation based on the accurate distribution of the index  $S_{pk}$ . A case study is illustrated to validate the effectiveness and practicality of the proposed scheme.

## 1. Introduction

With the advances in manufacturing and control technology, process quality can be constant at required level for the in-control process. Multiple constraints, such as quality, cost, risk, and environmental adaptability, are imposed on the process control. The constraints are originated from the intensifying market competition and the demand of sustainable development. Many kinds of process control tools are developed to monitor and improve the process quality under various constraints for the in-control process. As far as we know, the process control tool is designed for satisfying only one control constraint. For example, the continuous sampling plan (CSP) is commonly adopted in process control; various CSPs are presented for meeting only the quality constraint. The process yield index ( $S_{pk}$ ) is proposed to only express the process capability. No process control tool is developed for solving the optimal problem of multiple constraints.

Various CSPs are developed for satisfying the quality requirement and used to improve and control the process

quality. The current literature on CSP can be generally classified into three categories. Various CSPs with different inspecting procedures were presented in the first kind of literature. The single level CSP was designated as CSP-1 and was widely adopted for process quality control in manufacturing [1]. To meet different demands in process control, some reduced CSPs were presented with the objective of decreasing the number of units inspected when the probability of nonconformity for the process was very low [2–5]. Some tightened CSPs have been designed to guarantee that the outgoing quality meets stringent quality requirements stipulated by the customer [6–8]. Nevertheless, CSPs designed under quality constraint have not taken other constraint, such as cost and risk constraints, into account. The category II literature has investigated the influence of the change of the inspection scheme in CSPs on the inspection cost. From a cost perspective, it was demonstrated that implementing a CSP for a stable production process was inappropriate [9]. Considering the economic objective, the inspection scheme in CSP-1 could work most effectively when the probability

of nonconformity was roughly two-thirds of the value of the average outgoing quality limit (AOQL) [10]. Moreover, the different methods have been proposed for optimizing the parameters of inspection schemes in CSP-1 in a view of cost [11–14]. However, there is no agreed conclusion about how to identify the inspection scheme in CSPs whose inspection cost is the minimum. The third kind of literature has proposed the integrated control scheme between CSP and other process control tools, such as preventive maintenance and specification limit [15–21]. The integrated schemes were designed with the objective of minimizing the cost function. Nevertheless, no integrating process control scheme is presented for simultaneously meeting the quality and cost constraints by combining the CSP and other process control tools.

The in-control process is commonly regarded as stochastic process. There are two kinds of risk when process control tools are adopted in process control. One risk is the probability of making type I error that an in-control process with good quality is rejected. The other is the probability of making type II error that an in-control process with bad quality is accepted. CSPs have an inherent deficiency that the two risks are high [22–25]. The process yield index  $S_{pk}$  is an effective performance measure for reflecting the influence of the machining centre drift and process deviation fluctuation on the probability of nonconformity. The index  $S_{pk}$  had a one-to-one relationship with the probability of conformity and nonconformity [26]. It has been demonstrated that the natural estimator of  $\hat{S}_{pk}$  was asymptotically normally distributed [27]. The accuracy of the natural estimator has been investigated with a simulation technique [28]. The process yields for some specified cases, such as the imprecise sample data, circular profiles, autocorrelation between linear profiles, and multiple stream processes, have been analyzed in some literature [29–32]. In the above researches, the index  $S_{pk}$  was used to only reflect the process capability. In recent years, some variable sampling plans based on the process capability indices have been proposed with the objective of building the determination rules for the acceptance or rejection of product lots [33–36]. Two risks are simultaneously taken into consideration in the proposed variable sampling plans by utilizing the inference property of the natural estimator  $\hat{S}_{pk}$ . Nevertheless, no literature has been devoted to constructing the risk control strategy with the index  $S_{pk}$  under quality and cost constraints for in-line process control.

Distinguished from the objective of minimizing cost in the sampling inspection for lot acceptance decision [34], the in-line process control aims to reach the control goal of minimizing total cost. The total cost constraint restricts the in-constant process to keep constant at right capability level. Thus, the reasonable inspection cost is permitted in process control, which is called cost constraint. The integrated process control scheme needs to be presented with simultaneously considering the quality constraint and the inspection cost constraint. The inspection cost constraint can be translated into the maximum affordable inspected fraction for the in-line process control. In this work, an optimal scheme for process quality and cost control by combining CSP-1 and the index  $S_{pk}$  is presented. CSP-1 is adopted in the integrated scheme due to its simplicity and practicability in operation.

In the proposed plan, there are four parameters (the clearance number, the sampling fraction, the sampling size, and the critical value of the index  $S_{pk}$ ) under the quality and cost constraints. The two parameters (the clearance number and the sampling fraction) in CSP-1 are utilized to guarantee that the average outgoing quality is conforming for the in-control process. The index  $S_{pk}$  is used to construct the risk control scheme. The two risks are concurrently controlled at the stipulated level by constructing two nonlinear inequations with the accurate distribution of the estimator  $\hat{S}_{pk}$ .

The rest of this paper is organized as follows. In Section 2, the concept and estimator of the index  $S_{pk}$  are introduced. In Section 3, the optimal scheme for process quality and cost control and the operating procedure are presented. The method of identifying the plan parameters is provided. In Section 4, the values of the parameters of the optimal scheme for three different quality constraints are tabulated for practical purposes. Comparisons of the operating characteristic (OC) curves between CSP-1 and the optimal scheme are given to present the advantages of the integrated scheme. In Section 5, an example of application is provided to validate the effectiveness and practicality of the integrated control plan. Finally, Section 6 concludes the paper.

## 2. Process Fraction Nonconforming and Process Yield Index

The process fraction nonconforming,  $p$ , is a crucial performance measure for in-line process quality control. For a process that is well controlled, the value of  $p$  can be taken as a constant. However, the index of  $p$  is an unknown variable and needs to be estimated. Let  $F(\cdot)$  be the cumulative distribution function (CDF) of the quality characteristic interested, so  $p = 1 - [F(USL) - F(LSL)]$  for the in-control process with a two-sided specification limits, where  $USL$  is the upper specification limit and  $LSL$  is the lower specification limit. If the quality characteristic follows a normal distribution, we get

$$p = 1 - \left\{ \Phi \left( \frac{USL - \mu}{\sigma} \right) - \Phi \left( \frac{LSL - \mu}{\sigma} \right) \right\} \quad (1)$$

where  $\mu$  and  $\sigma$  are the process mean and the process standard deviation, respectively, and  $\Phi(\cdot)$  is the CDF of the standard normal distribution,  $N(0, 1)$ . Unfortunately, there is no literature devoted to a study of the distribution properties of the index of  $p$ .

Boyles [13] proposed the use of the process yield index,  $S_{pk}$ , to obtain an exact measure of the process yield for a process with a normal distribution. There is a one-to-one relationship between  $S_{pk}$  and the process yield. The proposed index  $S_{pk}$  is defined as

$$S_{pk} = \frac{1}{3} \Phi^{-1} \left\{ \frac{1}{2} \Phi \left( \frac{USL - \mu}{\sigma} \right) + \frac{1}{2} \Phi \left( \frac{\mu - LSL}{\sigma} \right) \right\} \quad (2)$$

where  $\Phi^{-1}(\cdot)$  is the inverse function of the CDF,  $\Phi(\cdot)$ , of the standard normal distribution. Let  $C_p = (USL - LSL)/(6\sigma)$  and  $C_a = 1 - |\mu - M|/d$ , where  $M$  is the middle point of the whole tolerance range,  $M = (USL + LSL)/2$ , and  $d$  is half the

tolerance range,  $d = (USL - LSL)/2$ . Equation (2) can also be expressed as

$$S_{pk} = \frac{1}{3} \Phi^{-1} \left\{ \frac{1}{2} \Phi(3C_p C_a) + \frac{1}{2} \Phi(3C_p(2 - C_a)) \right\} \quad (3)$$

The formula for the relationship between  $p$  and  $S_{pk}$  can be obtained from (1) and (2):

$$p = 2 - 2\Phi(3S_{pk}), \quad S_{pk} > 0 \quad (4)$$

Table 1 shows the one-to-one correspondence between process yield, process fraction nonconforming, and process yield index.

The process mean  $\mu$  and the process standard deviation  $\sigma$  are usually unknown variables and need to be estimated using the sample mean  $\bar{x}$  and the sample standard deviation  $s$ , where  $\bar{x} = (1/n) \sum_{i=1}^n x_i$  and  $s = \sqrt{(1/(n-1)) \sum_{i=1}^n (x_i - \bar{x})^2}$ . Thus, the estimator of  $S_{pk}$  can be written as follows:

$$\begin{aligned} \hat{S}_{pk} &= \frac{1}{3} \Phi^{-1} \left[ \frac{1}{2} \Phi \left( \frac{USL - \bar{x}}{s} \right) + \frac{1}{2} \Phi \left( \frac{\bar{x} - LSL}{s} \right) \right] \\ &= \frac{1}{3} \Phi^{-1} \left\{ \frac{1}{2} \Phi(3\hat{C}_p \hat{C}_a) + \frac{1}{2} \Phi(3\hat{C}_p(2 - \hat{C}_a)) \right\} \end{aligned} \quad (5)$$

The estimator  $\hat{S}_{pk}$  is such a complicated function that it is impossible to obtain its exact cumulative distribution function and probability density function. Lee et al. [14] furnished a useful approximation to the distribution of  $\hat{S}_{pk}$  under a normal distribution as

$$\hat{S}_{pk} \approx S_{pk} + \frac{1}{6\sqrt{n}} \frac{W}{\phi(3S_{pk})} \quad (6)$$

where  $\phi$  is the probability density function of the standard normal distribution  $N(0, 1)$  and

$W$

$$= \begin{cases} \frac{\sqrt{n}}{2} \left[ \frac{a(s^2 - \sigma^2)}{\sigma} \right] - \frac{\sqrt{n}b(\bar{x} - \mu)}{\sigma} & \text{for } \mu < M, \\ \frac{\sqrt{n}}{2} \left[ \frac{a(s^2 - \sigma^2)}{\sigma} \right] + \frac{\sqrt{n}b(\bar{x} - \mu)}{\sigma} & \text{for } \mu > M, \end{cases} \quad (7)$$

where  $a$  and  $b$  are defined as functions of  $\mu$  and  $\sigma$  (or  $C_p$  and  $C_a$ ):

$$\begin{aligned} a &= \frac{1}{\sqrt{2}} \left\{ \frac{USL - \mu}{\sigma} \phi \left( \frac{USL - \mu}{\sigma} \right) \right. \\ &\quad \left. + \frac{\mu - LSL}{\sigma} \phi \left( \frac{\mu - LSL}{\sigma} \right) \right\} \\ &= \frac{1}{\sqrt{2}} \left\{ 3C_p(2 - C_a) \phi(3C_p(2 - C_a)) \right. \\ &\quad \left. + 3C_p C_a \phi(3C_p C_a) \right\} \end{aligned}$$

TABLE 1: The process yield, process fraction nonconforming, and corresponding value of the process yield index  $S_{pk}$ .

Yield (%)	fraction nonconforming (%)	$S_{pk}$
99.3066052	0.6933948	0.90
99.5628077	0.4371923	0.95
99.7300204	0.2699796	1.00
99.8367295	0.1632705	1.05
99.9033152	0.0966848	1.10

$$\begin{aligned} b &= \phi \left( \frac{USL - \mu}{\sigma} \right) - \phi \left( \frac{\mu - LSL}{\sigma} \right) = \phi[3C_p(2 - C_a)] \\ &\quad - \phi(3C_p C_a) \end{aligned} \quad (8)$$

The estimator  $\hat{S}_{pk}$  approximately follows a normal distribution  $N(S_{pk}, (a^2 + b^2)/36n(\phi(3S_{pk}))^2)$ . Thus, the probability density function of  $\hat{S}_{pk}$  can be obtained as

$$\begin{aligned} f_{\hat{S}_{pk}}(x) &= \sqrt{\frac{18n}{\pi}} \\ &\quad \cdot \frac{\phi(3S_{pk})}{\sqrt{a^2 + b^2}} \exp \left[ -\frac{18n(\phi(3S_{pk}))^2}{a^2 + b^2} (x - S_{pk})^2 \right], \quad (9) \\ &\quad -\infty < x < +\infty \end{aligned}$$

### 3. The Optimal Scheme for Process Quality and Cost Control

**3.1. Quality and Cost Constraints.** Two constraints, the average outgoing quality limit (AOQL) and the inspection capability limit ( $AFI_L$ ), are predetermined by the practitioner.  $AFI$  is the average fraction inspected.  $AFI_L$  is the cost constraint and represents the maximum affordable inspection workload. The value of  $AFI_L$  is converted from the planned inspection cost. The value of AOQL is generally stipulated by the product designer or the customer. For the in-control process, the proportion of nonconformance  $p$  can be taken as constant. Obviously, there exists a specified in-control process, named limit quality process with quality limit  $p_{IQL}$ , for which conforming outgoing quality can be achieved only under the inspection capability limit  $AFI_L$ . This means that  $AFI_L = 1 - AOQL/p_{IQL}$  for the limit quality process with the quality limit  $p_{IQL}$ . It needs to be noted that generally  $p_{IQL} > AOQL$ .

**3.2. Determination of Parameters for the Quality Control Scheme in CSP-1.** The quality control scheme in CSP-1 should guarantee that the two constraints, AOQL and  $AFI_L$ , are simultaneously satisfied when  $p$  ranges from 0 to  $p_{IQL}$ . According to the performance formulas in CSP-1 proved by

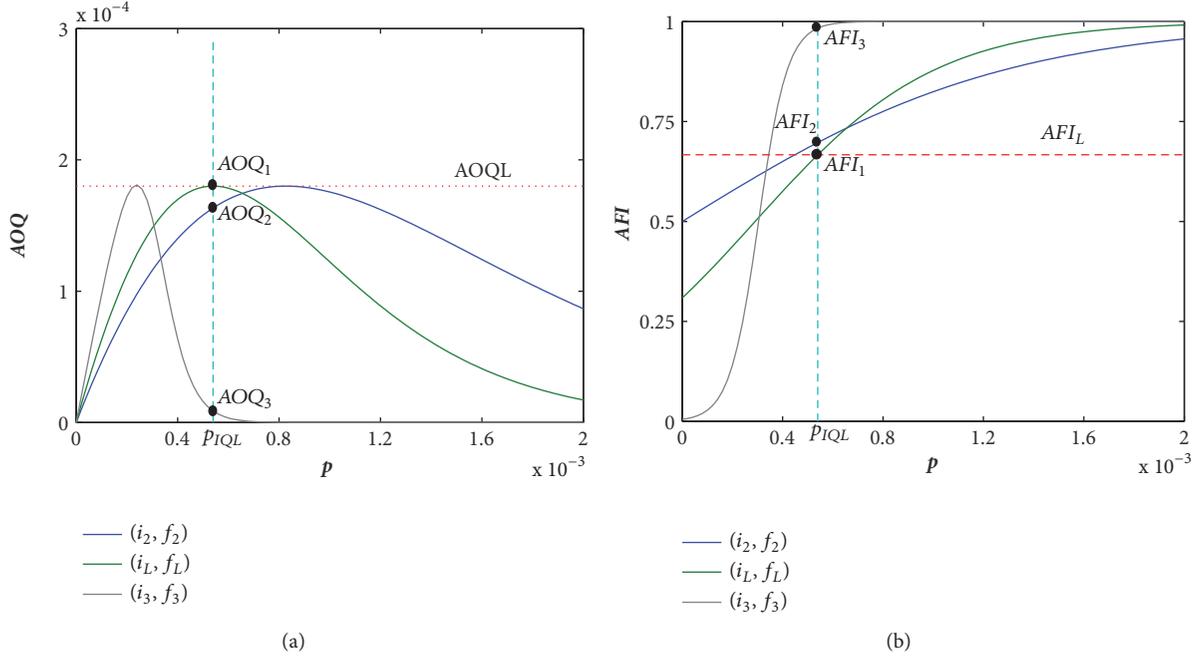


FIGURE 1: Comparisons of the curves of (a) AOQ and (b) AFI between  $(i_L, f_L)$  and other two types of the AOQL contour schemes.

Yang [37], the set of inequalities can be constructed as follows for the quality and cost constraints, respectively,

$$\frac{(1-f)pq^i}{(f+(1-f)q^i)} \leq \text{AOQL} \quad (10)$$

$$\frac{f}{(f+(1-f)q^i)} \leq \text{AFI}_L \quad (11)$$

where  $p \leq p_{IQL}$ ,  $q = 1 - p$ ,  $0 < f < 1$ ,  $i$  is the clearance number and only takes positive integer, and  $f$  is the inspection fraction.

Obviously, all AOQL contour schemes can meet inequality (10) when  $p$  ranges from 0 to 1. Figure 1 shows the curves of the performance measures AOQ and AFI for all AOQL contour schemes under the given constraints AOQL and  $\text{AFI}_L$ , and AOQ is the average outgoing quality. The inspection schemes  $(i_2, f_2)$  represent the type of schemes with  $p_L > p_{IQL}$ ,  $(i_3, f_3)$  with  $p_L > p_{IQL}$ ,  $(i_L, f_L)$  with  $p_L = p_{IQL}$ ,  $p_L$  is the point where  $\max(\text{AOQ}) = \text{AOQL}$  occurs. There are infinite schemes, respectively, included in the two types of the inspection schemes  $(i_2, f_2)$  and  $(i_3, f_3)$ . The inspection scheme  $(i_L, f_L)$  with  $p_L = p_{IQL}$  has only one included.

It can be seen from Figure 1(b) that only the inspection scheme  $(i_L, f_L)$  can meet the inequality (11) when  $p \leq p_{IQL}$ . Thus, the inspection scheme  $(i_L, f_L)$  is the sole scheme which can simultaneously satisfy the two inequalities (10) and (11), named the optimal scheme.

It can be concluded from Figure 1 that the value of AOQ increases gradually to the quality constraint AOQL and the inspection workload also increases gradually to the inspection capability limit  $\text{AFI}_L$  when the process quality is close to the quality limit  $p_{IQL}$  for the optimal scheme  $(i_L, f_L)$ .

The scheme  $(i_L, f_L)$  can achieve the conforming outgoing quality  $\text{AOQ} < \text{AOQL}$  with a smaller inspection workload than  $\text{AFI}_L$  for an in-control process with  $p < p_{IQL}$  and  $\text{AOQ} = \text{AOQL}$  under the inspection capability limit  $\text{AFI}_L$  for an in-control process with  $p = p_{IQL}$ .

The values of the parameters  $i_L$  and  $f_L$  cannot be easily achieved with inequalities (10) and (11). Li et al. [38] proposed the formulas to solve the parameters of the specified AOQL contour scheme for the specified  $p$ . Thus, the parameters  $i_L$  and  $f_L$  can be obtained as follows:

$$i_L = \frac{q_{IQL}}{p_{IQL} - \text{AOQL}} \quad (12)$$

$$f_L = \frac{q_{IQL}^i}{q_{IQL}^i + p_{IQL}i q_{IQL}^{-1} - 1} \quad (13)$$

where  $q_{IQL} = 1 - p_{IQL}$ .

Table 2 shows various inspection schemes  $(i_L, f_L)$  under three quality constraints and various cost constraints. The values of  $i_L$  decrease and the values of  $f_L$  increase when the values of the quality limit  $p_{IQL}$  become greater under the given quality requirement AOQL.

In addition, it can be observed from Figure 1(b) that, for the process with  $p > p_{IQL}$ , the minimum inspection workload demanded for meeting the quality constraint exceeds the inspection capability limit  $\text{AFI}_L$ ; the production should be stopped. Nevertheless, both the probabilities of making the type I and II errors are high when the process is controlled with the AOQL contour scheme in CSP-1. The risk control scheme for controlling concurrently the two risks in the optimal scheme should be redesigned under the quality and cost constraints.

TABLE 2: The quality control scheme  $(i_L, f_L)$  in the optimal scheme under specified AOQL and  $AFI_L(p_{iQL})$ .

$AFI_L$	AOQL = 0.00018			AOQL = 0.00143			AOQL = 0.0122		
	$p_{iQL}$	$i_L$	$f_L$	$p_{iQL}$	$i_L$	$f_L$	$p_{iQL}$	$i_L$	$f_L$
0.5000	0.00036	5554	0.1192	0.00286	697	0.1195	0.0244	80	0.1218
0.6667	0.00054	2776	0.3086	0.00429	348	0.3092	0.0366	39	0.3145
0.7500	0.00072	1851	0.4417	0.00572	232	0.4425	0.0488	26	0.4498
0.8000	0.00090	1388	0.5342	0.00715	174	0.5351	0.0610	19	0.5437
0.8333	0.00108	1110	0.6011	0.00858	139	0.6022	0.0732	15	0.6117
0.8571	0.00126	925	0.6515	0.01001	115	0.6527	0.0854	12	0.6629
0.8750	0.00144	793	0.6908	0.01144	99	0.6920	0.0976	11	0.7028
0.8889	0.00162	693	0.7222	0.01287	86	0.7235	0.1098	9	0.7347

### 3.3. Determination of Parameters for the Risk Control Scheme.

The in-control process with  $p \leq \text{AOQL}$  should be accepted with a higher probability than  $1 - \alpha$ ;  $\alpha$  is the probability of making the type I error that the process with high quality level is rejected. The in-control process with  $p > p_{IQL}$  should be rejected with a higher probability than  $1 - \beta$ ;  $\beta$  is the probability of making the type II error that the process with low quality level is accepted. However, the two types of risks cannot be simultaneously controlled with the AOQL contour schemes. Thus, the risk control scheme based on the index  $S_{pk}$  is designed under the quality and cost constraints.

There are two key points,  $(\text{AOQL}, \alpha)$  and  $(p_{IQL}, \beta)$ , that need to be considered at the same time in the risk control scheme. The OC curve should pass through the two designated points to meet the two constraints AOQL and  $AFI_L$ . Let  $S_{\text{AOQL}}$  be the value of the process yield index corresponding to the quality level  $p = \text{AOQL}$  and let  $S_{IQL}$  be the value of the process yield index corresponding to the quality level  $p = p_{IQL}$ . Therefore, the two key points  $(\text{AOQL}, \alpha)$  and  $(p_{IQL}, \beta)$  for the OC function can be designated as  $(S_{\text{AOQL}}, \alpha)$  and  $(S_{IQL}, \beta)$ . It means that if the estimator of  $\hat{S}_{pk}$  for the in-control process is greater than the given value of  $S_{\text{AOQL}}$ , the probability of accepting the in-control process will be greater than  $1 - \alpha$ , and if the estimator of  $\hat{S}_{pk}$  for the in-control process is lower than the fixed value of  $S_{IQL}$ , the probability of accepting the in-control process will be less than the given value of  $\beta$ .

For the quality characteristic following a normal distribution and having a lower specification limit  $LSL$  and upper specification limit  $USL$ , the OC function with  $S_{pk} = S'$  can be given as

$$P(\hat{S}_{pk} \geq s_0 | S_{pk} = S') = \int_{s_0}^{\infty} \sqrt{\frac{18n}{\pi}} \cdot \frac{\phi(3S')}{\sqrt{a^2 + b^2}} \exp\left[-\frac{18n(\phi(3S'))^2}{a^2 + b^2}\right] \times (x - S')^2 dx, \quad -\infty < x < +\infty \quad (14)$$

Thus, the two nonlinear inequalities specified by the two risks can be obtained as follows:

$$P_1(\hat{S}_{pk} \geq s_0 | S_{pk} = S_{\text{AOQL}}) = \int_{s_0}^{\infty} \sqrt{\frac{18n}{\pi}} \cdot \frac{\phi(3S_{\text{AOQL}})}{\sqrt{a_{\text{AOQL}}^2 + b_{\text{AOQL}}^2}} \exp\left[-\frac{18n(\phi(3S_{\text{AOQL}}))^2}{a_{\text{AOQL}}^2 + b_{\text{AOQL}}^2}\right] \times (x - S_{\text{AOQL}})^2 dx \geq 1 - \alpha \quad (15)$$

$$P_2(\hat{S}_{pk} \geq s_0 | S_{pk} = S_{IQL}) = \int_{s_0}^{\infty} \sqrt{\frac{18n}{\pi}} \cdot \frac{\phi(3S_{IQL})}{\sqrt{a_{IQL}^2 + b_{IQL}^2}} \exp\left[-\frac{18n(\phi(3S_{IQL}))^2}{a_{IQL}^2 + b_{IQL}^2}\right] \times (x - S_{IQL})^2 dx \leq \beta \quad (16)$$

The two parameters, the critical value  $s_0$  and the sample size  $n$ , must simultaneously satisfy the two nonlinear inequalities (15) and (16). There are infinite values of the combination  $(n, s_0)$  which can meet inequalities (15) and (16). The specified value of  $(n, s_0)$  which can simultaneously satisfy (17) and (18) is the boundary of all the feasible combinations  $(n, s_0)$ . It means that the combination of  $(n, s_0)$  meeting (17) and (18) can satisfy inequalities (15) and (16).

$$\int_{s_0}^{\infty} \sqrt{\frac{18n}{\pi}} \cdot \frac{\phi(3S_{\text{AOQL}})}{\sqrt{a_{\text{AOQL}}^2 + b_{\text{AOQL}}^2}} \exp\left[-\frac{18n(\phi(3S_{\text{AOQL}}))^2}{a_{\text{AOQL}}^2 + b_{\text{AOQL}}^2}\right] \times (x - S_{\text{AOQL}})^2 dx = 1 - \alpha \quad (17)$$

$$\int_{s_0}^{\infty} \sqrt{\frac{18n}{\pi}} \cdot \frac{\phi(3S_{IQL})}{\sqrt{a_{IQL}^2 + b_{IQL}^2}} \exp\left[-\frac{18n(\phi(3S_{IQL}))^2}{a_{IQL}^2 + b_{IQL}^2}\right] \times (x - S_{IQL})^2 dx = \beta \quad (18)$$

The solution  $(n, s_0)$  for (17) and (18) is sole. For example, under two constraints,  $\text{AOQL} = 0.00018 (S_{\text{AOQL}} = 1.2485)$  and  $p_{IQL} = 0.0009 (S_{IQL} = 1.1067)$ , and two given risks,  $\alpha = 0.1$  and  $\beta = 0.1$ , the sole solution  $(n, s_0) = (225, 1.1730)$  can be obtained from (17) and (18). It implies that the value of the estimator  $\hat{S}_{pk}$  is calculated with the 225 sample data. If  $\hat{S}_{pk} \geq 1.1730$ , the process is judged to be controllable with the current inspection scheme  $(i_L, f_L)$  under the quality and cost constraints, and the current inspection scheme  $(i_L, f_L)$  will continue. If  $\hat{S}_{pk} < 1.1730$ , the inspection with the inspection scheme  $(i_L, f_L)$  will become uneconomic under the given constraints AOQL and  $AFI_L$ , and the production should be stopped or the maintenance will be triggered.

**3.4. Operation Procedure.** The roles of CSP-1 and the index  $S_{pk}$  in the optimal scheme are independent and complementary when considering simultaneously the two constraints AOQL and  $AFI_L$  for the process quality and cost control. The objective of employing an inspection scheme  $(i_L, f_L)$  is to reduce the proportion of nonconformance and achieve the conforming outgoing quality. The combination of  $(n, s_0)$  is used to control simultaneously the two types of risks and can

be regarded as a new stopping rule for CSP-1. The operating procedure for the optimal scheme is as follows.

*Step 1* (process quality control). Calculate the value of the parameters  $(i_L, f_L)$  under the given values of AOQL and  $AFI_L$ . Implement the inspection scheme  $(i_L, f_L)$  according to the procedure in CSP-1 for the process with the quality characteristics interested.

*Step 2* (process risk control). Calculate the value of the parameters  $(n, s_0)$  under the given values of AOQL and  $AFI_L$ . Keep the latest number  $n$  of inspection data consecutively in the order of production when the inspection scheme  $(i_L, f_L)$  in CSP-1 is performed. Calculate the value of  $\widehat{S}_{pk}$  with the latest number  $n$  of inspection data. The determination is carried out as follows:

- (i) continue Step 1 if  $\widehat{S}_{pk} \geq s_0$ ,
- (ii) stop production if  $\widehat{S}_{pk} < s_0$ .

where  $s_0$  is the threshold used to guarantee that the two risks are controlled at the stipulated level.  $n$  is the number of the last inspection data recorded consecutively during inspection, including the screening inspection stage and the fraction inspection stage in CSP-1. The number  $n$  of inspection data is used to calculate the value of  $\widehat{S}_{pk}$ .

#### 4. Analyses and Comparisons

In the proposed optimal scheme for process quality and cost control, process quality control can be achieved by performing the inspection scheme  $(i_L, f_L)$ , and process risk control can be attained with the combination of  $(n, s_0)$ . The combination of  $(n, s_0)$  also plays the role of stopping rule in the inspection scheme  $(i_L, f_L)$ . The values of  $n$  and  $s_0$  are different for various values of  $\alpha$  and  $\beta$  for specified constraints AOQL and  $AFI_L$ . Tables 3–5 show the values of the combination of  $(n, s_0)$  for  $\alpha = 0.01, 0.05, \text{ and } 0.1$ , and  $\beta = 0.01, 0.05, \text{ and } 0.1$  under three quality requirements, AOQL = 0.00018, 0.00143 and 0.0122, and three inspection capability limits,  $AFI_L = 0.6667, 0.8 \text{ and } 0.8571$ . According to the one-to-one correspondence between  $AFI_L, p_{IQL}$ , and  $S_{IQL}$ , three different process yield limits,  $S_{IQL} = 1.1534, 1.1067 \text{ and } 1.0750$  ( $AFI_L = 0.6667, 0.8 \text{ and } 0.8571$ ) under  $S_{AOQL} = 1.2485$  (AOQL = 0.00018);  $S_{IQL} = 0.9520, 0.8966 \text{ and } 0.8585$  ( $AFI_L = 0.6667, 0.8 \text{ and } 0.8571$ ) under  $S_{AOQL} = 1.0628$  (AOQL = 0.00143); and  $S_{IQL} = 0.6967, 0.6245, \text{ and } 0.5734$  ( $AFI_L = 0.6667, 0.8 \text{ and } 0.8571$ ) under  $S_{AOQL} = 0.8354$  (AOQL = 0.0122), respectively, are considered to examine the behavior of  $(n, s_0)$ . It can be observed from Tables 3–5 that the value of the sample size  $n$  becomes smaller as the  $\alpha$  or the  $\beta$  becomes larger. This phenomenon can be interpreted to mean that if the practitioner reduces the expected values at which high quality processes are rejected and/or low quality processes are accepted, the sample size for the judgement on the quality and capability of the processes will reduce. For a given  $S_{AOQL}, \alpha$ , and  $\beta$ , the sample size becomes smaller as the process yield limit decreases (the value of  $S_{IQL}$  becomes smaller). For a fixed  $\alpha, \beta$ , and  $S_{IQL}$ , the value of  $n$  becomes

smaller when the value of AOQL becomes larger (the value of  $S_{AOQL}$  becomes smaller). We can explain these phenomena by saying that the determination of the quality and capability of the processes can be done easily using a smaller number of inspection data when the difference between the quality limit  $p_{IQL}$  and the quality constraint AOQL becomes larger. For example, for  $S_{AOQL} = 1.2485, \alpha = 0.01, \text{ and } \beta = 0.01$ , the required number of inspection data is 1688 for  $S_{IQL} = 1.1534$  and only 485 for  $S_{IQL} = 1.0750$ .

Figures 2–4 show the OC curves to depict a comparison of the optimal scheme with other two AOQL contour inspection schemes in CSP-1. It can be seen from Figures 2–4 that, for various given values of the quality constraint AOQL and various values of the process quality limit  $p_{IQL}(S_{IQL})$  (which represents the cost constraint), the OC curves for the proposed integrated schemes are more ideal than the OC curves for the other two schemes in CSP-1. When  $\alpha$  and  $\beta$  take larger values, for example, when  $\alpha = 0.1$  and  $\beta = 0.1$ , the OC curves still have a more ideal shape than the curves for the other two schemes in CSP-1 at a higher yield level of  $S_{IQL}$ . When  $\alpha$  and  $\beta$  are both given smaller values, for example,  $\alpha = 0.01$  and  $\beta = 0.01$ , the OC curves are more ideal than the curves for CSP-1 at various quality limit levels of  $p_{IQL}(S_{IQL})$ . The OC curves for the optimal schemes move towards the right when the values of  $p_{IQL}$  become bigger (the values of  $S_{IQL}$  become smaller), which shows that the risk control scheme can supply the more rational probabilities of the acceptance and rejection for the quality control scheme in the optimal scheme than the AOQL contour scheme in CSP-1.

It can be noted that the inspection workload has not increased in the proposed integrated scheme because the number of inspection data  $n$  used to calculate the estimator of  $\widehat{S}_{pk}$  is recorded during the CSP-1.

#### 5. Example Application

In order to present the way in which the proposed optimal scheme can be applied in practice, the following example taken from a compressor manufacturing enterprise is considered. A cylinder is the key functioning part of an air conditioning compressor. Cylinder thickness is an important dimension for ensuring compressor performance. Based on the quality requirements given by the designer, the tolerance range of the cylinder thickness is set to  $27.784 \pm 0.002$  and the quality requirement AOQL is set to 0.00018. The current inspection scheme,  $(i, f) = (1540, 0.5)$ , is adopted in CSP-1 to control the outgoing quality. For using the optimal scheme, the values of the four constraints are specified as AOQL = 0.00018 ( $S_{AOQL} = 1.2485$ ),  $AFI_L = 0.8571$  ( $S_{IQL} = 1.0750$ ),  $\alpha = 0.05$ , and  $\beta = 0.05$ . From Table 3, the optimal scheme can be found as  $(i_L, f_L, n, s_0) = (925, 0.6515, 242, 1.1553)$ . Using the proposed control scheme, the probability of accepting an in-control process with high quality level (e.g., the non-conforming fraction of the in-control process is lower than the value of AOQL = 0.00018) is greater than  $1 - \alpha = 0.95$ . The probability of accepting an in-control process with low quality level (e.g., the value of the index  $\widehat{S}_{pk}$  is lower than the value of  $S_{IQL} = 1.0750$ ) is less than the value of  $\beta = 0.05$ . Performing the quality control scheme  $(i_L, f_L) = (925, 0.6515)$ ,

TABLE 3: The values of  $(n, s_0)$  at  $S_{AOQL} = 1.2485$  (AOQL = 0.00018) and three levels of  $S_{IQL}(AFI_L)$  for  $\alpha = 0.01, 0.05,$  and  $0.1$  and  $\beta = 0.01, 0.05,$  and  $0.1$ .

$\alpha$	$\beta$	$S_{IQL} = 1.1534$ ( $AFI_L = 0.6667$ ) $(i_L, f_L) = (2776, 0.3086)$		$S_{IQL} = 1.1067$ ( $AFI_L = 0.8$ ) $(i_L, f_L) = (1388, 0.5342)$		$S_{IQL} = 1.0750$ ( $AFI_L = 0.8571$ ) $(i_L, f_L) = (925, 0.6515)$	
		$n$	$s_0$	$n$	$s_0$	$n$	$s_0$
		0.01	0.01	1688	1.1985	740	1.1730
	0.05	1271	1.1895	551	1.1610	362	1.1406
	0.10	1046	1.1850	459	1.1520	305	1.1308
0.05	0.01	1229	1.2060	523	1.1850	345	1.1702
	0.05	843	1.1985	370	1.1730	242	1.1553
	0.10	673	1.1925	296	1.1640	195	1.1446
0.1	0.01	996	1.2120	431	1.1940	279	1.1808
	0.05	661	1.2045	290	1.1820	188	1.1661
	0.10	512	1.1985	225	1.1730	147	1.1553

TABLE 4: The values of  $(n, s_0)$  at  $S_{AOQL} = 1.2485$  (AOQL = 0.00018) and three levels of  $S_{IQL}(AFI_L)$  for  $\alpha = 0.01, 0.05,$  and  $0.1$  and  $\beta = 0.01, 0.05,$  and  $0.1$ .

$\alpha$	$\beta$	$S_{IQL} = 0.9520$ ( $AFI_L = 0.6667$ ) $(i_L, f_L) = (348, 0.3092)$		$S_{IQL} = 0.8966$ ( $AFI_L = 0.8$ ) $(i_L, f_L) = (174, 0.5351)$		$S_{IQL} = 0.8585$ ( $AFI_L = 0.8571$ ) $(i_L, f_L) = (115, 0.6527)$	
		$n$	$s_0$	$n$	$s_0$	$n$	$s_0$
		0.01	0.01	874	1.0035	371	0.9720
	0.05	655	0.9945	281	0.9585	181	0.9327
	0.10	554	0.9885	236	0.9480	153	0.9214
0.05	0.01	639	1.0125	266	0.9870	168	0.9674
	0.05	437	1.0035	186	0.9720	120	0.9498
	0.10	358	0.9975	150	0.9615	97	0.9373
0.1	0.01	508	1.0200	214	0.9960	135	0.9799
	0.05	346	1.0110	144	0.9825	92	0.9624
	0.10	265	1.0035	113	0.9720	73	0.9498

TABLE 5: The values of  $(n, s_0)$  at  $S_{AOQL} = 1.2485$  (AOQL = 0.00018) and three levels of  $S_{IQL}(AFI_L)$  for  $\alpha = 0.01, 0.05,$  and  $0.1$  and  $\beta = 0.01, 0.05,$  and  $0.1$ .

$\alpha$	$\beta$	$S_{IQL} = 0.6967$ ( $AFI_L = 0.6667$ ) $(i_L, f_L) = (39, 0.3145)$		$S_{IQL} = 0.6245$ ( $AFI_L = 0.8$ ) $(i_L, f_L) = (19, 0.5437)$		$S_{IQL} = 0.5734$ ( $AFI_L = 0.8571$ ) $(i_L, f_L) = (12, 0.6629)$	
		$n$	$s_0$	$n$	$s_0$	$n$	$s_0$
		0.01	0.01	324	0.7590	129	0.7140
	0.05	246	0.7470	99	0.6960	61	0.6590
	0.10	206	0.7395	85	0.6855	52	0.6452
0.05	0.01	229	0.7710	89	0.7320	53	0.7026
	0.05	163	0.7590	65	0.7140	39	0.6800
	0.10	133	0.7500	53	0.7012	32	0.6647
0.1	0.01	187	0.7800	71	0.7455	42	0.7188
	0.05	127	0.7680	49	0.7267	29	0.6961
	0.10	100	0.7590	39	0.7140	24	0.6800

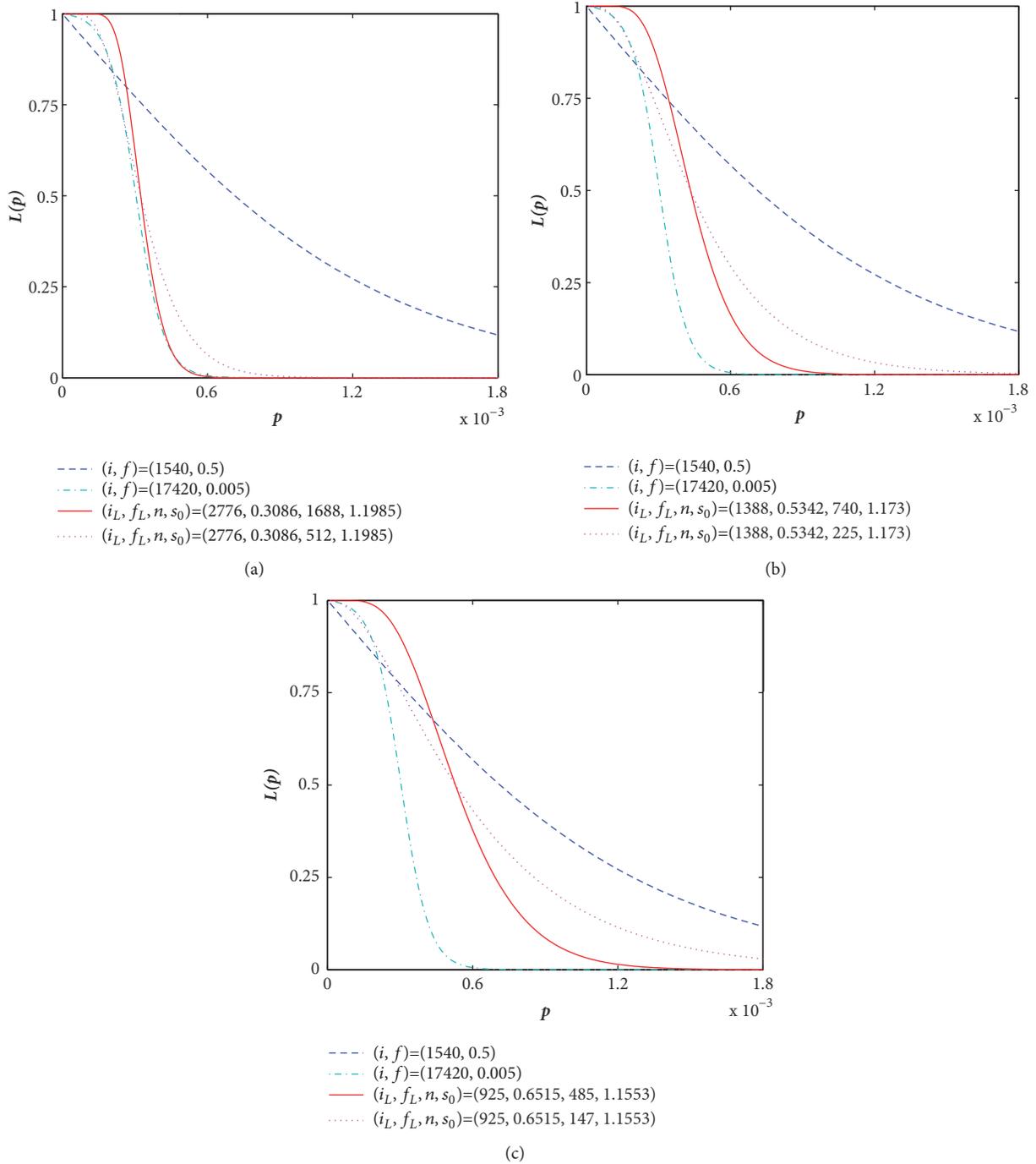


FIGURE 2: Comparisons of OC curves for the optimal schemes and inspection schemes in CSP-1 under AOQL = 0.00018 (a) at  $AFI_L = 0.6667$ , (b) at  $AFI_L = 0.8$ , and (c) at  $AFI_L = 0.8571$ .

242 sample items are recorded consecutively, as shown in Table 6. The normal probability plot of the 242 sample items is displayed in Figure 5. Figure 6 shows the histogram of the 242 observed data with the lower and upper specification limits. Obviously, based on the normality test in Figures 5 and 6, the in-control process is shown to be close to a normal distribution.

Carrying out the calculation with the 242 inspection data, we get  $\bar{x} = 27.7842$ ,  $s = 0.000517$ ,  $\hat{S}_{pk} = 1.2144$ . Obviously, the estimator of  $\hat{S}_{pk} = 1.2144$  is greater than the value of  $s_0 = 1.1553$ . The optimal scheme which can simultaneously meet the two constraints AOQL = 0.00018 and  $AFI_L = 0.8571$  will be continued. Table 7 shows a comparison of the

TABLE 6: The 242 observations collected consecutively from the inspection process in CSP-1.

277834	277844	277839	277834	277844	277838	277851	277836	27784	277842
277834	277842	277836	277832	277846	27785	277844	277837	277844	277840
277844	277844	277835	277834	277839	277852	277845	277836	277841	277848
277844	277849	277835	277836	277838	277846	27785	277841	277845	277854
277849	277842	277837	277834	277846	277847	277844	277844	277841	277845
277844	277841	277838	277830	277840	277843	277844	277839	277844	277843
277847	277844	277844	27783	277842	277845	277840	277837	277843	277849
277849	277835	277846	277834	277839	277849	277839	277844	277835	277845
277849	277836	277834	277844	277842	277846	277837	27784	277842	277846
277849	277836	277830	277842	277845	277844	277847	277844	277843	277848
277847	277846	277831	277841	277844	277851	277836	277844	277842	277849
277846	277836	277832	277842	277843	277846	277836	277842	277834	277843
277849	277836	277834	277838	277845	277845	277838	277842	277834	277840
277844	277839	277832	277841	277847	277847	277841	277842	277834	277840
277846	277841	277829	277842	277844	277849	277844	277845	277834	277837
277842	277837	277835	277844	277846	277852	277846	277843	277835	277844
277849	277839	277832	277844	277849	277849	277842	277844	277835	277841
277844	277838	277834	277841	277844	277848	277848	277844	277835	277843
277844	277834	277833	277844	277844	277847	277841	277844	277847	277844
277842	277841	277843	277844	277844	277837	277842	277843	27784	277844
277836	277836	277836	277842	277842	277845	277845	277844	277844	277843
277846	277840	277845	277842	277838	277836	277836	277836	277836	277837

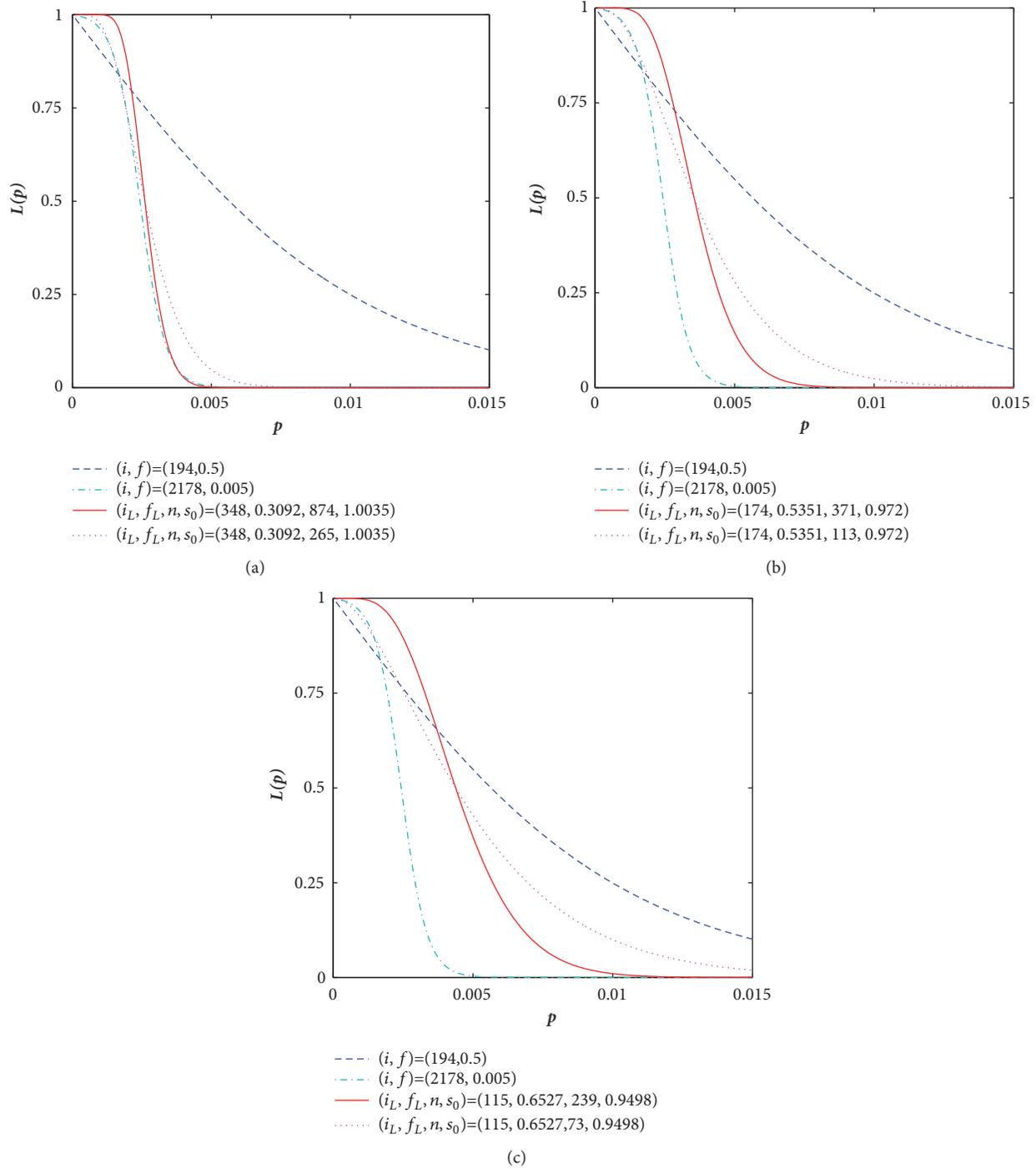


FIGURE 3: Comparisons of OC curves for the optimal schemes and inspection schemes in CSP-1 under  $AOQL = 0.00143$  (a) at  $AFI_L = 0.6667$ , (b) at  $AFI_L = 0.8$ , and (c) at  $AFI_L = 0.8571$ .

performances between the original inspection scheme and the optimal scheme, where  $L(p)$  represents the probability of acceptance. It can be seen from Table 7 that the performance  $AFI$  increases by 0.103172,  $AOQ$  decreases by 0.000028, and the value of the probability of acceptance increases by 0.062855.

## 6. Conclusions

An optimal scheme for the process quality and cost control is proposed to monitor the process capability and improve process quality. The CSP-1 and  $S_{pk}$ , which play independent and complimentary roles, are integrated in the optimal scheme.

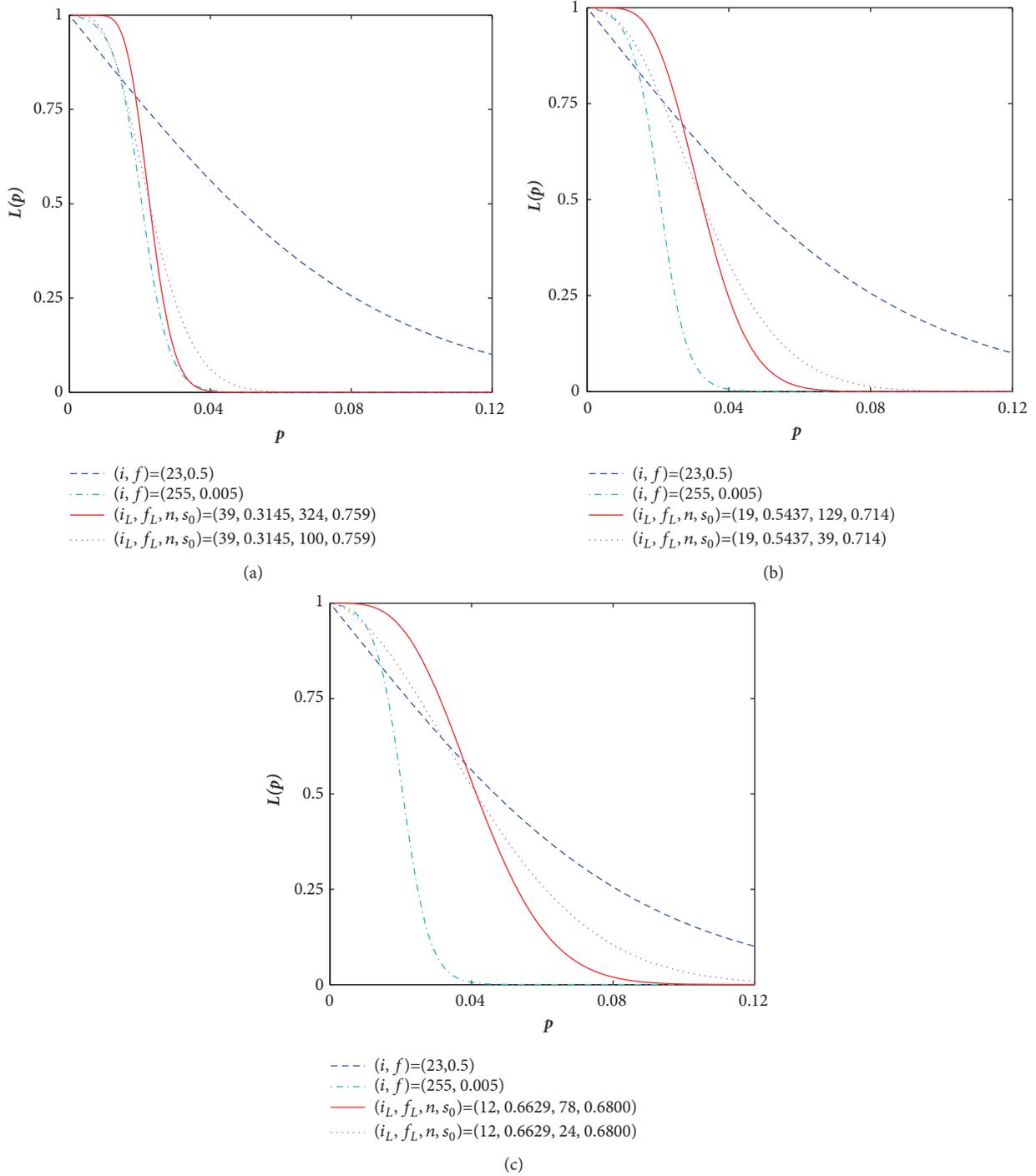


FIGURE 4: Comparisons of OC curves for the optimal schemes and inspection schemes in CSP-1 under AOQL = 0.0122 (a) at  $AFI_L = 0.6667$ , (b) at  $AFI_L = 0.8$ , and (c) at  $AFI_L = 0.8571$ .

The sole feasible inspection scheme in CSP-1 for meeting concurrently the quality and cost constraints is one of the AOQL contour schemes in CSP-1, which occurs at the point of  $p_L = p_{IQL}$ . Two types of risk under quality and cost constraints are simultaneously controlled at the stipulated level with the risk control scheme. The risk control scheme is constructed with two nonlinear inequalities based on

the accurate distribution of the natural estimator  $\hat{S}_{pk}$ . The combination of the two risk control parameters  $(n, s_0)$  plays the role of the stopping rule in the inspection scheme  $(i_L, f_L)$ . There is a one-to-one correspondence between the values of the four parameters  $(i_L, f_L, n, s_0)$  and the given values of the four constraints (quality, cost, and the two risks). The proposed optimal scheme shows the advantages over the

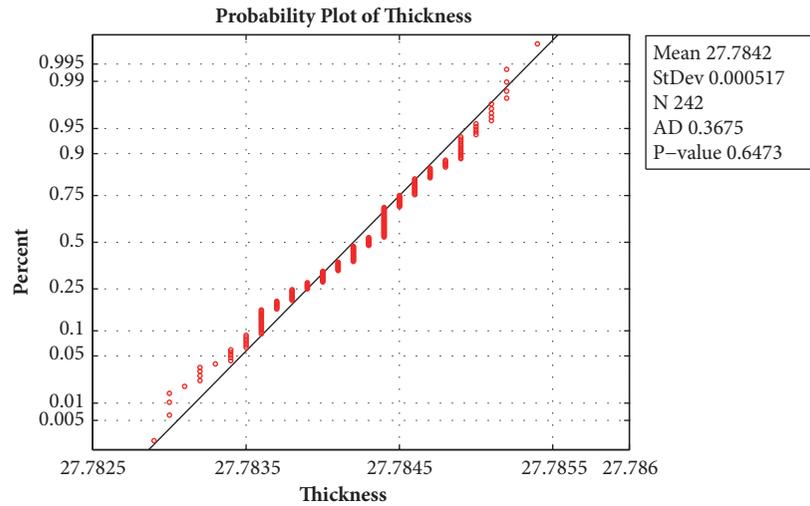


FIGURE 5: The normal probability plot of the 242 observations.

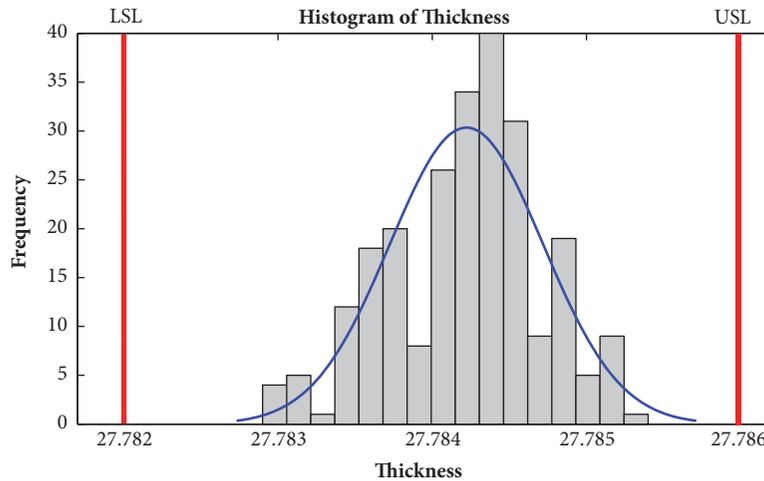


FIGURE 6: The histogram of the 242 observations with double specification limits.

TABLE 7: Comparison of the performances of  $AFI$ ,  $AOQ$ , and  $L(p)$ .

	$(i, f) = (1540, 0.5)$	$(i_L, f_L, n, s_0) = (925, 0.6515, 242, 1.1553)$
$AFI$	0.602509	0.705681
$AOQ$	0.000107	0.000079
$L(p)$	0.794982	0.857837

original process control tools in the measures of  $AFI$ ,  $AOQ$ , and the probability of acceptance.

**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.

**References**

- [1] H. F. Dodge, "A sampling inspection plan for continuous production," *Annals of Mathematical Statistics*, vol. 14, pp. 264–279, 1943.
- [2] K. Govindaraju and C. Kandasamy, "Design of generalized CSP-C continuous sampling plan," *Journal of Applied Statistics*, vol. 27, no. 7, pp. 829–841, 2000.
- [3] G. J. Lieberman and H. Solomon, "Multi-level continuous sampling plans," *Annals of Mathematical Statistics*, vol. 26, pp. 686–704, 1955.
- [4] P. Guayjarernpanishk and T. Mayureesawan, "The design of two-level continuous sampling plan MCSP-2-C," *Applied Mathematical Sciences*, vol. 6, no. 89-92, pp. 4483–4495, 2012.

- [5] S. Balamurali and K. Subramani, "Modified CSP-C continuous sampling plan for consumer protection," *Journal of Applied Statistics*, vol. 31, no. 4, pp. 481–494, 2004.
- [6] K. Govindaraju and S. Balamurali, "Tightened single-level continuous sampling plan," *Journal of Applied Statistics*, vol. 25, no. 4, pp. 451–461, 1998.
- [7] C. Kandasamy and K. Govindaraju, "Selection of tightened two-level continuous sampling plans," *Journal of Applied Statistics*, vol. 20, no. 2, pp. 271–284, 1993.
- [8] V. S. S. Kumar, "A tightened M-level continuous sampling plan for Markov-dependent production processes," *IIE Transactions*, vol. 16, no. 3, pp. 257–261, 1984.
- [9] C. R. Cassady, L. M. Maillart, I. J. Rehmert, and J. A. Nachlas, "Demonstrating deming's kp rule using an economic model of the CSP-1," *Quality Engineering*, vol. 12, no. 3, pp. 327–334, 2000.
- [10] A. J. Duncan, *Quality control and industrial statistics*, Revised ed, Richard D. Irwin, Inc., Homewood, Ill., 1959.
- [11] M. Eleftheriou and N. Farmakis, "Expected cost for continuous sampling plans," *Communications in Statistics—Theory and Methods*, vol. 40, no. 16, pp. 2969–2984, 2011.
- [12] H. F. Yu, W. C. Yu, and W. P. Wu, "A mixed inspection policy for CSP-1 and precise inspection under inspection errors and return cost," *Computers & Industrial Engineering*, vol. 57, no. 3, pp. 652–659, 2009.
- [13] A. Haji and R. Haji, "The optimal policy for a sampling plan in continuous production in terms of the clearance number," *Computers & Industrial Engineering*, vol. 47, no. 2-3, pp. 141–147, 2004.
- [14] C. H. Chen and C. Y. Chou, "Economic design of continuous sampling plan under linear inspection cost," *Journal of Applied Statistics*, vol. 29, no. 7, pp. 1003–1009, 2002.
- [15] W. Xu, Y. Y. Yu, and Q. S. Zhang, "An evaluation method of comprehensive product quality for customer satisfaction based on intuitionistic fuzzy number," *Discrete Dynamics in Nature Society*, vol. 2018, no. 1, p. 12, 2018.
- [16] B. Bouslah, A. Gharbi, and R. Pellerin, "Joint economic design of production, continuous sampling inspection and preventive maintenance of a deteriorating production system," *International Journal of Production Economics*, vol. 173, pp. 184–198, 2016.
- [17] J. Zhang, "Optimal control problem of converter steelmaking production process based on operation optimization method," *Discrete Dynamics in Nature and Society*, Art. ID 483674, 13 pages, 2015.
- [18] R. Mehdi, R. Nidhal, and C. Anis, "Integrated maintenance and control policy based on quality control," *Computers & Industrial Engineering*, vol. 58, no. 3, pp. 443–451, 2010.
- [19] C. H. Chen, C. Y. Chou, and T. S. Cheng, "Joint design of continuous sampling plans and specification limits," *The International Journal of Advanced Manufacturing Technology*, vol. 21, pp. 235–237, 2003.
- [20] C. H. Chen and C.-Y. Chou, "Joint design of economic manufacturing quantity, sampling plan and specification limits," *Economic Quality Control*, vol. 17, no. 2, pp. 145–153, 2002.
- [21] C. H. Chen and C. Y. Chou, "Design a continuous sampling plan based on quadratic quality loss function," *Asia Pacific Management Review*, vol. 6, no. 4, pp. 485–489, 2001.
- [22] S. Muthulakshmi, "Stopping rules to limit inspection effort in CSP-C continuous sampling plan," *International Journal of Mathematical Archive*, vol. 3, pp. 656–662, 2012.
- [23] Y. L. Fan, "Two more excellent stopping rules for continuous sampling plan," *Journal of Applied Statistics*, vol. 19, no. 6, pp. 24–30, 2000.
- [24] Y. L. Fan, "Continuous sampling plan CSP-1 with stopping inspection rule ( $n^*-1$ )," *Acta Mathematicae Applicatae Sinica*, vol. 19, no. 1, pp. 135–143, 1995.
- [25] R. B. Murphy, "Stopping rules with CSP-1 sampling plans," *Industrial Control*, vol. 16, no. 5, pp. 10–16, 1959.
- [26] R. A. Boyles, "Process capability with asymmetric tolerances," *Communications in Statistics—Simulation and Computation*, vol. 23, no. 3, pp. 615–643, 1994.
- [27] J. C. Lee, H. N. Hung, W. L. Pearn, and T. L. Kueng, "On the distribution of the estimated process yield index Spk," *Quality and Reliability Engineering International*, vol. 18, pp. 111–116, 2002.
- [28] C. W. Wu and M.-Y. Liao, "Estimating and testing process yield with imprecise data," *Expert Systems with Applications*, vol. 36, no. 8, pp. 11006–11012, 2009.
- [29] W. L. Pearn and C. C. Chuang, "Accuracy analysis of the estimated process yield based on Spk," *Quality and Reliability Engineering International*, vol. 20, no. 4, pp. 305–316, 2004.
- [30] F. K. Wang, "Process Yield for Multiple Stream Processes with Individual Observations and Subsamples," *Quality and Reliability Engineering International*, vol. 32, no. 2, pp. 335–344, 2016.
- [31] F. K. Wang, "Measuring the process yield for circular profiles," *Quality and Reliability Engineering International*, vol. 31, no. 4, pp. 579–588, 2015.
- [32] F. K. Wang and Y. Tamirat, "Process yield analysis for autocorrelation between linear profiles," *Computers & Industrial Engineering*, vol. 71, no. 1, pp. 50–56, 2014.
- [33] A. Lepore, B. Palumbo, and P. Castagliola, "A note on decision making method for product acceptance based on process capability indices  $C_{pk}$  and  $C_{pmk}$ ," *European Journal of Operational Research*, vol. 267, no. 1, pp. 393–398, 2018.
- [34] M. S. Fallah Nezhad and S. Seifi, "Repetitive group sampling plan based on the process capability index for the lot acceptance problem," *Journal of Statistical Computation and Simulation*, vol. 87, no. 1, pp. 29–41, 2017.
- [35] Y. Tamirat and F. K. Wang, "Sampling Plan based on the Exponentially Weighted Moving Average Yield Index for Autocorrelation within Linear Profiles," *Quality and Reliability Engineering International*, vol. 32, no. 5, pp. 1757–1768, 2016.
- [36] I. Negrin, Y. Parmet, and E. Schechtman, "Developing a sampling plan based on Cpk-unknown variance," *Quality and Reliability Engineering International*, vol. 27, no. 1, pp. 3–14, 2011.
- [37] G. L. Yang and G. L. Yang, "A renewal-process approach to continuous sampling plans," *Technometrics*, vol. 25, no. 1, pp. 59–67, 1983.
- [38] C. Z. Li, S. R. Tong, and K. Q. Wang, "Optimal CSP-1 boundary scheme based on the estimator of the proportion of conformance for specified in-control process," *Quality Technology & Quantitative Management*, 2018.

## Research Article

# Data-Driven Recovery Potential Analysis and Modeling for Batteries Recovery Operations in Electric Bicycle Industry

Ping Zhang  and Guangfu Liu

School of Economics & Management, Tongji University, Shanghai 200092, China

Correspondence should be addressed to Ping Zhang; 1310362@tongji.edu.cn

Received 7 September 2018; Accepted 22 October 2018; Published 11 November 2018

Academic Editor: Lu Zhen

Copyright © 2018 Ping Zhang and Guangfu Liu. 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.

To help the government manage waste lead-acid batteries in a more targeted and sustainable way, accurately forecasting the number of waste lead-acid batteries and analyzing their recovery potential play a key role. In China, electric bicycles are one of the most common means of transportation. As of the end of 2017, the social holding quantity of electric bicycles in China was over 250 million and that of electric tricycles was over 50 million. The quantity is equal to the total number of electric bicycles manufactured between 2011 and 2017. Currently, 90% of electric bicycles adopt lead-acid batteries as their power batteries. However, there are a few studies on the lead-acid batteries used in electric bicycles as power batteries. In this paper, we have selected lead-acid batteries used in electric bicycles as the subject of research as such kind of batteries enjoys the widest user base, the most single-battery consumption volume, and the strongest mobility. Based on the output and sales of electric bicycles, we have obtained the quantity of power lead-acid batteries. We have then estimated the annual waste quantity of lead-acid batteries used in electric bicycles in 2000-2022 using the “market supply A model” and the “Stanford Model”, respectively, and based on the proportion of raw materials contained in lead-acid batteries and the proportion between reclaimed and discarded lead-acid batteries, we have estimated the recovery potential of discarded lead-acid batteries in 2000-2022. We estimate that the lead-acid batteries used in electric bicycles only have great recovery potential and there are abundant potential resources for recovery. The research data and results can help decision-makers make more effective and more accurate management measures and policies.

## 1. Overview of Waste Lead-Acid Batteries

On May 15, 2018, the mandatory national standard *Technical Code for Safety of Electric Bicycles* (hereinafter referred to as *Technical Code*) was approved for issuance by the State Administration for Market Regulation and the Standardization Administration of the People's Republic of China through Standard Notice of the People's Republic of China (No. 72018) and will be officially implemented since April 15, 2019. The *Technical Code* states that electric bicycles shall have pedal power, the maximum design speed shall be no more than 25km/h, the total weight (including battery) shall be no more than 55kg, the motor power shall be no more than 400W, and the nominal battery voltage shall be no more than 48V. The period from May 15, 2018, to April 14, 2019, is the transitional period, during which manufacturers are encouraged to organize production according to the *Technical Code*, sales companies to

sell products that meet the requirements specified in the *Technical Code*, and consumers to buy products that meet the requirements specified in the *Technical Code*. After the *Technical Code* is officially implemented, products that do not meet the requirements specified in the *Technical Code* shall not be manufactured, sold, or imported. The issuance of the *Technical Code* has not only accelerated the transformation and upgrading of the electric bicycle industry, but also greatly simulated the demand of the electric bicycle industry for the small, light lithium batteries with stable voltage because of mandatory norms on speed, motor power, and voltage, especially regulations on the weight of batteries. In the meantime, as the absolute dominate products, lead-acid batteries account for over 90% of battery products in the electric bicycle market [1]. The cheap but large and heavy lead-acid batteries as power batteries of electric bicycles are faced with accelerated replacement or direct elimination. To properly recycle and process large

quantities of waste lead-acid batteries will be a big challenge.

In foreign countries, waste lead-acid batteries have become a hot topic in circular economy because of its relatively high recycling value. On the other hand, if they are improperly handled, they can easily cause serious environmental pollution and threaten our health. Therefore, waste lead-acid batteries are internationally recognized as hazardous waste. As early as in the 1990s, the EU issued Directive 91/157/EEC on batteries containing hazardous substances, which include lead-acid batteries. The Directive sets relevant directive on battery labeling, margin system and environmental protection input, and so on [2]. Nowadays the recovery rate of lead-acid batteries in developed countries has basically reached 100% [3]. Developed countries attach great importance to the recycling of waste batteries and the production of secondary lead. The average annual output of secondary lead enterprises in developed countries is as high as more than 70,000 tons. In 1998, the total lead output in Western countries was 4.896 million tons, of which that of secondary lead was 2.846 million tons, accounting for 58.13% of the total lead output; the total lead output of the United States was 1.422 million tons, of which that of secondary lead was 1.083 million tons, accounting for 76.3%; the output density of secondary lead in countries like Germany, France, and Sweden was all over 50% [4]. In the 1960s, the world output of primary lead started to decline, while that of secondary lead gradually rose. In the 1990s, the world output of secondary lead exceeded that of primary lead. The main raw materials used to manufacture secondary lead are used lead-acid batteries. Currently lead-acid batteries already account for over 85% of raw materials of secondary lead [2]. In 2017, the global lead-acid battery market size was about USD42.9 billion, up 1% YoY. The lead consumption of lead-acid batteries in the United States accounts for over 95% of the country's total lead consumption. Thanks to sound regulations and effective management, the lead emission from manufacturing of lead-acid batteries accounts for only 1.5% of total emissions. In 2008, the US government removed lead-acid battery manufacturing from main lead pollution sources [5].

China is the largest manufacturer and seller of electric bicycles and also the world's largest manufacturer and consumer of lead accumulators. According to the data of the National Bureau of Statistics, by the end of 2017, the number of electric bicycles registered or in use was over 250 million, and that of electric tricycles was over 50 million (Ye. X. H. 2018) [6]. Since 2011, the production of electric bicycles has grown by 30 million to 37 million per year and that of electric tricycles by 7.6 million to 10 million a year [7]. The newly increased production of the two each year is nearly 40 million to 50 million. For a long time over 95% of the power batteries used in electric bicycles in China are lead-acid batteries. Currently lead-acid batteries are still the primary power batteries in the electric bicycle industry. Even though lithium batteries are better in performance, energy density, and service life, as their prices are relatively high, the proportion of lithium batteries in commercially available vehicle models is less than 10% [8].

The service life of the lead-acid batteries used as power batteries of electric bicycles is generally 1 to 3 years and the average service cycle is 2 years. With increasing production of lead-acid batteries and increasing number of electric bicycles registered or in use, the number of scrapped and replaced lead-acid batteries is huge. According to incomplete statistics, the total weight of waste lead-acid batteries in China every year is around 4 million tons and it is growing at an annual rate of about 15% (Li X. Z. 2016) [9]. Around 2000, about 50 million lead-acid batteries or over 300,000 tons were scrapped in China every year (Ma Y. G. 2000) [10]; the weight of waste lead batteries produced in 2015 was more than 2.6 million tons [11]; in 2016 there were about 3.5 million tons of waste batteries [12], 11.7 times that in 2000, and the annual growth rate of waste lead-acid batteries in 2016 was over 30%. However, according to statistics, the number of waste lead-acid batteries recycled with proper methods is less than 30% [13]. That is to say, over 70% of waste lead-acid batteries are illegally recycled and dismantled.

The acid solution in lead-acid batteries contains various heavy metals such as lead, zinc, manganese, and cadmium. Improperly dismantling or processing waste lead-acid batteries can cause lead dust pollution and blood lead poisoning (acute or chronic injection of lead into the body can cause nerve metabolic, reproductive, and mental diseases and can even result in death). As for improper pouring of waste lead solution, the heavy metals resolved from the acid solution and toxic waste solution will cause severe pollution to soil, plants, rivers, surface water, ground water, air, etc., seriously threatening ecological balance and human health.

On the other hand, seen from the global lead consumption structure, 86% of the downstream lead demands are used to manufacture batteries, while those used in paint, boards, and alloy account for 5%, 4%, and 2%, respectively. Therefore, the battery industry is the main downstream application of secondary lead [14]. In China around 80% of refined lead production is used to manufacture lead-acid batteries every year (Li S. L. 2018) [15]. The Chinese secondary lead industry developed slowly. In the 1950s, the annual production of secondary lead hovered around the thousand tonnage. In 1990 its production was 28,200 tons, in 1994 it reached 95,000 tons, and in 1995 it broke 100,000 tons. Between 1990 and 1993, the production of secondary lead accounted for about 10% of the total production of refined lead, and it increased to around 20% since 1994 (Ma Y. G. 2000) [16]. In 2013, the production of secondary lead in China was 1.5 million tons, up 7.1% year on year, and it is estimated that it accounted for over 30% [17] of the consumption of refined lead that year. In 2015, the production of refined lead in China was about 4.7 million tons and that of secondary lead was about 1.6 million tons, and the proportion of secondary lead production was about 35% [18]. This shows that after nearly 30 years' development, the use ratio of secondary lead has only increased from 10% in the early 1990s to about 35% today. Waste lead-acid batteries are main raw materials of secondary lead. According to incomplete statistics, the use ratio of secondary lead is 90% in the United States, 85% in Japan and 80%-90% in Europe, while in China the ratio of

secondary lead actually used is less than 50% (Xue X. 2016) [19].

Moreover, waste lead batteries are composed of 74% lead and its compound, 20% sulfuric acid, and 6% plastics, which have relatively high resource recovery value [24]. According to statistics, in 2012 the global battery market size was USD75.975 billion, of which the market size of lead-acid batteries was USD39.294 billion, which was the largest. And 97% of the lead in lead-acid batteries can be recycled, making them the consumables with the highest reclamation rate (Lv X. L. 2013) [25].

Therefore, if we can further improve the recovery of waste lead-acid batteries and increase the use ratio of secondary lead, the “urban mine” of waste lead-acid batteries will have irreplaceable strategic significance to our protection of primary metal resources, utilization of secondary metal resources, development of the secondary lead industry, and national metal resource security.

## 2. Related Works

With low prices, simple production technology, high recovery rate, great recovery value, and a high use ratio of secondary lead, lead-acid batteries have incomparable advantages in the secondary battery field. Renowned lead-acid battery scientist Detchko Pavlov writes in the preface of his book *Lead-acid Batteries: Science and Technology* that because they use a high proportion of secondary lead and are easy to produce, lead-acid batteries are currently the lowest-cost chemical power source. For decades lead-acid batteries have taken up 65%-70% (Pavlov D. 2015) [26] of global chemical power source production. Lead-acid batteries are cheap, easy to produce, and easy to recycle, and the resources for their manufacture are practically unlimited (Pavlov D. 2015) [27]. Waste lead-acid batteries have very high recovery value. The grids in waste lead-acid batteries and the lead slime containing  $PbO_2$  and  $PbSO_4$  are main sources of secondary lead. In China over 85% of the raw materials of secondary lead are from waste lead-acid batteries, and 50% of secondary lead is used to produce storage batteries (Zhu S.R. 2002) [28]. Therefore, to be able to reasonably, scientifically, and accurately predict the generation amount of waste lead-acid batteries and estimate the amount of secondary lead and other renewable resources recycled from waste lead-acid batteries is of great significance to the waste metal resource recovery, ecological environment protection, metal resource safety, and sustainable development of the (secondary) lead industry of China.

As lead-acid batteries are very similar and comparable to electronic and electrical products in service characteristics, scrapping cycle, and recovery method, the studies that mainly use waste electronic and electrical products as the subject of forecast can be used for reference. Liu Xiaoli et al. (2005) [29] estimated the annual waste quantity of five major categories of electronic and electrical products in 2000-2010 in China using the Stanford Model. Liang Xiaohui et al. (2009) [30] forecasted the waste quantity, recovery amount, storage amount, circulating amount, and filling amount of

five categories of electronic products in 2008-2012 using the Carnegie Mellon Model. Tang Hongxia (2009) [31] forecasted the generation amount of waste electrical and electronic products in Shanghai using the “expert estimation” model. Tong Xin et al. (2013) [32] compared the forecast results of the generation amount of electronic waste in China based on the Stanford Model and the holding quantity coefficient method with the actual recovery data in the pilot areas of the “old for new” activity of home appliances in China. Wang Qi et al. (2014) [33] conducted a comparative study of the four most representative estimation models for waste electronic products, namely, the Carnegie Mellon Model, the market supply model, the Stanford Model, and the time sequence model, and presented concrete calculation examples to show their advantages, disadvantages, and applicable scope. Zhang Han et al. (2016) [34] established a model and estimated and compared the annual waste quantity of primary batteries, lead-acid batteries, lithium batteries, and Ni-MH batteries in China between 2011 and 2020.

Meanwhile, researchers have conducted meaningful studies with various research methods to forecast the amounts of different wastes in the world. For example, V. P. Ulnikovica et al. (2012) [35] presented a methodology for the assessment of waste material quantities that was developed as part of the Technological Development Project TR 21037 of the Republic of Serbia. Required information on the amount of traffic, vessel types, and numbers as well as the number of dockings was extracted from questionnaires and interviews with watermen and researchers to determine the quantity of vessel-generated waste. V. Bijayashree et al. (2014) [36] forecasted municipal solid waste quantity and composition, respectively, through a multiple linear regression model and system dynamics model and forecasted the generation rate, amount, and waste stream of solid waste in India's capital Delhi in 2011-2014. R. Intharathirat et al. (2015) [37] forecasted MSW collected in Thailand with prediction interval in long term period by using the optimized multivariate grey model which is the mathematical approach. Li Xin et al. (2017) [38] by setting the consumption intensity, recovery intensity, and life distribution functions of minerals and analyzing the historical experience of industrialized countries such as the United States, the United Kingdom, Germany, Japan, and China on the consumption, storage, and recovery of metal products between 1949 and 2015, taking copper, steel, and aluminum as examples, predicted the consumption and waste recovery variation trend of the three metals in China between 2016 and 2030. P. van Der Werf et al. (2018) [39] adopted direct measurement of waste streams through waste composition studies to estimate the quantity of food waste disposed in the garbage stream by households in southern Ontario, Canada, and determine if this common methodology could be expanded and serve as the basis of a standardized and rigorous household food waste measurement methodology. A. K. Awasthi et al. (2018) [40] revealed the presence of a strong linear correlation among global e-waste generation and Gross Domestic Product by comparing the best fit for data relationship between e-waste collected volumes and GDP PPS. They also held that because e-waste contains valuable metals such as copper, gold, and silver and their

content is higher in precious metals than in mineral ores, the better collection of e-waste acts an important role concerning the circular economy.

Regarding “recovery potential” of waste, scholars also have studied it from multiple aspects. Some scholars studied the quantity of resources with recovery value in municipal solid waste and the economic benefit they can bring. Some scholars estimated the content of precious metals in waste pipes and cables that remain underground in cities and concluded that they have great potential for urban mining. Some scholars studied the types and recovery potential of landfill mined plastic wastes and determined the feasibility of landfill mining projects based on that. Some scholars determined the theoretical recovery potential of 57 elements a complete survey of the sewage sludge ash (SSA) from mono-incineration facilities and concluded that SSA is an important secondary resource of P. M. Alamgir et al. (2007) [41] analyzed the contents of components in solid waste such as organics, paper, and plastic, evaluated the potential for recovery and reduction based on the waste characteristics, and predicted the economic benefit that can be earned from recycling and composting of municipal solid waste, based on the study of the types and generation amount of solid waste in six major cities in Bangladesh in 2005. B. Wallsten et al. (2013) [42] analyzed the pipes and cables that remain in the ground after being taken out of use or disconnected for a number of reasons using the GIS-based MFA method and found that they contain rich mineral resources such as copper, aluminum, and iron. They believe that these infrastructures “cold spots” are hibernating stock with a significant potential for urban mining. C. Zhou. et al. (2014) [43] study the characteristics of the landfill mined plastic wastes and their recovery potential to determine the feasibility of landfill mining project. O. Krüger C. Adam (2015) [44] conducted a complete survey of the sewage sludge ash (SSA) from German mono-incineration facilities and determined the theoretical recovery potential of 57 elements. German SSA contains up to 19,000t/a P which equals approximately 13% of phosphorus applied in the German agriculture in form of phosphate rock based mineral fertilizers. Thus, SSA is an important secondary resource of P. Stijn van Ewijk et al. (2018) [45] estimated the recovery potential of waste in global paper life cycle using the life-cycle assessment method, and they evaluated the use of global paper materials and the ideal waste recovery potential using the “recovery potential” index.

Life-cycle assessment (LCA) is a standardized method (Ciroth A et al. 2011) [46] to assess environmental impacts associated with all the stages of a product's, process's or activity's life from raw material extraction through materials processing, manufacture, distribution, use, recovery, maintenance, and disposal. Its research focus is to establish the life-cycle environmental impact assessment model and collect data or use LCA database data to analyze the environmental impacts of the subject of study in its whole life-cycle. For example, H. A. Arafat et al. (2015) [47] assessed the environmental impacts of five municipal solid waste (MSW) treatment processes using the life-cycle assessment (LCA) tool.

As far as the subject is concerned, there are almost no studies on the subject of the waste quantity of lead-acid batteries in China, especially the lead-acid batteries used as power batteries in electric bicycles. In terms of research methods, the common estimation models used to estimate waste quantity are the market supply model, the market supply A model, the Stanford Model, the Carnegie Mellon Model, the time sequence model, the expert estimation model, the holding quantity coefficient method, the ICER model, etc. Due to limited data, early researcher obtained some data needed for calculation from indirect calculation of other data that they could collect at that time. Moreover, with the development of society, some key data needed for calculation such as product service life were adjusted due to rapid social development, quick changes in industries and the issue of particular policies. Currently the holding quantity of electric bicycles (including tricycles) in China is more than 300 million [7], and annual sales volume is nearly 40 million [7]. Each electric bicycle is equipped with 3 to 4 battery sets, while over 90% of electric bicycles use lead-acid batteries as power batteries. Therefore, lead-acid batteries take up a huge market share, no matter in production, sales or holding quantity, and waste quantity and have great recovery potential and, meanwhile, will face a seriously pollution prevention situation. It is thus strongly necessary to study the lead-acid batteries used in electric bicycles, adopt the latest statistical data, and select the optimal estimation model to forecast and study their waste quantity and provide scientific support for waste recovery and reclamation in China.

### 3. Research Method and Data Source

#### 3.1. Research Method

*3.1.1. Introduction to Estimation Models.* In the estimation of the waste quantity of an electronic product, the production, sales, and service life of the product are generally considered. There are mainly 7 estimation models (Simon W. et al. 2011) [48].

(1) Market supply model: the model is a method to estimate electronic waste based on product sales data and average service life. The assumption is that the sold electronic product is completely discarded at the end of its service life and can still be used before the end of its service life and the average service life of the product is relatively stable. The estimation formula of the annual waste quantity of a certain electronic product using the model is

$$Q_w = S_n \quad (1)$$

$Q_w$  is quantity of electronic waste;  $S_n$  is sales of the electronic product  $n$  years ago;  $n$  is average service life of the electronic product.

(2) Market supply A model (Yamasue E et al. 2006) [49]: on the basis of the market supply model, the model adopts the distribution value of average service life of a product. It is assumed that the product is subject to several different service lives every year and gives each service life a certain proportion. According to relevant research, the service lives

of the product are in normal distribution around the average service life. The estimation formula of electronic waste using the market supply A model is

$$Q_w = \sum_{i=0}^n S_i P_i \quad (2)$$

$Q_w$  is quantity of electronic waste;  $S_i$  is sales of the electronic product  $i$  years ago beginning this year;  $P_i$  is percentage of the electronic product with  $i$  years of service life;  $i$  is time life of the electronic product.

(3) Stanford Model (Yang and Williams 2009) [50]: the model uses the changes in the sales in a certain period of time after entering society and the social holding quantity in that period of time to calculate the quantity of electronic waste. Its calculation method is similar to that of the market supply A model, except that  $P_i$  in the market supply A model is a constant value while  $P_i$  in the Stanford Model is variable. The model assumes that every year the product that is being sold is subject to several different service lives according to its usage. The formula is

$$Q_w = \sum_{i=0}^n S_i P_i \quad (3)$$

$Q_w$  is quantity of electronic waste;  $S_i$  is sales of the electronic product  $i$  years ago beginning this year;  $P_i$  is the percentage of the electronic product with  $i$  years of service life;  $i$  is the time life of the electronic product.

(4) Carnegie Mellon Model: the model has corrected the market supply method by taking into account the disposal methods after discarding. When making forecast, it takes into account how consumers treat and handle unused electronic products. On the basis of analyzing consumers' handling of electronic waste, it has set four different handling scenarios when an electronic product is obsolete, namely, refurbishing and reselling it, laying it aside, dismantling and restoring it, and disposing it as waste, and gives each handling method a certain proportion. The Carnegie Mellon Model is suitable for large waste electrical appliances with a relatively longer service life.

(5) Time gradient model: the model starts with holding quantity, takes into account the number of home appliances entering and exiting the holding quantity statistics, and estimates waste production based on sales data and private holding quantity and industrial holding quantity level. The formula is

$$P_t = \sum_{i=t_1}^t S_n - \sum_{i=t_1}^t P_n - (H_t - H_{t_1}), \quad (t_1 < t) \quad (4)$$

$P_t$  is production of electronic waste in the  $t$  year;  $P_n$  is production of electronic waste in the  $n$  year;  $S_n$  is sales of the electronic product in the  $n$  year;  $H_t$  is social stock of electronic production in the  $t$  year;  $H_{t_1}$  is social stock of electronic production in the  $t_1$  year.

(6) "Estimation" model: It mainly adopts social holding quantity and average service life. The estimation formula is

$$Q_w = \frac{\text{Holding quantity}_{(\text{private}+\text{industrial})}}{n} \quad (5)$$

$Q_w$  is quantity of electronic waste;  $n$  is average service life of the electronic product.

(7) ICER model: the model uses the replacement rate of the estimated product to estimate its waste quantity. On the basis of the market supply model and the Stanford Model, Chinese scholars have established a prediction model based on the production of waste electronic information products in fixed and dynamic cycles.

The aforesaid models are mainly used to estimate the waste quantity of electronics and electric appliances. We suggest selecting a suitable model to estimate the quantity of waste lead-acid batteries based on the service cycle and waste characteristics of lead-acid batteries.

### 3.1.2. Selection of Prediction Model

(1) *Types and Characteristics of Lead-Acid Batteries.* Based on their field of application, lead-acid batteries can be divided into four types: starting batteries, power batteries, standby batteries, and energy storage batteries (see Table 1). The service life of lead-acid batteries varies depending on their purpose of use. Therefore, the discarding time and annual discarding quantity also vary. Taking the statistical data in 2012, for example, the production of lead-acid batteries used in electric bicycles as power batteries accounted for less than 37% of the production of the whole lead-acid battery industry. However, some scholars (Zhang H. et al. 2016) [34] did not make such a strict distinction. When estimating the quantity of waste lead-acid batteries, they adopted the data of the whole industry as the production and sales data for calculation, while only adopting the service life of a certain type of batteries for calculation, which can easily lead to deviation of estimation data.

Globally starting lead-acid batteries take up the largest proportion, which is 48%, followed by power lead-acid batteries with 28%. The proportion of standby and energy storage lead-acid batteries is 15%, and that of other lead-acid batteries is 9% (Lu L.Q. 2018) [51]. The proportions of the four types of lead-acid batteries are similar to those in the Chinese market.

In addition, in terms of waste recovery, lead-acid batteries in different application fields are faced different recovery dilemmas (see Table 1). The lead-acid batteries used as power batteries are the focus and difficulty in recovery management, while the lead-acid batteries used in electric bicycles as power batteries are inevitably the priority in research.

(2) *Selection of Prediction Model.* The aforesaid 7 models are mainly applicable to the estimation of the quantity of waste electronic and electrical products, including TVs, refrigerators, air conditioners, computers, washing machines, and mobile phones (Gao Y. N. et al. 2010) [52]. Because of the generality between battery-using products and electronic and electrical products in terms of service cycle, discarding cycle and replacement frequency, some scholars used some of the models to estimate the quantity of waste batteries, including primary batteries, lead-acid batteries, lithium batteries, and Ni-MH batteries.

TABLE 1: Types and recovery management characteristics of lead-acid batteries.

Battery type	Application field	Battery			Recovery management characteristics
		production (2012) [20] (10,000 KVAH)	Proportion	Service life (year)	
Starting battery	Used as storage batteries for starting, ignition and lighting (SLI) and so on in vehicles with an internal combustion engine such as buses, trucks and motorcycles and other vehicles	7495	52%	1.5-3	The users are widely distributed. The production is high, the single use quantity is relatively huge, and most of such batteries are replaced through distributors or maintenance service providers. The recovery management focuses are distributors and maintenance service providers.
Power battery	Used as power batteries in electric vehicles such as electric bicycles, electric special-purpose vehicles (electric sightseeing buses, golf carts, patrol cars, forklifts, etc.), low-speed electric passenger cars and hybrid electric vehicles	5320	37%	1-3	The users are especially disperse, the production is relatively large, the single use quantity is large and the replacement frequency is high, but the recovery is irregular. Illegal recovery vendors are main clients. The recovery management focuses are consumers, distributors and maintenance service providers.
Standby battery	Used as standby power in telecommunication base stations, UPS, emergency lighting power and other standby power	1601	11%	3-6	The batteries have built-in GPS, the production is low, but the single use quantity is huge. They can be recycled through proper channels and are easy to manage. Recovery management focuses are distributors and maintenance service providers.
Energy storage battery	equipment, wind turbines and other renewable energy resources				

TABLE 2: Sales-output ratio of electric bicycles in 2015-2017 (unit: 10,000 sets.).

Year	Output of electric bicycles	Sales of electric bicycles	Sales-output ratio
2015	3257	3398	104.3%
2016	3215	2600	80.9%
2017	3113	3200	102.8%

Source: National Bureau of Statistics (www.stats.gov.cn) and www.chyxx.com.

TABLE 3: Sales-output ratio of electric tricycles in 2014-2017 (unit: 10,000 sets).

Year	Output of electric tricycles	Sales of electric tricycles	Sales-output ratio
2014	1122	898	80%
2015	1163	980	84.2%
2016	1153	930	80.7%
2017	850	800	94.1%

Source: www.chyxx.com.

In view of the characteristics of lead-acid batteries, especially the power batteries used in electric bicycles as the subject of study of the thesis, we have selected the “market supply A model” and the “Stanford Model” to forecast the waste quantity of power batteries used in electric bicycles.

### 3.2. Source of Data

3.2.1. Calculation of Sales of Lead-Acid Batteries Used in Electric Bicycles as Power Batteries. Both the “market supply A model” and the “Stanford Model” adopt product sales ( $S_i$ ), product service life ( $i$ ), and service life distribution ( $P_i$ ) as the calculation basis to estimate waste quantity.

In this thesis, we have adopted the annual production of electric bicycles and electric tricycles as the basis data to estimate the waste quantity of power lead-acid batteries used in electric bicycles. We have collected and sorted out the overall data of the electric bicycle industry, and based on the industry structure, we have obtained the annual net increment of electric bicycles (including electric tricycles) using lead-acid batteries as power using formula (6). Based on the annual net increment of lead-acid battery powered electric bicycles (according to national regulations, a common electric bicycle generally uses four 48V12ah or four 48V20ah batteries), we have then obtained the annual increment of lead-acid batteries through calculation and have used it as the annual newly increased sales of power lead-acid batteries used in electric bicycles.

$$\begin{aligned}
 & \underline{\text{Annual net increment of}} \text{ lead-acid battery powered electric bicycles (including tricycles)} \\
 & = \text{production of electric bicycles} \\
 & + \text{production of electric tricycles} - \text{production of lithium battery powered bicycles} \\
 & + \text{import volume of electric bicycles} - \text{export volume of electric bicycles}
 \end{aligned} \tag{6}$$

(1) Calculation of Sales-Output Ratios of Three Types of Electric Bicycles. According to the acquired data, calculation, and analysis, the average sales-output ratio of electric bicycles in 2015-2017 (see Table 2) is 96%, that of electric tricycles in 2014-2017 (see Table 3) is 85%, and that of lithium battery powered bicycles (see Table 4) is 87%. The three types of electric bicycles basically ensured dynamic balance between sales and output. Therefore, the output of the three types of electric bicycles is deemed equal to their sales in this thesis.

(2) Calculation of Import Volume of Electric Bicycles. Due to limited data and influence of domestic policy and international market, the import and export data of electric bicycles have been fluctuating widely over the years. To ensure the

accuracy and coherence of estimation data, regarding the missing import data of electric bicycles in 2016 and 2017, we have excluded the data in abnormal years (2012, 2013, and 2014) and adopted the average value of the remaining data as the import volume of electric bicycles in 2016 and 2017 (see Table 5).

(3) Calculation of Output of Lithium Battery Powered Bicycles. Due to limited data, the output of lithium battery powered bicycles in 2017 is missing. Through the analysis of the data on lithium battery powered bicycles over the years, we have found that the proportion of the output of lithium battery powered bicycles in the bicycle industry is on a rising trend since 2006. The fitted equation on the proportion of lithium battery powered bicycles in electric bicycles in 2006-2016 is

TABLE 4: Sales-output ratio of lithium battery powered bicycles in 2006-2015 (unit: 10,000 sets).

Year	Output of lithium battery powered bicycles	Sales of lithium battery powered bicycles	Sales-output ratio
2006	5	5	100%
2007	13.5	8	59.2%
2008	22.3	12	53.8%
2009	29.84	17	57%
2010	40.2	25	62%
2011	65	70	107%
2012	149.9	160	106%
2013	231.1	231.1	100%
2014	301	350	116%
2015	292	300	106%

Source: www.evtimes.cn/, www.escn.com.cn/, and www.chyxx.com/industry/201702/493958.html.

TABLE 5: Import and export volume of electric bicycles (unit: 10,000 sets.).

Year	Export volume of electric bicycles	Import volume of electric bicycles
2009	122.4	457
2010	163.4	488
2011	141.1	435
2012	128.9	835
2013	151.9	2456
2014	211.3	886
2015	649.4	441
2016	133.9	(455)
2017	730.1	(455)

Source: General Administration of Customs (www.customs.gov.cn/), www.chinaidr.com/news/2017-02/110769.html, and www.chinaidr.com/news/2017-02/110770.html.

$$y = -0.00022x^3 + 0.00477x^2 - 0.01925x + 0.02324 \tag{7}$$

$$R^2 = 0.96559$$

With the equation, we have obtained that the proportion in 2017 is 9.90%, which is highly consistent with industry experts' forecast on the proportion of lithium battery powered bicycles in 2017, which is that "seen from data performance, the market share of lithium battery powered bicycles is increasing year by year. Even though official data is currently unavailable, we can be sure that the market share of the sales of lithium battery powered bicycles in 2017 is around one tenth" [53]. Based on that, we have obtained the output of lithium battery powered bicycles in 2017, which is 3.0806 million (see Table 6).

(4) *Calculation of Annual Net Increment of Lead-Acid Battery Powered Bicycles (Including Tricycles).* The annual net increment of lead-acid battery powered bicycles (including tricycles) in 2000-2017 is obtained using formula (6) (see Table 7).

(5) *Forecast of Annual Net Increment of Lead-Acid Battery Powered Bicycles (Including Tricycles) in 2018-2022.* Based on the data change in the annual net increment of lead-acid battery powered bicycles in 2000-2017, we have analyzed its characteristics and variation trend and selected the annual

net increment of lead-acid batteries in the last five years (2013-2017). We have adopted the fitted equation:

$$y = (1,973.54328) \ln(x) + 6,598.11029 \tag{8}$$

$$R^2 = 0.90909$$

To estimate the annual net increment of lead-acid battery powered bicycles (including tricycles) in 2018-2022 (see Table 8), based on the fact that the most common electric bicycle uses four 48V12ah storage batteries and the weight of each storage battery is about 4.3±0.2kg, we have forecasted the annual net increment of lead-acid batteries used in electric bicycles in 2000-2022 (see Table 9).

3.2.2. *Estimation with "Market Supply A Model".* During our visit and interview, an owner who has been selling electric bicycles (including old for new business of storage batteries) for more than a decade said that even though the motives to buy an electric bicycle (household use, delivering goods, and express delivery) and the using frequency and time (the frequency is 2-6 times a day when using it to go to and get off work or take children to school or pick them up from school, and the using time is about 1-1.5 hours; for delivery of goods and parcels, the frequency is over 30 times a day, and the using time is more than 4 hours) vary, most storage batteries (lead-acid batteries) used in electric bicycles need to be replaced

TABLE 6: Output of lithium battery powered bicycles (unit: 10,000 sets.).

Year	Output of electric bicycles	Output of lithium battery powered bicycles	Proportion of lithium battery powered bicycles in electric bicycles
2006	1950	5	0.26%
2007	2138	13.5	0.63%
2008	2188	22.3	1.02%
2009	2369	29.84	1.26%
2010	2954	40.2	1.36%
2011	3096	65	2.10%
2012	3505	149.9	4.28%
2013	3695	231.1	6.25%
2014	3551	301	8.48%
2015	3257	292	8.97%
2016	3215	282	8.77%
2017	3113	(308.06)	(9.90%)

Source: [www.evtimes.cn/](http://www.evtimes.cn/) and [www.escn.com.cn/](http://www.escn.com.cn/).

TABLE 7: Annual net increment of lead-acid battery powered bicycles (including tricycles) in 2000-2017 (unit: 10,000 sets).

Year	Output of electric bicycles	Output of electric tricycles	Output of lithium battery powered bicycles	Export volume of electric bicycles	Import volume of electric bicycles	Annual net increment of lead-acid battery powered bicycles (including tricycles)
2000	29	-	-	-	-	29.00
2001	59	-	-	-	-	59.00
2002	159	-	-	-	-	159.00
2003	400	-	-	-	-	400.00
2004	676	-	-	-	-	676.00
2005	1211	-	-	-	-	1211.00
2006	1950	-	5	-	-	1945.00
2007	2138	-	13.5	-	-	2124.50
2008	2188	-	22.3	-	-	2165.70
2009	2369	603	29.84	122.4	457	3276.77
2010	2954	678	40.2	163.4	488	3916.36
2011	3096	761	65	141.1	435	4085.92
2012	3505	892	149.9	128.9	835	4953.22
2013	3695	1021	231.1	151.9	2456	6788.98
2014	3551	1122	301	211.3	886	5046.71
2015	3257	1163	292	649.4	441	3919.60
2016	3215	1153	282	133.9	(455)	4407.10
2017	3113	850	(308.06)	730.1	(455)	3379.84

Source: National Bureau of Statistics ([www.stats.gov.cn/](http://www.stats.gov.cn/)).

in 2 years if they are properly used. Through the survey of the families using electric bicycles in the surroundings, we have found that most families buy electric bicycles as a means of transportation to go to and get off work and send their children to school or pick them up by the way they generally

use them for no more than 30 minutes in a single time and 2-3 times a day and replace batteries in 2-3 years.

Take a 48V12ah lead-acid battery for example. According to the product information when the storage battery is sold, the cycle index is 400-600. If it is charged once a day, which

TABLE 8: Annual net increment of lead-acid battery powered bicycles (including tricycles) in 2018-2022 (unit: 10,000 sets.).

Year	Annual net increment of lead-acid battery powered bicycles (including tricycles)
2013	6788.98
2014	5046.71
2015	3919.60
2016	4407.10
2017	3379.84
<b>2018</b>	<b>3062.00</b>
<b>2019</b>	<b>2757.77</b>
<b>2020</b>	<b>2494.24</b>
<b>2021</b>	<b>2261.79</b>
<b>2022</b>	<b>2053.86</b>

means 1 cycle a day, it can be used for 1-1.5 years; if it is charged once in two days, it can be used for 2-3.5 years; if it is charged once in three days, it can be used for 3-5 years.

According to industry standard, the service life of the lead-acid batteries used in electric bicycles should ensure 350 times of discharging based on 70% nominal capacity. However, in actual use, the cycle life of some batteries is as high as 600 cycles, the total capacity released is 6151ah, and the corresponding accumulative mileage is about 24,600km. Based on that, it can be used for over 2 years (Guo Z.Q. 2003) [54].

Therefore, through social research and literature reading, we have found that the average service life of the lead-acid batteries powering electric bicycles is 2 years and the lead-acid batteries with a service life under 1 year or above 4 years are in the minority; that is, most lead-acid batteries are discarded after used for 2-3 years. We have selected the average value  $X = 2$  of the service life of lead-acid batteries as the  $\mu$  value and obtained  $\sigma = 63$ ,  $\sigma^2 = 2/3$  using the equation  $\sigma = Xi + \bar{X}2/n$ . After the service life of lead-acid batteries complies with the normal distribution  $N(2, 2/3)$ , we can obtain the service life distribution proportion of lead-acid batteries by querying the normal distribution table: the lead-acid batteries with a service life  $< 1$  year account for 11.12%; those with a service life of 1-2 years account for 38.88%; those with a service life of 2-3 years account for 38.88%, and those with a service life  $> 3$  years account for 11.12%; that is,  $P_1 = 11.12\%$ ,  $P_2 = 38.88\%$ ,  $P_3 = 38.88\%$ , and  $P_4 = 11.12\%$ .

After putting  $S_i$ ,  $P_i$  in the market supply A model, respectively, we have obtained the waste quantity of lead-acid batteries used in electric bicycles in 2000-2022 (see Table 10).

### 3.2.3. Estimation Using "Stanford Model"

(1) *Model Assumption.* The main difference between the Stanford Model and the market supply A model lies in that the Stanford Model mainly adopts the service life of products and service life distribution proportion to estimate waste quantity. That is to say, the  $P_i$  (percentage of waste) in the market supply A model is a constant value while the  $P_i$  (percentage

of waste) in the Stanford Model is variable. According to the literature (Liu M. L. et al. 2015) [55], the model is based on the following assumptions.

First, the lead-acid battery used in the eclectic bicycle is always being used till the end of its service life, and it is completely discarded after its service life ends. Second, the service life of the annual net incremental lead-acid batteries used in electric bicycles is in normal distribution centering on the average service life by different proportions. Third, the service life of the annual net incremental lead-acid batteries used in electric bicycles remains unchanged.

To the lead-acid batteries studied in this thesis, we have assumed that the probability of discarding of the lead-acid batteries after  $n_1, n_2, \dots, n_m$  years is  $p_1, p_2, \dots, p_m$  and we have used the annual net increment of lead-acid batteries used in electric bicycles obtained through calculation to forecast the future waste quantity. If the longest service life of the lead-acid batteries used in electric bicycles is  $M$ , the waste quantity is  $S_t$  in the  $t$  year and the scrap rate is  $P_k$  after  $k$  years, then the waste quantity  $Q_j$  in the year  $j$  is

$$Q_j = \sum_k^M S_{j-k} P_k \quad (9)$$

(2) *Model Calculation.* On the basis of social survey, literature search and study of product use information, and the use characteristics of the lead-acid batteries used in electric bicycles, we have determined the service life, discarding time, and scrap rate of such lead-acid batteries every year (see Table 11).

After putting the data in Table 11 into the Stanford Model for calculation, we have obtained the waste quantity of lead-acid batteries used in electric bicycles in 2000-2022 (see Table 12).

## 4. Analysis of Predictions

In this thesis we have adopted the "market supply A model" and the "Stanford Model" to predict the annual waste quantity of lead-acid batteries used in electric bicycles in 2000-2022. As shown in Tables 10 and 12, the use quantity of lead-acid batteries changes with the sales or annual net increment of electric bicycles, while the waste quantity of lead-acid batteries is closely related to their service life. By analyzing the sales of lead-acid battery powered electric bicycles in 2000-2017, we have found that the sales of lead-acid batteries increased steadily as the sales of electric bicycles increased in 2000-2004, saw explosive growth in 2005-2013, reached the peak in 2013, and steadily fell down since 2013 till 2017 in which the sales are close to that in 2010. The forecast of the sales of lead-acid battery powered electric bicycles also shows that the downward trend will last till 2022. By comparing and analyzing the results of the two prediction models, we have found that because the  $P_i$  in the Stanford Model is variable and the model has higher requirements on data and is more comprehensive, the quantity of waste lead-acid batteries estimated with the model is slightly smaller than that estimated using the market supply A model; however, the

TABLE 9: Annual net increment of lead-acid batteries used in electric bicycles in 2000-2022.

Year	Annual net increment of lead-acid battery powered bicycles (including tricycles) (unit: 10,000 sets.)	Annual net increment of lead-acid batteries used in electric bicycles		
		(10,000 sets.)	(10,000 t)	(10,000 KVAH)
2000	29	116.0	0.5	66.8
2001	59	236.0	1.0	135.9
2002	159	636.0	2.7	366.3
2003	400	1600.0	6.9	921.6
2004	676	2704.0	11.6	1557.5
2005	1211	4844.0	20.8	2790.1
2006	1945	7780.0	33.5	4481.3
2007	2124.5	8498.0	36.5	4894.8
2008	2165.7	8662.8	37.3	4989.8
2009	3276.8	13107.1	56.4	7549.7
2010	3916.4	15665.4	67.4	9023.3
2011	4085.9	16343.7	70.3	9414.0
2012	4953.2	19812.9	85.2	11412.2
2013	6789.0	27155.9	116.8	15641.8
2014	5046.7	20186.8	86.8	11627.6
2015	3919.6	15678.4	67.4	9030.8
2016	4407.1	17628.4	75.8	10154.0
2017	3379.8	13519.4	58.1	7787.1
2018	3062.0	12248.0	52.7	7054.8
2019	2757.8	11031.1	47.4	6353.9
2020	2494.2	9977.0	42.9	5746.7
2021	2261.8	9047.2	38.9	5211.2
2022	2053.9	8215.4	35.3	4732.1

predictions of the two models on the overall trend of the waste are consistent.

Moreover, due to the issuance and implementation of the mandatory national standard *Technical Code for Safety of Electric Bicycles* (GB17761-2018) since May 15, 2018, lead-acid battery powered electric bicycles will certainly be faced with either situation: to innovate own technology or to be replaced by lithium battery powered bicycles on a large scale. If lead-acid batteries are replaced by lithium batteries on a large scale, the sales of new electric bicycles powered by lead-acid batteries will reduce significantly, while the quantity of waste lead-acid batteries will increase sharply. Therefore, the lead-acid battery industry will be faced with industry transformation or technology upgrading, and recovery and processing enterprises of waste lead-acid batteries will be faced with new opportunities.

## 5. Analysis of Recovery Potential of Waste Lead-Acid Batteries

The average service life of the lead-acid batteries used in electric bicycles is about 2 years. Generally each electric bicycle is equipped with four 48V12ah lead-acid batteries. According to the prediction above, the waste quantity of lead-acid batteries used in electric bicycles only in 2017 will be nearly 200 million, and the weight 850,000 tons. Because

in each lead-acid battery, 73.26% is lead and 11.46% is pure sulfuric acid, the 850,000-ton waste lead-acid batteries will contain 620,000 tons of lead and 9.74 million tons of pure sulfuric acid.

We have studied the list of raw materials needed in the production phase of lead-acid batteries and the list of treatment links of waste lead-acid batteries based on the life-cycle assessment (LCA) method. The LCA method can be used to not only assess the environmental impacts and the consumption of environmental resources in the production, sales, transportation and use, and disposal and treatment of lead-acid batteries but also forecast the recovery potential of lead-acid batteries in the ideal state.

The materials used to manufacture lead-acid batteries mainly include lead (73.26%), pure sulfuric acid (11.46%), plastic (12.09%), antimony, arsenic, tin (2.51%), rubber (0.35%), copper (0.35%), etc. (see Table 14). Based on Table 10 ("Market supply A model" based annual waste quantity of lead-acid batteries used in electric bicycles in 2000-2022), Table 12 ("Stanford Model" based annual waste quantity of lead-acid batteries used in electric bicycles in 2000-2022), Table 13 (List of raw materials of lead-acid batteries), and Table 14 (Recovery list of disposal link of lead-acid batteries), we have obtained the recovered resource amount of lead-acid batteries used in electric bicycles as power batteries in 2000-2022 (see Tables 15 and 16) through calculation.

TABLE 10: “Market supply A model” based annual waste quantity of lead-acid batteries used in electric bicycles in 2000-2022.

Year	Annual waste quantity of lead-acid batteries used in electric bicycles		
	(10,000 sets.)	(10,000 t)	(10,000 KVAH)
2000	-	-	-
2001	12.9	0.1	7.4
2002	71.3	0.3	41.1
2003	207.6	0.9	119.6
2004	529.9	2.3	305.2
2005	1196.3	5.1	689.1
2006	2282.8	9.8	1314.9
2007	3977.7	17.1	2291.2
2008	6153.9	26.5	3544.6
2009	7830.8	33.7	4510.6
2010	8994.8	38.7	5181.0
2011	11151.1	47.9	6423.0
2012	13967.5	60.1	8045.3
2013	16105.9	69.3	9277.0
2014	18819.4	80.9	10840.0
2015	22323.7	96.0	12858.4
2016	22353.5	96.1	12875.6
2017	18924.4	81.4	10900.5
2018	16697.8	71.8	9617.9
2019	15215.7	65.4	8764.2
2020	13205.3	56.8	7606.2
2021	11663.7	50.2	6718.3
2022	10536.0	45.3	6068.7

Recovery potential is the ratio between the recoverable mass of each material of waste lead-acid batteries and the mass of the raw material contained in lead-acid batteries. It can be used to access the recovery value of waste lead-acid batteries.

The recovery potential of each material of lead-acid batteries can be obtained based on the data in Tables 13 and 14: the recovery potential of lead is 86.9%, that of pure sulfuric acid is 85%, and that of plastic is 87.7%.

Taking 2017, for example, the waste quantity of lead-acid batteries used in electric bicycles is about 110 million KVAH, which contains 3,000 tons of lead. As the recovery potential of lead is 86.9%, recoverable lead is 2,607 tons. According to the international lead price of USD2, 600 per ton in 2017, the economic value is over USD6.7 million. The lead-acid battery recovery industry thus has enormous economic benefit.

## 6. Conclusions

In the thesis, through statistical analysis of the data in the electric bicycle industry, with the output of electric bicycles as the entry point, we have collected import and export data of electric tricycles, lithium battery powered bicycles, and electric bicycles. And based on the characteristics of the electric bicycle industry, as there are currently two main

power batteries used in the electric bicycle industry, lead-acid batteries and lithium batteries, we have eliminated the output of lithium battery powered electric bicycles and taken into account the import and export value of electric bicycles and have finally obtained the annual net increment of lead-acid battery powered electric bicycles through calculation. With the annual net increment of lead-acid battery powered electric bicycles as the basic data to study the lead-acid batteries used in electric bicycles, we have then obtained the annual net increment of lead-acid batteries used in electric bicycles and forecasted increment in 2018-2022. Then, with the “market supply A model” and the “Stanford Model”, we have estimated the quantity of discarded lead-acid batteries used in electric bicycles between 2000 and 2022, respectively, and based on the proportion of raw materials contained in lead-acid batteries and the proportion between reclaimed and discarded lead-acid batteries, we have estimated the recovery potential of waste lead-acid batteries in 2000-2022. We estimate that the lead-acid batteries used in electric bicycles only have great recovery potential and there are abundant potential resources for recovery.

(1) Based on the forecast of the output of lead-acid batteries in 2018-2022, the output of lead-acid batteries tends to fall significantly. As shown by the data, due to restrictions of economic environment and environmental protection policy, the electric bicycles using lead-acid batteries as power

TABLE 11: Service life distribution proportion of lead-acid batteries powering electric bicycles.

Year	Annual net increment of lead-acid batteries used in electric bicycles (10,000 sets)	Average service life (year)	Service life distribution (%) of lead-acid batteries			
			$P_4$	$P_3$	$P_2$	$P_1$
2000	116.0	2.3	10%	30%	40%	20%
2001	236.0	2.3	10%	30%	40%	20%
2002	636.0	2.3	10%	30%	40%	20%
2003	1600.0	2.3	10%	30%	40%	20%
2004	2704.0	2.3	10%	30%	40%	20%
2005	4844.0	2.3	10%	30%	40%	20%
2006	7780.0	2.5	15%	35%	35%	15%
2007	8498.0	2.5	15%	35%	35%	15%
2008	8662.8	2.5	15%	35%	35%	15%
2009	13107.1	2.5	15%	35%	35%	15%
2010	15665.4	2.5	15%	35%	35%	15%
2011	16343.7	2.5	15%	35%	35%	15%
2012	19812.9	2.7	20%	40%	30%	10%
2013	27155.9	2.7	20%	40%	30%	10%
2014	20186.8	2.7	20%	40%	30%	10%
2015	15678.4	2.7	20%	40%	30%	10%
2016	17628.4	2.7	20%	40%	30%	10%
2017	13519.4	2.7	20%	40%	30%	10%
2018	12248.0	2.7	20%	40%	30%	10%
2019	11031.1	2.7	20%	40%	30%	10%
2020	9977.0	2.7	20%	40%	30%	10%
2021	9047.2	2.7	20%	40%	30%	10%
2022	8215.4	2.7	20%	40%	30%	10%

will be faced with a big risk of elimination, while the lithium battery powered bicycles that currently only account for 10% of the output of the electric bicycle industry show great development momentum. The storage battery manufacturers that mainly produce lead-acid batteries supplied to electric bicycles will face transformation and upgrading.

(2) The implementation of the mandatory national standard *Technical Code for Safety of Electric Bicycles* and the rise of lithium battery powered electric bicycles will jointly force lead-acid battery powered electric bicycles to make technological innovation. Lithium batteries will gradually replace lead-acid batteries and become main power source of electric bicycles, and a large number of waste lead-acid batteries will (before end of service life) be eliminated and discarded. How to properly recycle and dispose waste lead-acid batteries will be a big challenge.

(3) According to the research data presented in the thesis, waste lead-acid batteries have great economic benefit and environmental protection benefit. However, the recovery industry is currently faced with the dilemma of how to get rid of the recovery of lead-acid batteries used in electric bicycles by “irregulars” such as petty dealers, eliminate illegal smelting of secondary lead and improper disposal of waste sulfuric acid, improve the recovery quantity by regular secondary enterprises through regular channels, increase circulation of lead resources, and ensure normal development of secondary

lead enterprises. Those will be difficulties in government management. “Extended Producer Responsibility” (EPR) [56] will play one of the most important roles to solve those difficulties.

## 7. Limitations and Prospects

Limited by data, the calculation results in the thesis are not obtained directly using the data of lead-acid batteries, so there is still some uncertainty in the waste quantity of lead-acid batteries and the forecast of their recovery potential, and the deviation between the waste quantity of lead-acid batteries forecasted using models and the actual generation data is unavoidable. Moreover, the service life of lead-acid batteries is critical to the forecast of waste quantity. However, their service life is affected by many factors such as whether the charger is matched, whether the user uses the battery correctly and the quality of the battery itself. In future research, we will further improve data accuracy and adjust the adaptability of models.

## Data Availability

All the data used to support the findings of this study are included in our manuscript and can be accessed freely from the references.

TABLE 12: “Stanford Model” based annual waste quantity of lead-acid batteries used in electric bicycles in 2000-2022.

Year	Annul waste quantity of lead-acid batteries used in electric bicycles		
	(10,000 sets)	(10,000 t)	(10,000 KVAH)
2000	-	-	-
2001	-	-	-
2002	-	-	-
2003	-	-	-
2004	656.8	2.8	378.3
2005	1395.2	6.0	803.6
2006	2594.0	11.2	1494.1
2007	4075.8	17.5	2347.7
2008	5721.3	24.6	3295.5
2009	7481.1	32.2	4309.1
2010	9139.3	39.3	5264.3
2011	11244.0	48.3	6476.5
2012	13821.4	59.4	7961.1
2013	15150.6	65.1	8726.7
2014	16729.6	71.9	9636.2
2015	20542.2	88.3	11832.3
2016	22448.8	96.5	12930.5
2017	19972.3	85.9	11504.0
2018	16949.2	72.9	9762.7
2019	15467.6	66.5	8909.4
2020	13710.9	59.0	7897.5
2021	11910.1	51.2	6860.2
2022	10759.8	46.3	6197.7

TABLE 13: List of raw materials of lead-acid batteries (Yu Y. J. et al. 2010 [21]; Zhang H. et al. 2013 [22]).

Name of raw material	Mass of raw material contained in 1000 KVAH lead-acid batteries (kg)	Ratio of mass of raw material in total mass
Lead	26.85	73.26%
Pure sulfuric acid	4.20	11.46%
Plastic	4.43	12.09%
Antimony, arsenic, tin	0.92	2.51%
Rubber	0.13	0.35%
Copper	0.12	0.35%

TABLE 14: Recovery list of disposal link of lead-acid batteries (Zhang H. Wang et al. 2013) [23].

Type	Name	Quantity (1,000kg)	Unit
Recovered material	Lead	636.91	Kg
	Total sulfur (100% sulfuric acid)	97.41	Kg
	Plastic recovery quantity	106.06	Kg
Energy	Coal	138.23	Kg
	Electric power	21.40	Kg
Emission	Lead fume	12.88	Kg
	Residue-containing lead	10.55	Kg
	SO <sub>2</sub> emission	267.07	Kg

TABLE 15: Recovery quantity of lead-acid batteries used in electric bicycles as power batteries in 2000-2022 (market supply A model).

Year	Waste quantity (10,000 KVAH)	Annual waste quantity and recovery quantity of lead-acid batteries used in electric bicycles						
		Lead (t)	Pure sulfuric acid (t)	Plastic (t)	Antimony, arsenic, tin (t)	Rubber (t)	Copper (t)	
2000	-	2.0	0.3121	0.3291	0.0684	0.0097	0.0089	
2001	7.4	11.0	1.7	1.8	0.3781	0.0534	0.0493	
2002	41.1	32.1	5.0	5.3	1.1	0.1554	0.1435	
2003	119.6	81.9	12.8	13.5	2.8	0.3968	0.3662	
2004	305.2	185.0	28.9	30.5	6.3	0.8958	0.8269	
2005	689.1	353.0	55.2	58.2	12.1	1.7	1.6	
2006	1314.9	615.2	96.2	101.5	21.1	3.0	2.7	
2007	2291.2	951.7	148.9	157.0	32.6	4.6	4.3	
2008	3544.6	1211.1	189.4	199.8	41.5	5.9	5.4	
2009	4510.6	1391.1	217.6	229.5	47.7	6.7	6.2	
2010	5181.0	1724.6	269.8	284.5	59.1	8.3	7.7	
2011	6423.0	2160.2	337.9	356.4	74.0	10.5	9.7	
2012	8045.3	2490.9	389.6	411.0	85.3	12.1	11.1	
2013	9277.0	2910.5	455.3	480.2	99.7	14.1	13.0	
2014	10840.0	3452.5	540.1	569.6	118.3	16.7	15.4	
2015	12858.4	3457.1	540.8	570.4	118.5	16.7	15.5	
2016	12875.6	2926.8	457.8	482.9	100.3	14.2	13.1	
2017	10900.5	2582.4	404.0	426.1	88.5	12.5	11.5	
2018	9617.9	2353.2	368.1	388.3	80.6	11.4	10.5	
2019	8764.2	2042.3	319.5	337.0	70.0	9.9	9.1	
2020	7606.2	1803.9	282.2	297.6	61.8	8.7	8.1	
2021	6718.3	1629.4	254.9	268.8	55.8	7.9	7.3	
2022	6068.7							

TABLE 16: Recovery quantity of lead-acid batteries used in electric bicycles as power batteries in 2000–2022 (Stanford Model).

Year	Waste quantity (10,000 KVAH)	Annual waste quantity and recovery quantity of lead-acid batteries used in electric bicycles						
		Lead (t)	Pure sulfuric acid (t)	Plastic (t)	Antimony, arsenic, tin (t)	Rubber (t)	Copper (t)	
2000	-	-	-	-	-	-	-	-
2001	-	-	-	-	-	-	-	-
2002	-	-	-	-	-	-	-	-
2003	-	-	-	-	-	-	-	-
2004	378.3	101.6	15.9	16.8	3.5	0.4918	0.4540	
2005	803.6	215.8	33.8	35.6	7.4	1.04	0.9644	
2006	1494.1	401.2	62.8	66.2	13.7	1.9	1.8	
2007	2347.7	630.3	98.6	104.0	21.6	3.1	2.8	
2008	3295.5	884.8	138.4	146.0	30.3	4.3	4.0	
2009	4309.1	1157.0	181.0	190.9	39.6	5.6	5.2	
2010	5264.3	1413.5	221.1	233.2	48.4	6.8	6.3	
2011	6476.5	1738.9	272.0	286.9	59.6	8.4	7.8	
2012	7961.1	2137.6	334.4	352.7	73.2	10.3	9.6	
2013	8726.7	2343.1	366.5	386.6	80.3	11.3	10.5	
2014	9636.2	2587.3	404.7	426.9	88.7	12.5	11.6	
2015	11832.3	3177.0	497.0	524.2	108.9	15.4	14.2	
2016	12930.5	3471.8	543.1	572.8	119.0	16.8	15.5	
2017	11504.0	3088.8	483.2	509.6	105.8	15.0	13.8	
2018	9762.7	2621.3	410.0	432.5	89.8	12.7	11.7	
2019	8909.4	2392.2	374.2	394.7	82.0	11.6	10.7	
2020	7897.5	2120.5	331.7	349.9	72.7	10.3	9.5	
2021	6860.2	1842.0	288.1	303.9	63.1	8.9	8.2	
2022	6197.7	1664.1	260.3	274.6	57.0	8.1	7.4	

## 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 a major project financed by the National Social Science Fund of China (Approval no. 15ZDC030) and a key project financed by the National Social Science Fund of China (Approval no. 12AZD104).

## References

- [1] H. T. Zhang, "White Book on Development of China Electric Vehicle Industry 2017," <http://www.zol.com.cn>, 2017 & Table 6 Output of lithium battery powered bicycles.
- [2] Z. M. Zhang, "Current situation of waste lead-acid battery recovery industry in developed countries," *World Nonferrous Metals*, no. 11, pp. 80-81, 2008.
- [3] <https://www.mining120.com/tech/show-htm-itemid-13366.html>.
- [4] H. M. Wang and Q. Liu, "Current situation and macro-management research of disposal of waste lead-acid batteries in china," in *Proceedings of the 2012 Annual Academic Conference of the Chinese Society of Environmental Sciences*, pp. 2441-2443, 2012.
- [5] <http://www.escn.com.cn/news/show-540299.html>.
- [6] X. H. Ye, "Suggestions on strengthening recovery management," 2018, <http://www.sohu.com>.
- [7] <http://www.chyxx.com>.
- [8] G. Calabrese, "White book on development of china electric vehicle industry 2017," 2017, <http://www.zol.com.cn>.
- [9] X. Z. Li, "Yuguang Model' of recovery of waste lead-acid batteries under new situation," *China Nonferrous Metals*, no. 8, pp. 35-39, 2016.
- [10] Y. G. Ma, "Recovery of waste lead-acid batteries and production of secondary lead in China," *Chinese Journal of Power Sources*, no. 6, pp. 165-168, 2000.
- [11] "China becomes the largest lead battery market; waste lead battery recovery market has great development potential," *The World of Power Supply*, no. 7, p. 4, 2015.
- [12] "Waste lead batteries increase; recovery industry enters in-depth adjustment period," 2017, <http://www.hbzhan.com>.
- [13] "China Lead and Zinc," no. 5, pp. 40-42, 2017.
- [14] "Present situation of lead supply and demand in China in 2016 and analysis and forecast of market growth and prospect," 2016, <http://www.chyxx.com>.
- [15] S. L. Li, "Main problems in present secondary lead industry," 2018, <http://www.chinalbaogao.com>.
- [16] Y. G. Ma, "Recovery of waste lead-acid batteries and production of secondary lead in China," *Chinese Journal of Power Sources*, no. 6, pp. 165-167, 2000.
- [17] "Proportion of output of secondary lead in primary lead is expected to exceed 40% in 2015," 2014, <http://www.askci.com>.
- [18] "Current situation and development of secondary lead resource recovery technology in China," 2013, <http://www.china-mcc.com>.
- [19] X. Xue, "Yuguang model: chinese way of green recovery of waste storage batteries," *China Nonferrous Metals*, vol. 8, no. 8, pp. 29-34, 2016.
- [20] L. F. Wang, "Lead-acid battery market application and future prospects in China," *Chinese LABAT Man*, vol. 51, no. 4, pp. 171-175, 2014.
- [21] Y. J. Yu and Y. Chen, "Life-cycle assessment of typical secondary batteries," *Environmental Pollution & Control*, vol. 10, no. 10, pp. 15-18, 2010.
- [22] H. Zhang, H. M. Wang, C. L. Ma, and Y. Xu, "Life-cycle assessment of lead-acid batteries," *Chinese Journal of Environmental Management*, no. 3, p. 44, 2013.
- [23] H. Zhang, H. M. Wang, C. L. Ma, and Y. Xu, "Life-cycle Assessment of Lead-acid Batteries," *Chinese Journal of Environmental Management*, no. 3, p. 46, 2013.
- [24] "Recovery of waste lead batteries in China," 2017, <http://www.hbzhan.com>.
- [25] X. L. Lv, "Current development situation of global lead-acid battery market and future market trend," in *Proceedings of the 2013 Sino Lead-acid Battery Summit*, 2013.
- [26] D. Pavlov, *Lead-acid Batteries: Science and Technology*, X. C. Duan and S. Yuan, Eds., vol. 5, China Machine Press, 1st, part IV edition, 2015.
- [27] D. Pavlov, *Lead-acid Batteries: Science and Technology*, vol. 5, China Machine Press, 1st, part VII edition, 2015.
- [28] S. R. Zhu, *Battery Manual*, Tianjin University Press, Tianjin, China, 2002.
- [29] X. L. Liu, J. X. Yang, and R. S. Wang, "Estimation of generation of main electronic waste in China," *China Population Resources and Environment*, vol. 15, no. 5, pp. 113-117, 2005.
- [30] X. H. Liang, G. M. Li, W. Z. He, and J. W. Huang, "Prediction of quantity of waste electronic products in China," *Environmental Pollution & Control*, vol. 31, no. 7, pp. 82-91, 2009.
- [31] H. X. Tang, "Prediction of generation of waste electrical and electronic products in China and research on recovery capacity," *Chinese Journal of Environmental Management*, no. 3, pp. 4-7, 2011.
- [32] X. Tong, Y. F. Cai, and L. Yan, "Estimation method to evaluate generation of electronic waste based on recovery data of 'Old for New Home Appliances'," *Ecological Economy*, no. 7, pp. 38-42, 2013.
- [33] Q. Wang, Y. Yang, and H. Y. Ma, "Research on prediction model of resource potential of waste electronic products," *China Population Resources and Environment*, vol. 24, no. 11, pp. 147-153, 2014 (Chinese).
- [34] H. Zhang, H. M. Wang, C. Q. Xi, and Z. Zhou, "Estimation and analysis of generation of waste batteries in China," *China Population Resources and Environment*, vol. 26, no. 5, pp. 98-101, 2016.
- [35] V. Presburger Ulnikovic, M. Vukic, and R. Nikolic, "Assessment of vessel-generated waste quantities on the inland waterways of the Republic of Serbia," *Journal of Environmental Management*, vol. 97, no. 1, pp. 97-101, 2012.
- [36] B. Vivekananda and A. K. Nema, "Forecasting of solid waste quantity and composition: A multilinear regression and system dynamics approach," *International Journal of Environment and Waste Management*, vol. 13, no. 2, pp. 179-198, 2014.
- [37] R. Intharathirat, P. Abdul Salam, S. Kumar, and A. Untong, "Forecasting of municipal solid waste quantity in a developing country using multivariate grey models," *Waste Management*, vol. 39, pp. 3-14, 2015.
- [38] X. Li, X. D. Cui, Y. A. Liang, M. X. Wang, T. Dai, and Q. Y. Liu, "Consumption strength and recovery potential of metal minerals in China," *China Population Resources and Environment*, vol. 7, no. 7, pp. 53-59, 2017.

- [39] P. van der Werf, J. A. Seabrook, J. A. Gilliland, and J. Hewitt, "The quantity of food waste in the garbage stream of southern Ontario, Canada households," *PLoS ONE*, vol. 13, no. 6, Article ID e0198470, 2018.
- [40] A. K. Awasthi, F. Cucchiella, I. D'Adamo et al., "Modelling the correlations of e-waste quantity with economic increase," *Science of the Total Environment*, vol. 613-614, pp. 46-53, 2018.
- [41] M. Alamgir and A. Ahsan, "Municipal solid waste and recovery potential: Bangladesh perspective," *Journal of Environmental Health Science and Engineering*, vol. 4, no. 2, pp. 67-76, 2007.
- [42] B. Wallsten, A. Carlsson, P. Frändegård, J. Krook, and S. Svanström, "To prospect an urban mine - Assessing the metal recovery potential of infrastructure cold spots in Norrköping, Sweden," *Journal of Cleaner Production*, vol. 55, pp. 103-111, 2013.
- [43] C. Zhou, W. Fang, W. Xu, A. Cao, and R. Wang, "Characteristics and the recovery potential of plastic wastes obtained from landfill mining," *Journal of Cleaner Production*, vol. 80, pp. 80-86, 2014.
- [44] O. Krüger and C. Adam, "Recovery potential of German sewage sludge ash," *Waste Management*, vol. 45, pp. 400-406, 2015.
- [45] E. Stijin, Y. P. Joo, and R. C. Marian, "Quantifying the system-wide recovery potential of waste in global paper life cycle," *Resources, Conservation & Recycling*, vol. 134, pp. 48-60, 2018.
- [46] M. Ciroth, J. Finkberer, and J. Hildenbrand, *Towards A Live Cycle Sustainability Assessment Making Informed Choices on Products*, United Nations Environment Program (UNEP), 2011.
- [47] H. A. Arafat, K. Jijakli, and A. Ahsan, "Environmental performance and energy recovery potential of five processes for municipal solid waste treatment," *Journal of Cleaner Production*, vol. 105, pp. 233-240, 2015.
- [48] W. Simon, D. Noel, and C. Matt, "Environment protection agency of Ireland, waste from electric and electronic equipment in Ireland," in *A Statues Report*, pp. 68-81, Ireland, 2011.
- [49] E. Yamasue, K. Nakajima, H. Okumura, and K. N. Ishihara, "Estimation of the number of discarded home electric appliances considering new alternative products," *Nippon Kinzoku Gakkaishi*, vol. 70, no. 8, pp. 611-614, 2006.
- [50] Y. Yang and E. Williams, "Logistic model-based forecast of sales and generation of obsolete computers in the U.S.," *Technological Forecasting & Social Change*, vol. 76, no. 8, pp. 1105-1114, 2009.
- [51] L. Q. Lu, "Analysis on development trend of global lead-acid battery industry in 2018," 2018, <http://www.escn.com.cn>.
- [52] Y. N. Gao, H. Xu, and X. J. Lu, "Research on prediction of quantity of waste mobile phones based on market supply a model," in *Proceedings of Annual Meeting of Chinese Society for Environmental Sciences*, pp. 3597-3601, 2010.
- [53] "Annual survey: market share rose from 0 to 10%! lithium battery powered vehicles grew rapidly in 2017," 2017, <http://www.evtimes.cn/html/201712/66856.html>.
- [54] Z. Q. Guo, "Overview and development prospect of electric bicycles," in *Proceedings of Academic Seminar on Electric Vehicles and New-type Batteries*, pp. 40-45, 2003.
- [55] M. L. Liu, H. L. Zhong, and Y. Y. Wang, "Research on prediction of quantity of electronic waste based on GM (1, 1) Stanford Model," *China Market*, no. 12, pp. 95-98, 2015.
- [56] J. Wang, Y. Wang, J. liu, S. Zhang, and M. Zhang, "Effects of fund policy incorporating Extended Producer Responsibility for WEEE dismantling industry in China," *Resources, Conservation & Recycling*, vol. 130, pp. 44-50, 2018.

## Research Article

# Implementation Flexibility of Multiperiod Rail Line Design with Consideration of Uncertainties in Population Distribution

Kunpeng Zhang,<sup>1</sup> Binglei Xie,<sup>1</sup> and Ding Liu <sup>2</sup>

<sup>1</sup>Shenzhen Key Laboratory of Urban Planning and Decision Making Simulation, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen 518055, China

<sup>2</sup>College of Transport and Communications, Shanghai Maritime University, Shanghai 201306, China

Correspondence should be addressed to Ding Liu; [stunker@126.com](mailto:stunker@126.com)

Received 20 April 2018; Revised 8 July 2018; Accepted 22 July 2018; Published 1 November 2018

Academic Editor: Lu Zhen

Copyright © 2018 Kunpeng 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.

This paper aims to investigate the implementation flexibility of multiperiod rail line design in a linear monocentric city. Three alternatives (fast-tracking, deferring, and do-nothing-alternative (DNA) of a candidate rail line project) are examined, based on an in-depth uncertainties analysis of the demand side for this candidate rail line project. Conditions for the three alternatives of fast-tracking, deferring, and DNA are analytically explored and an illustrative example is given to demonstrate the application of the proposed models. Insightful findings are reported on the interrelationship between the rail line length and spatial and temporal correlation of population distribution as well as the implication of the correlation in practice. Sensitivity analyses are carried out in several scenarios in another numerical example to show the proposed conditions of three alternatives.

## 1. Introduction

In traditional transportation planning models, population distribution at each residential location was assumed to follow independence of irrelevant alternatives (IIA) (see, e.g., [1–4]). This assumption was generally acceptable since the traditional transportation planning models were commonly static and proposed for European and American cities with relatively low population densities.

The IIA assumption of population distribution, however, may result in inaccurate results and the ignorance of the correlation between random variables (Shao et al., 2012). Particularly, for cities with relatively high population densities, like Shanghai and Hong Kong in China, the correlation of travel demand was found to play a major role in road network expected total travel time [5]. Zhao and Kockelman [6] concluded that ignorance of correlation of travel demand would ultimately affect policy-making and infrastructure decisions. Yip et al. [7] confirmed the existence of the correlation of travel demand using the data in Hong Kong. Travel demand between two towns and urban areas in a two-5-year

period, 1996–2001 and 2001–2006, was deployed for investigation in their study.

With the IIA assumption of population distribution, uncertainties of population distribution cannot be fully explored. Uncertainties of population distribution can be classified into variations of population distribution year by year and spatial-temporal correlation of population distribution [8–10]. The year-by-year variations of population distribution can be captured by a stochastic variable of annual population growth rate, with its mean value and standard deviation [9, 11, 12]. The spatial-temporal correlation of population distribution can be described by the spatial and temporal correlation coefficient of population densities [8, 9, 13].

Uncertainties of population distribution directly affect the travel demand of rail service and further the implementation of a candidate rail transit line. For instance, while the population distribution is higher than the forecasting result of the candidate rail transit line in original feasible report, the candidate rail transit line can be fast-tracked. While the spatial-temporal correlation of population distribution is

TABLE 1: Comparison between some closely concerned models with the model proposed in this paper.

Citation	Variation of population distribution	Spatial-temporal correlation of population distribution	Multi-period NDP	Implementation flexibility
Vuchi and Newell (1968)	No	No	No	No
Chien and Qin (2004)	No	No	No	No
Szeto and Lo [19]	Yes	No	Yes	No
Ukkusuri and Patil (2009)	Yes	No	Yes	Yes
Ma and Lo [20]	Yes	No	Yes	No
Li et al. [2]	No	No	No	No
Shao et al. [9]	Yes	Yes	No	No
Liu [14]	Yes	Yes	No	No
Peng et al. [22]	Yes	No	No	No
This paper	Yes	Yes	Yes	Yes

very closely, the candidate rail transit line may be fast-tracked [8, 14].

The implementation flexibility problem of a candidate rail transit line is investigated in this paper. Specifically, the following questions are explored with the proposed model:

- (i) Under which condition, a rail line project should be fast-tracked or deferred?
- (ii) How long this rail line should be built and what are the mean and standard deviation of the rail line length?
- (iii) What effect exists between spatial and temporal correlation of population distribution and the rail line length?

The implementation flexibility problem of a candidate rail transit line is essentially a network design problem (NDP). In terms of time dimensions, NDP can be classified into two types of single-period NDP and multiperiod NDP. An NDP in most previous studies is typically examined in a single specific period. Lo and Szeto [15] extended a single-period NDP into a multiperiod NDP. Szeto and Lo [16, 17] incorporated time-dependent tolling into a multiperiod NDP. Szeto and Lo [18] took into account equity for multiperiod NDP, whereas Lo and Szeto [19] examined a multiperiod NDP with cost-recovery constraints. Ma and Lo [20] investigated time-dependent integrated transport supply and demand strategies and their impact on land use patterns.

Ukkusuri and Patil (2009) introduced flexibility into a multiperiod NDP, in which future investment could be deferred or abandoned. Flexibility is defined as the ability of the system to adapt to external changes, while maintaining satisfactory system performance [21]. Flexibility gives authorities and/or operators to fast-track or defer the future investment in a rail line system for several years, if necessary.

Table 1 compares some closely concerned models with the model proposed in this paper, with respect to variation of population distribution, spatial-temporal correlation of population distribution, multiperiod NDP, and implementation flexibility. In contrast with other studies, variation of

population distribution and spatial-temporal correlation of population distribution are both considered in this paper. To explore the implementation flexibility of the candidate rail transit line, a multiperiod model is proposed in this paper.

Three major extensions to the related literature are proposed in this paper: (1) the over-year uncertainties of population distribution are considered while conducting the implementation flexibility analysis of the candidate rail transit line; (2) the spatial-temporal correlation of population distribution is incorporated into the proposed model; (3) the benefit of fast-tracking a project and the penalty for deferring a project for several years are considered analytically. The proposed model has the potential to help authorities and/or operators implementing candidate rail line projects in an appropriate year, in accordance with the yearly varied travel demand of the rail service and time-dependent construction cost of the rail line projects.

The reminder of this paper is organized as follows: basic considerations are presented in Section 2. The conditions of fast-tracking or deferring the rail line project for several years are then explored in Section 3. Section 4 gives two numerical examples to show the contributions of the proposed models. Conclusions and further work are given in Section 5.

## 2. Basic Considerations

As shown in Figure 1,  $x_1$  and  $x_2$  represent the residential locations.  $L(t_1)$  and  $L(t_2)$  are the length of a candidate rail line in year  $t_1$  and  $t_2$ , which can be determined endogenously with the proposed model.  $\tilde{P}(x, t)$  is the yearly varied population density distributed in residential location  $x$  in year  $t$ , with  $\forall x \in \{x_1, x_2\}$  and  $\forall t \in \{t_1, t_2\}$ .

To facilitate the presentation of the essential ideas, the notations are summarized as follows. The notations are partitioned into the two types of deterministic variables and stochastic variables.

- (i) *Deterministic Variables*

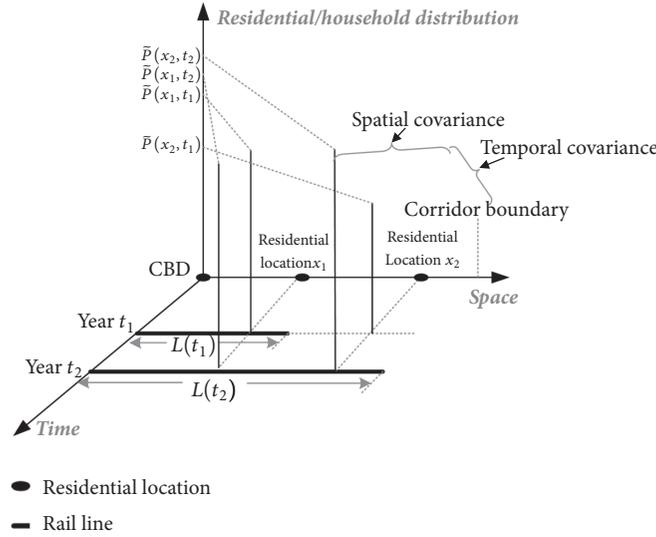


FIGURE 1: Configuration of a candidate rail line over years in a linear monocentric city.

$c(x, t)$ : the yearly nominal generalized travel cost for traveling from residential location  $x$  to the CBD in year  $t$ .

$D_s(t)$ : average station spacing of the candidate rail transit line in year  $t$ .

$f_0$ : fixed component of fare for using the rail service.

$f_1$ : variable component of fare per unit distance for using the rail service.

$h(q(x, t))$ : headway of train operation during morning peak hour in year  $t$ .

$H(x, t)$ : the nominal housing supply at residential location  $x$  in year  $t$ .

$L(t)$ : rail length in year  $t$ .

$P(x, t)$ : nominal population distribution at location  $x$  in year  $t$ .

$p(x, t)$ : the probability of choosing to live at residential location  $x$  in year  $t$ .

$q(x, t)$ : nominal travel demand of rail service at location  $x$  in year  $t$ .

$r(x, t)$ : the nominal housing rent at residential location  $x$  in year  $t$ .

$U(x, t)$ : the nominal disutility for residential location  $x$  in year  $t$ .

(ii) *Stochastic Variables*

$\tilde{P}(x, t)$ : yearly varied population distribution at location  $x$  in year  $t$ .

$\tilde{q}(x, t)$ : yearly varied travel demand of rail service at location  $x$  in year  $t$ .

$\tilde{U}(x, t)$ : yearly varied disutility for residential location  $x$  in year  $t$ .

To facilitate the presentation of the essential ideas, without loss of generality, the following assumptions are made in this paper:

A1: the candidate rail line project is assumed to be linear and start from the CBD and then be built along a linear monocentric city, as shown in Figure 1 [1, 2, 23]. The candidate rail line project in each period is assumed to finish on time and the rail service is expected to be provided at the end of each period [19].

A2: the standard deviation (SD) of population distribution is assumed to be an increasing function with respect to its mean value. This function is referred to as the stochastic population distribution function. Besides, the stochastic population distribution function is assumed as a nondecreasing function with respect to its mean value [9, 14].

A3: travelers' responses to the quality of rail service are measured by a generalized travel cost that is a weighted combination of in-vehicle time, access time, waiting time, and the fare [24]. Travelers are assumed to be homogeneous and have the same preferred arrival time at workplace located in the CBD. This study focuses mainly on travelers' home-based work trips, which are compulsory activities, and thus the number of trips is not affected by a variety of factors, such as income level [2].

A4: a rail project is worthy of investment, if benefit is not less than the penalty in its planning and operation horizon. The study period is assumed to be a one-hour period, for instance, the morning peak hour, which is usually the most critical period in the day [25].

### 3. Model Formulation

Population distribution is closely related to the planning procedure of a rail line project. This data and its growth rate over years are used to make strategic decisions including whether to introduce, defer, or fast-track new rail lines; how long the rail should be built; and how to determine effective

train operation parameters, such as the number of carriages in each train, the headway between trains, and the fares.

*3.1. Uncertainties of Population Distribution.* To allow for the yearly uncertainty of population distribution, it is assumed that there exists a perturbation in the population distribution. The yearly varied population distribution  $\tilde{P}(x, t)$  is given by the following equation [11]:

$$\tilde{P}(x, t) = P(x, t) + \varepsilon, \quad (1)$$

where  $P(x, t)$  is the nominal population distribution,  $E[\tilde{P}(x, t)] = P(x, t)$ ;  $\varepsilon$  is a random term, such that  $E[\varepsilon] = 0$ . Note that the nominal population distribution  $P(x, t)$  is deterministic.

In previous studies, population distribution was assumed to follow a Poisson distribution (e.g., [26]), normal distribution (e.g., Fu et al., 2012), or multivariate normal (MVN) distribution (e.g., [27]). For the assumption of Poisson distribution, equality of mean and variance was very rigid and may be inappropriate in practice [28]. The normal distribution was commonly assumed to maintain tractability of the model (Fu et al., 2012). Following the assumption of MVN distribution, it was important to propose efficient algorithms to avoid excessive computational costs [27, 29].

In terms of A2, the SD of population distribution can be expressed as [9, 14]

$$\sigma_p(x, t) = \sqrt{\text{var}[\tilde{P}(x, t)]} = \sqrt{\text{var}[\varepsilon]} = \varphi(P(x, t)), \quad (2)$$

where  $\varphi(\cdot)$  is defined as the stochastic population distribution function, which represents the functional relationship between the mean value and the variance of the stochastic population distribution.

To take spatial and temporal correlation of population distribution into account, the following spatial and temporal covariance is defined as [8]

$$\begin{aligned} \sigma_p(x_1, t_1; x_2, t_2) &= \text{cov}[\tilde{P}(x_1, t_1), \tilde{P}(x_2, t_2)] \\ &= \rho_{x_1, t_1}^{x_2, t_2} \varphi(P(x_1, t_1)) \varphi(P(x_2, t_2)), \end{aligned} \quad (3)$$

where  $\rho_{x_1, t_1}^{x_2, t_2}$  ( $-1 \leq \rho_{x_1, t_1}^{x_2, t_2} \leq 1$ ) is the correlation coefficient, which is an important measurement reflecting the statistical correlation between  $\tilde{P}(x_1, t_1)$  and  $\tilde{P}(x_2, t_2)$ . There are three correlation coefficient cases: negative, positive, or zero, representing negative, positive statistical dependence or statistical independence of population distribution. Specifically, with  $x_1 = x_2$ , and  $t_1 = t_2$ , the spatial and temporal covariance becomes the SD value.

The spatial and temporal correlation of population distribution cannot be taken into account under the assumptions of independent Poisson distribution and normal distribution. For instance, under independent normal distribution, the probability density function (PDF) of population distribution is given by  $f(P(x, t)) = (1/\sqrt{2\pi}\sigma_p(x, t))\exp(-(\tilde{P}(x, t) - P(x, t))^2/2(\sigma_p(x, t))^2)$ . Under MVN distribution, the PDF of population distribution is

$f(\mathbf{P}(\mathbf{x}, \mathbf{t})) = (1/(\sqrt{2\pi})^{|\mathbf{R}|/2} |\Sigma^{\mathbf{P}(\mathbf{x}, \mathbf{t})}|^{1/2}) \exp(-(1/2)(\tilde{\mathbf{P}}(\mathbf{x}, \mathbf{t}) - \mathbf{P}(\mathbf{x}, \mathbf{t}))^T (\Sigma^{\mathbf{P}(\mathbf{x}, \mathbf{t})})^{-1} (\tilde{\mathbf{P}}(\mathbf{x}, \mathbf{t}) - \mathbf{P}(\mathbf{x}, \mathbf{t})))$ , where  $\mathbf{P}(\mathbf{x}, \mathbf{t}) = (P(x_1, t_1), P(x_1, t_2), \dots, P(x_n, t_m))$  ( $n$  is the residential location number,  $m$  is the planning time horizon) is a  $|\mathbf{R}|$ -vector,  $|\Sigma^{\mathbf{P}(\mathbf{x}, \mathbf{t})}|$  is the determinant of  $\Sigma^{\mathbf{P}(\mathbf{x}, \mathbf{t})}$ , and  $\Sigma^{\mathbf{P}(\mathbf{x}, \mathbf{t})}$  is the covariance matrix of  $\mathbf{P}(\mathbf{x}, \mathbf{t})$ .

To consider the yearly uncertainty of rail service travel demand, it is assumed that a perturbation exists in this travel demand, and the yearly perturbed travel demand  $\tilde{q}(x, t)$  is given by the following equation:

$$\tilde{q}(x, t) = q(x, t) + \varepsilon_1, \quad (4)$$

where  $q(x, t)$  is the nominal travel demand at location  $x$  in year  $t$  and  $\varepsilon$  is a random term, with  $E[\tilde{q}(x, t)] = q(x, t)$  and  $E[\varepsilon_1] = 0$ .

The nominal travel demand of rail service is assumed to be a function of population distribution and generalized travel cost from residential locations to the CBD destination in terms of A3. Without loss of generality, an exponential function is used as follows [30]:

$$\begin{aligned} q(x, t) &= P(x, t) \exp(-\theta_1 c(x, t)), \\ \forall x \in [0, L], t \in [0, T], \end{aligned} \quad (5)$$

where  $q(x, t)$ ,  $P(x, t)$ , and  $c(x, t)$  are the yearly nominal travel demand of rail service, population distribution, and generalized travel cost, respectively;  $\theta_1$  is sensitivity parameter in travel demand function.

Population distribution  $P(x, t)$  is closely concerned with the disutility perceived by travelers residing at residential location  $x$  in year  $t$ . Because of the spatial and temporal correlation of population distribution, population distributed at different locations and years are not independent or irrelevant. To facilitate the computational advantage of the use of a closed-form analytical expression, the C-Logit model proposed by Cascetta et al. [31] was applied, owing to its relatively low levels of calibration influences and its rational behaviour consistent with random utility theory [32].

The C-Logit model of population distribution can be stated as follows [32]:

$$P(x, t) = P^t p(x, t), \quad (6)$$

where  $P^t$  is the total population along the rail line system in year  $t$ .  $p(x, t)$  is the probability of choosing to live at residential location  $x$  in year  $t$ :

$$p(x, t) = \frac{\exp[-\theta_2 (U(x, t) + CF(x, t))]}{\int_0^L \exp[-\theta_2 (U(w, t) + CF(w, t))] dw}, \quad (7)$$

where  $\theta_2$  is a dispersion parameter of population distribution. If  $\theta_2$  is large, population distribution along the candidate rail line is then assumed to be a decentralized type. If  $\theta_2$  is small, population is centrally distributed along the candidate rail line.  $CF(x, t)$  is the commonality factor for each residential location.

Several functional commonality factor forms have been proposed by Cascetta et al. [31]. These forms can be classified

into two types: flow-dependent and flow-independent. The flow-dependent functional form is used in this study as follows [11]:

$$CF(x, t) = \alpha_2 \ln \frac{U(x, t)}{\int_0^L U(w, t) dw}, \quad \forall t \in [0, T], \quad (8)$$

where  $\alpha_2$  indicates a constant parameter. Examples of the effect of changes to this parameter are given in the work of Cascetta et al. [31] and Prashker and Bekhor [33].

The population conservation equation can be expressed by

$$\int_0^L P(w, t) dw = P^t, \quad (9)$$

where  $P^t$  is the total population number along the candidate rail transit line in year  $t$ . To response the year-by-year change of the total population, a constant yearly growth rate is assumed [19]:

$$P^t = (1 + \gamma(t)) P^{t-1}, \quad (10)$$

where  $\gamma(t)$  is the average yearly growth rate of the total population number after year  $t$ . As  $\gamma(t)$  is positive, the implication is that the number of total population along the candidate rail transit line increases and vice versa.

In this paper, the disutility of travelers mainly consists of travel costs from residential location to CBD and housing rent. It is assumed that travelers put equal weighting on generalized travel cost and housing rent. Mathematically, it can be expressed as [2]

$$\tilde{U}(x, t) = U(x, t) + \varepsilon_2(x, t), \quad (11)$$

$$U(x, t) = c(x, t) + r(x, t), \quad (12)$$

where  $U(x, t)$  is nominal disutility,  $c(x, t)$  is the nominal generalized travel cost,  $r(x, t)$  is the nominal housing rent at residential location  $x$  in year  $t$ , and  $\varepsilon_2(x, t)$  is the associated random terms, where  $E[\varepsilon_2(x, t)] = 0$ .

The nominal generalized travel cost consists of fare, access cost for travelers from residential locations to rail stations, waiting cost for rail service at stations, and in-vehicle cost from rail stations to CBD, shown as follows [30]:

$$c(x, t) = f(x, t) + \mu_c t_c + \mu_w t_w + \mu_i t_i, \quad (13)$$

where  $\mu_c/\mu_w/\mu_i$  are values of access time, waiting time, and in-vehicle time, respectively;  $f(x, t)$  is distance-based fare for rail service,  $t_c$  is average access time for travelers from residential locations to the rail station,  $t_w$  is the average waiting time for rail service at stations, and  $t_i$  is average in-vehicle cost from rail stations to CBD. The distanced-based fare  $f(x, t)$  is given by

$$f(x, t) = f_0 + f_1 x, \quad (14)$$

where  $f_0$  is the fixed fare component and  $f_1$  is the variable fare component per kilometer. Waiting time  $t_w$  is closely

concerned with travel demand and supply of the rail service. For long-term planning, this value can be estimated using the following function:

$$t_w = \alpha h(q(x, t)), \quad (15)$$

where  $\alpha$  is a calibration parameter which depends on the distribution of train headway and travelers arrival time, and  $h(q(x, t))$  is the average headway [34].

The nominal rent for a house rises in proportion to the number of families wanting to live in that house. For instance, the landlord can increase the rent in accordance with the demand. Hence, in theory of the housing supply matches demand, the nominal rent is likely to be reduced. In response to this factor, housing rent is assumed to be given by the following function [35]:

$$r(x, t) = \alpha_1 \left( 1 + \beta_1 \frac{P(x, t)}{H(x, t) - P(x, t)} \right), \quad (16)$$

where  $\alpha_1$  and  $\beta_1$  are parameters and  $H(x, t)$  is the nominal housing supply. Calibration of the housing function parameters is indispensable.

*3.2. Payoff of Fast-Tracking or Deferring a Candidate Rail Project.* A candidate rail transit line project can be implemented in three different ways:

- (1) It can be implemented as planned (do-nothing-alternative (DNA)).
- (2) It can be fast-tracked several years.
- (3) It can be deferred several years.

Which alternative should be used depends on the payoff with respect to benefit and penalty, as shown in Table 2.

For the DNA, benefit comes from the convenience of rail service supplied to travelers during the operation horizon of this rail line system. This benefit can be measured by customer surplus. Mathematically, it can be expressed as (Ukkusuri and Patil, 2009):

$$\begin{aligned} \tilde{\pi}(\gamma(t)) = \sum_{t=0}^T \int_0^{L(t)} \left[ \int_0^{q(x,t)} (q(w, t))^{-1} dw \right. \\ \left. - q(x, t) c(x, t) \right] dx = \sum_{t=0}^T \int_0^{L(t)} \frac{q(x, t)}{\theta_1} dx. \end{aligned} \quad (17)$$

Since  $q(x, t)$  is an increasing function of  $P(x, t)$  and  $P(x, t)$  is an increasing function of  $\gamma$ ,  $\tilde{\pi}(\gamma(t))$  is an increasing function of  $\gamma(t)$ .

The daily average penalty of the DNA is comprised by construction cost, operation cost, and maintenance cost [36]. This penalty can be expressed as

$$\begin{aligned} \hat{\pi}(i(t)) \\ = \frac{C_l L(t) + C_s L(t) / D_s(t) + C_{om} PWF(i(t), T)}{365}, \end{aligned} \quad (18)$$

where  $C_l$  is the construction cost for rail line per kilometer if this project is implemented as schedule,  $C_s$  is the construction

TABLE 2: The benefit and penalty of each alternative.

Alternative	Project start time	Interest rate	Growth factor of population density per year	Benefit	Penalty
DNA	0	$i(t)$	$\gamma(t)$	$\tilde{\pi}(\gamma(t))$	$\tilde{\pi}(i(t))$
Fast-tracked	$-t_1$	$i_f(t)$	$\gamma_f(t)$	$\tilde{\pi}_f(\gamma_f(t))$	$\tilde{\pi}_f(i_f(t))$
Deferred	$t_2$	$i_d(t)$	$\gamma_d(t)$	$\tilde{\pi}_d(\gamma_d(t))$	$\tilde{\pi}_d(i_d(t))$

Notes:

- (1) The original project start time as schedule is set as year 0.
- (2) The planning and operation horizon is assumed to be same for the three alternatives.
- (3) Normally, the following equations are held:  
 $\gamma_f(t) > \gamma(t) > \gamma_d(t)$  or  $i_d(t) > i(t) > i_f(t)$ .

cost of each rail station,  $D_s(t)$  is the average station spacing of this rail transit line, and  $C_{om}$  is the annual operation and maintenance cost of this rail transit line project. The number "365" is used to convert the annual average penalty to daily average penalty.

If the alternative option of fast-tracking this rail line project is chosen, travelers would enjoy the benefits of rail service earlier. Meanwhile, the total cost of construction, operation, and maintenance costs of the system would increase

compared with those of the original plan, within the same project planning and operation horizon. Mathematically, the benefit can be expressed as

$$\tilde{\pi}_f(\gamma_f(t)) = \sum_{t=-t_1}^{T-t_1} \int_0^{L(t)} \frac{q(x,t)}{\theta_1} dx, \quad (19)$$

and the daily average penalty for this alternative is expressed as follows:

$$\tilde{\pi}_f(i_f(t)) = \frac{C_l CAF(i_f(t), t_1) L(t) + C_s CAF(i_f(t), t_1) L(t) / D_s(t) + C_{om} CAF(i_f(t), t_1) PWF(i_f(t), T)}{365}, \quad (20)$$

where  $CAF(i_f(t), t_1)$  is compound-amount factor, defined as  $(1+i_f(t))^{t_1}$ . This factor takes into account the time cost of each construction or operation cost component of rail project in terms of average interest rate  $i_f(t)$  and fast-tracking period  $t_1$ . The first term  $C_l CAF(i_f(t), t_1) L(t) / 365$  represents the daily construction cost of rail line, as the rail transit line project is fast-tracked  $t_1$  years. Accordingly, the second term in (21) represents the daily construction cost of all rail stations, and

the third term in (21) represents the daily operation and maintenance cost of this rail transit line project.

If the alternative of deferring this rail transit line project is chosen, the total cost of construction, operation, and maintenance costs would decrease compared with those of the original plan, within the same project planning and operation horizon. However, the travelers would suffer the penalty of traffic crowding until a better rail service is supplied. Mathematically, the daily average benefit is calculated by

$$\tilde{\pi}_d(\gamma_d(t)) = \frac{C_l PWF'(i_d(t), t_2) L(t) + C_s PWF'(i_d(t), t_2) L(t) / D_s(t) + C_{om} PWF'(i_d(t), t_2) PWF(i_d(t), T)}{365}, \quad (21)$$

where  $PWF'(i_d(t), t_2)$  is present-worth factor to obtain present value of a future value, defined as  $1/(1+i(t))^t$ . The first term  $C_l PWF'(i_d(t), t_2) L(t) / 365$  represents the daily construction cost of rail line, as the rail transit line project is deferred  $t_2$  years. Accordingly, the second term in (22) represents the daily construction cost of all rail stations, and the third term in (22) represents the daily operation and maintenance cost of this rail transit line project.

The daily average penalty of deferring the rail transit line project mainly comes from the travel inconvenience of travelers. For instance, traffic congestion cannot be eliminated while rail service is not supplied on time. This daily average

penalty can be measured by consumer surplus of travelers, expressed as

$$\tilde{\pi}_d(\gamma_d(t)) = \sum_{t=0}^{t_2} \int_0^{L(t)} \frac{q(x,t)}{\theta_1} dx. \quad (22)$$

where  $t_2$  is the deferred period and the integral term represents daily consumer surplus of travelers, namely, daily average penalty of deferring the rail project.

**3.3. Conditions for Fast-Tracking or Deferring a Candidate Rail Project.** In terms of A4, a rail transit line project is worthy

of investment, if benefit  $\tilde{\pi}(\gamma(t))$  is not less than the penalty  $\tilde{\pi}(i(t))$  in its planning and operation horizon  $t \in [0, T]$ . Specifically, this rail project is break-even, if benefit  $\tilde{\pi}(\gamma(t))$  is equal to penalty  $\tilde{\pi}(i(t))$ , namely,

$$\tilde{\pi}(\gamma(t)) = \tilde{\pi}(i(t)), \quad (23)$$

In (23), given interest rate  $i(t)$ , rail length  $L(t)$ , and average station spacing  $D_s(t)$ , only the growth rate of total population density  $\gamma(t)$  is unknown in terms of (1)–(21). By solving (23), a break-even growth rate  $\gamma^*(t)$  of total population along the candidate rail system can be obtained. As the growth rate of the total population in the candidate rail system is greater than  $\gamma^*(t)$ , this rail project is worthy of investment. Similarly, given the growth rate of total population  $\gamma(t)$ , rail length  $L(t)$ , and average station spacing  $D_s(t)$ , only the interest rate  $i(t)$  is unknown. By solving (23), the break-even interest rate  $i^*(t)$  can be determined. This  $i^*(t)$  is the internal rate of return (IRR), which makes this project just break-even. When the actual interest rate is lower than this IRR  $i^*(t)$ , the project is worthy of investment. With a given interest rate  $i^*(t)$ , a growth rate of total population density  $\gamma^*(t)$ , and average station spacing  $D_s(t)$ , the rail length  $L^*(t)$ , which makes this project break-even, can be determined year by year by solving (23).

As previously stated, the rail transit line project can be fast-tracked or deferred, while interest rate or growth rate of total population varies over years. Different scenarios of interest rate and growth rate of total population over years are investigated here. For each scenario, the increased benefit and

increased penalty or the consumer surplus loss and capital cost saving should be firstly compared, so as to determine the suitable alternative. A detailed results summary is given in Table 3. The expressions of benefit and penalty in Table 3 are given by (18)–(23).

The values of  $\gamma_f(t) > \gamma(t) > \gamma_d(t)$  and  $i_d(t) > i(t) > i_f(t)$  in Table 3 are assumed to be given. The conditions for fast-tracking or deferring the rail project are sufficient conditions. If the values of  $\gamma_f(t) > \gamma(t) > \gamma_d(t)$  and  $i_d(t) > i(t) > i_f(t)$  are unknown, the necessary conditions for fast-tracking or deferring the rail projects are required. Propositions 3 and 4 present necessary conditions for fast-tracking or deferring the rail project. The values of  $\gamma_f^*(t)$ ,  $i_f^*(t)$ ,  $\gamma_d^*(t)$ , and  $i_d^*(t)$  can be determined endogenously.

**Proposition 1.** *Necessary condition for fast-tracking the rail project is summarized as follows.*

*A rail project is worthy of fast-tracking, only if the increased benefit  $(\partial\tilde{\pi}_f(\gamma_f(t))/\partial\gamma_f(t))(\gamma_f(t) - \gamma^*(t))$  is more than the increased penalty  $(\partial\tilde{\pi}_f(i_f(t))/\partial i_f(t))(i^*(t) - i_f(t))$ ; namely,*

$$\begin{aligned} & \frac{\partial\tilde{\pi}_f(\gamma_f(t))}{\partial\gamma_f(t)} (\gamma_f(t) - \gamma^*(t)) \\ & > \frac{\partial\tilde{\pi}_f(i_f(t))}{\partial i_f(t)} (i^*(t) - i_f(t)). \end{aligned} \quad (24)$$

where

$$\begin{aligned} \frac{\partial\tilde{\pi}_f(\gamma_f(t))}{\partial\gamma_f(t)} &= \frac{1}{\theta_1} \sum_{t=t_1}^{T-t_1} \int_0^L ((t-1)(1+\gamma^*(t))^{t-2} P^0 p(x,t)(1-\theta_1 c(x,t))) dx. \\ \frac{\partial\tilde{\pi}_f(i_f(t))}{\partial i_f(t)} &= \left( C_1 L(t) + \frac{C_s L(t)}{D_s(t)} \right) \frac{t_1 (1+i_f(t))^{t_1-1}}{365} \\ &+ C_{om} \frac{(T+t_1)i_f(t)(1+i_f(t))^T - [(1+i_f(t))^T - 1](1+i_f(t) + i_f(t)T) - t_1 i_f(t)}{(365)i_f^2(t)(1+i_f(t))^{T-t_1+1}}. \end{aligned} \quad (25)$$

*Proof.* The first-order derivative of benefit  $\tilde{\pi}_f(\gamma_f(t))$  with respect to growth rate  $\gamma_f(t)$  is derived as follows:

$$\begin{aligned} \frac{\partial\tilde{\pi}_f(\gamma_f(t))}{\partial\gamma_f(t)} &= \frac{1}{\theta_1} \sum_{t=t_1}^{T-t_1} \int_0^L \frac{\partial q(x,t)}{\partial\gamma_f(t)} dx, \\ \frac{\partial q(x,t)}{\partial\gamma_f(t)} &= \exp(-\theta_1 c(x,t)) \left[ \frac{\partial P(x,t)}{\partial\gamma_f(t)} - \theta_1 P(x,t) \frac{\partial c(x,t)}{\partial\gamma_f(t)} \right], \\ \frac{\partial P(x,t)}{\partial\gamma_f(t)} &= \frac{\partial P^t}{\partial\gamma_f(t)} P(x,t) + P^t \frac{\partial p(x,t)}{\partial\gamma_f(t)}, \end{aligned}$$

TABLE 3: Alternative choice with respect to different scenarios of growth factor of population density over years and interest rate ( $\gamma_f(t) > \gamma(t) > \gamma_d(t)$  and  $i_d(t) > i(t) > i_f(t)$ ).

	Interest rate	Growth factor	Results
Scenario 1	$i_f(t)$	$\gamma_d(t)$	$(\partial\tilde{\pi}(\gamma_d(t))/\partial\gamma_d(t))(\gamma^*(t) - \gamma_d(t)) < (\partial\tilde{\pi}(i_f(t))/\partial i_f(t))(i^*(t) - i_f(t))$ , fast-track
			$(\partial\tilde{\pi}(\gamma_d(t))/\partial\gamma_d(t))(\gamma^*(t) - \gamma_d(t)) = (\partial\tilde{\pi}(i_f(t))/\partial i_f(t))(i^*(t) - i_f(t))$ , DNA
			$(\partial\tilde{\pi}(\gamma_d(t))/\partial\gamma_d(t))(\gamma^*(t) - \gamma_d(t)) > (\partial\tilde{\pi}(i_f(t))/\partial i_f(t))(i^*(t) - i_f(t))$ , defer
Scenario 2	$i_f(t)$	$\gamma(t)$	Fast-track
Scenario 3	$i_f(t)$	$\gamma_f(t)$	Fast-track
Scenario 4	$i(t)$	$\gamma_d(t)$	Defer
Scenario 5	$i(t)$	$\gamma(t)$	DNA
Scenario 6	$i(t)$	$\gamma_f(t)$	Fast-track
Scenario 7	$i_d(t)$	$\gamma_d(t)$	Defer
Scenario 8	$i_d(t)$	$\gamma(t)$	Defer
Scenario 9	$i_d(t)$	$\gamma_f(t)$	$(\partial\tilde{\pi}(\gamma_f(t))/\partial\gamma_f(t))(\gamma_f(t) - \gamma^*(t)) > (\partial\tilde{\pi}(i_d(t))/\partial i_d(t))(i_d(t) - i^*(t))$ ,fast-track
			$(\partial\tilde{\pi}(\gamma_f(t))/\partial\gamma_f(t))(\gamma_f(t) - \gamma^*(t)) = (\partial\tilde{\pi}(i_d(t))/\partial i_d(t))(i_d(t) - i^*(t))$ ,DNA
			$(\partial\tilde{\pi}(\gamma_f(t))/\partial\gamma_f(t))(\gamma_f(t) - \gamma^*(t)) < (\partial\tilde{\pi}(i_d(t))/\partial i_d(t))(i_d(t) - i^*(t))$ , defer

$$\frac{\partial P^t}{\partial \gamma_f(t)} = (t-1)(1+\gamma(t))^{t-2} P^0,$$

$$\begin{aligned} & \frac{\partial p(x,t)}{\partial \gamma_f(t)} \\ &= \frac{-\theta_2 \exp[-\theta_2(U(x,t) + CF(x,t))] (\partial U(x,t)/\partial \gamma_f(t) + \partial CF(x,t)/\partial \gamma_f(t))}{\int_0^L \exp[-\theta_2(U(w,t) + CF(w,t))] dw} \\ &+ \frac{\theta_2 \exp[-\theta_2(U(x,t) + CF(x,t))] \int_0^L \exp[-\theta_2(U(w,t) + CF(w,t))] (\partial U(w,t)/\partial \gamma_f(t) + \partial CF(w,t)/\partial \gamma_f(t)) dw}{\left[ \int_0^L \exp(-\theta_2(U(w,t) + CF(w,t))) dw \right]^2}, \end{aligned} \quad (26)$$

$$\frac{\partial U(x,t)}{\partial \gamma_f(t)} = \frac{\partial c(x,t)}{\partial \gamma_f(t)} + \frac{\partial r(x,t)}{\partial \gamma_f(t)},$$

$$\frac{\partial CF(x,t)}{\partial \gamma_f(t)} = \alpha_2 \frac{\int_0^L \sqrt{U(w,t)} dw}{U(x,t)} \left[ \frac{\partial U(x,t)}{\partial \gamma_f(t)} \frac{1}{\int_0^L \sqrt{U(w,t)} dw} - \frac{U(x,t) \int_0^L (1/\sqrt{U(w,t)}) (\partial U(w,t)/\partial \gamma_f(t)) dw}{2 \left( \int_0^L \sqrt{U(w,t)} dw \right)^2} \right], \quad (27)$$

$$\frac{\partial c(x,t)}{\partial \gamma_f(t)} = \alpha \frac{\partial h(q(x,t))}{\partial q(x,t)} \frac{\partial q(x,t)}{\partial \gamma_f(t)} = 0,$$

$$\frac{\partial r(x,t)}{\partial \gamma_f(t)} = \frac{\beta H(x,t)}{(H(x,t) - P(x,t))^2} \frac{\partial P(x,t)}{\partial \gamma_f(t)} = 0,$$

Thus, we have

$$\frac{\partial \bar{\pi}(\gamma_f(t))}{\partial \gamma_f(t)} = \frac{1}{\theta_1}$$

$$\cdot \sum_{t=t_1}^{T-t_1} \int_0^L \left( \frac{\partial P(x,t)}{\partial \gamma_f(t)} (1 - \theta_1 P(x,t)) \frac{\partial c(x,t)}{\partial \gamma_f(t)} \right) dx$$

$$= \frac{1}{\theta_1} \sum_{t=t_1}^{T-t_1} \int_0^L \left( ((t-1)(1+\gamma^*(t))^{t-2} P^0 p(x,t)) \cdot (1 - \theta_1 c(x,t)) \right) dx.$$

(28)

The first-order derivative of penalty  $\bar{\pi}_f(i_f(t))$  with respect to interest rate  $i_f(t)$  is derived as

$$\begin{aligned} \frac{\partial \bar{\pi}_f(i_f(t))}{\partial i_f(t)} &= \frac{1}{365} \frac{\partial (C_l CAF(i_f(t), t_1) L(t) + C_s CAF(i_f(t), t_1) L(t) / D_s(t) + C_{om} CAF(i_f(t), t_1) PWF(i_f, T))}{\partial i_f} \\ &= \left( C_l L(t) + \frac{C_s L(t)}{D_s(t)} \right) \frac{t_1 (1 + i_f(t))^{t_1-1}}{365} \\ &\quad + C_{om} \frac{(T + t_1) i_f(t) (1 + i_f(t))^T - \left[ (1 + i_f(t))^T - 1 \right] (1 + i_f(t) + i_f(t) T) - t_1 i_f(t)}{(365) i_f^2(t) (1 + i_f(t))^{T-t_1+1}}. \end{aligned}$$

(29)

Given the values of  $i^*(t), i_f(t), \gamma^*(t), D_s(t)$  and rail length  $L^*(t)$ , solving the equation of  $(\partial \bar{\pi}_f(\gamma_f(t)) / \partial \gamma_f(t))(\gamma_f(t) - \gamma^*(t)) = (\partial \bar{\pi}_f(i_f(t)) / \partial i_f(t))(i^*(t) - i_f(t))$ , a threshold value  $\gamma_f^*(t)$  could be obtained. Once the expected growth factor of the total population density is larger than  $\gamma_f^*(t)$ , the project is worth fast-tracking. Similarly, given a value of  $i^*(t), \gamma^*(t), \gamma_f(t), D_s(t)$  and rail length  $L^*(t)$ , a threshold value  $i_f^*(t)$  could be obtained. Once the

interest rate is lower than  $i_f(t)$ , the project is worth fast-tracking.

Given the values of  $i^*(t), \gamma^*(t), D_s(t)$ , the growth rate of  $\gamma_f(t)$ , and interest rate  $i_f(t)$ , the year-by-year rail length  $L_f(t)$  can also be determined, shown as Corollary 2.  $\square$

**Corollary 2.** Given the values of  $i^*(t), \gamma^*(t), D_s(t)$ , the growth rate of  $\gamma_f(t)$ , and interest rate  $i_f(t)$ , the year-by-year rail length  $L_f(t)$  for fast-tracking the rail project  $t_1$  years is given by

$$\begin{aligned} L(t) &= \frac{365 \sum_{t=t_1}^{T-t_1} \int_0^L \left( ((t-1)(1+\gamma^*(t))^{t-2} P^0 p(x,t)) (1 - \theta_1 c(x,t)) \right) dx}{\theta_1 (C_l + C_s / D_s(t)) t_1 (1 + i_f(t))^{t_1-1}} \\ &\quad - \frac{C_{om}}{(C_l + C_s / D_s(t))} \frac{(T + t_1) i_f(t) (1 + i_f(t))^T - \left[ (1 + i_f(t))^T - 1 \right] (1 + i_f(t) + i_f(t) T) - t_1 i_f(t)}{t_1 i_f^2(t) (1 + i_f(t))^T}. \end{aligned}$$

(30)

Because the values of  $p(x,t)$  are a stochastic value, the year-by-year rail length  $L_f(t)$  is also a stochastic value. The expression of its standard deviation is given by

$$\begin{aligned} &365 \int_{-t_1}^{T-t_1} \int_{-t_1}^{T-t_1} \int_0^{L(t)} \int_0^{L(t)} \sigma_p(x_1, t_{11}; x_2, t_{12}) \sigma \left( (t_{11} - 1) (1 + \gamma^*(t_{11}))^{t_{11}-2} P(x_1, t_{11}) \right) \\ &\sigma(L(t)) = \frac{\times \sigma \left( (t_{12} - 1) (1 + \gamma^*(t_{12}))^{t_{12}-2} P(x_2, t_{12}) \right) dx_1 dx_2 dt_{11} dt_{12}}{\theta_1 (C_l + C_s / D_s(t)) t_1 (1 + i_f(t))^{t_1-1}}, \end{aligned}$$

(31)

where  $((t_{11} - 1)(1 + \gamma^*(t_{11}))^{t_{11}-2} P(x_1, t_{11}))$  and  $\sigma_p(x_1, t_{11}; x_2, t_{12})$  are given by (3) and (4).

*Proof.* In terms of Proposition 1, let  $(\partial \tilde{\pi}_f(\gamma_f(t))/\partial \gamma_f(t))(\gamma_f(t) - \gamma^*(t)) = (\partial \tilde{\pi}_f(i_f(t))/\partial i_f(t))(i^*(t) - i_f(t))$ ; we have

$$L(t) = \frac{365 \sum_{t=t_1}^{T-t_1} \int_0^L ((t-1)(1 + \gamma^*(t))^{t-2} P^0 p(x, t)) (1 - \theta_1 c(x, t)) dx}{\theta_1 (C_l + C_s/D_s(t)) t_1 (1 + i_f(t))^{t_1-1}} - \frac{C_{om}}{(C_l + C_s/D_s(t))} \frac{(T + t_1) i_f(t) (1 + i_f(t))^T - [(1 + i_f(t))^T - 1] (1 + i_f(t) + i_f(t)T) - t_1 i_f(t)}{t_1 i_f^2(t) (1 + i_f(t))^T}. \quad (32)$$

With given  $i^*, \gamma^*, i_f, \gamma_f$ , the standard deviation of  $L(t)$ ,  $\sigma(L(t))$  is calculated by

$$\sigma(L(t)) = \frac{365}{\theta_1 (C_l + C_s/D_s(t)) t_1 (1 + i_f(t))^{t_1-1}} \sigma \left( \sum_{t=t_1}^{T-t_1} \int_0^{L(t)} ((t-1)(1 + \gamma^*(t))^{t-2} P^0 p(x, t)) (1 - \theta_1 c(x, t)) dx \right) = \frac{365 \int_{-t_1}^{T-t_1} \int_{-t_1}^{T-t_1} \int_0^{L(t)} \int_0^{L(t)} \sigma_p(x_1, t_{11}; x_2, t_{12}) \sigma((t_{11} - 1)(1 + \gamma^*(t_{11}))^{t_{11}-2} P(x_1, t_{11})) \sigma((t_{12} - 1)(1 + \gamma^*(t_{12}))^{t_{12}-2} P(x_2, t_{12})) dx_1 dx_2 dt_{11} dt_{12}}{\theta_1 (C_l + C_s/D_s(t)) t_1 (1 + i_f(t))^{t_1-1}} \quad (33)$$

From (31), it can be seen that  $\sigma(L(t))$  is underestimated while both spatial and temporal covariance parameters are both positive and overestimated while both spatial and temporal covariance parameters are negative.  $\square$

**Proposition 3.** Necessary condition for deferring the rail project is summarized as follows.

A rail project is worth deferring by several years, if the loss of consumer welfare  $(\partial \tilde{\pi}_d(\gamma_d(t))/\partial \gamma_d(t))(\gamma^*(t) - \gamma_d(t))$  is less

than the saving in capital cost  $(\partial \tilde{\pi}_d(i_d(t))/\partial i_d(t))(i_d(t) - i^*(t))$ ; namely,

$$\frac{\partial \tilde{\pi}_d(\gamma_d(t))}{\partial \gamma_d(t)} (\gamma^*(t) - \gamma_d(t)) < \frac{\partial \tilde{\pi}_d(i_d(t))}{\partial i_d(t)} (i_d(t) - i^*(t)) \quad (34)$$

where

$$\frac{\partial \tilde{\pi}_d(i_d(t))}{\partial i_d(t)} = \left( C_l L(t) + \frac{C_s L(t)}{D_s(t)} \right) \frac{-t_2}{365 \cdot (1 + i_d(t))^{t_2+1}} + C_{om} \frac{T i_d(t) (1 + i_d(t))^T + ((1 + i_d(t))^T - 1) ((1 + i_d(t) + T i_d(t)) - t_2 i_d(t))}{(365) i_d^2(t) (1 + i_d(t))^{T+t_2+1}} \quad (35)$$

$$\frac{\partial \tilde{\pi}_d(\gamma_d(t))}{\partial \gamma_d(t)} = \frac{1}{\theta_1} \sum_{t=t_2}^{T+t_2} \int_0^L ((t-1)(1 + \gamma^*(t))^{t-2} P^0 p(x, t)) (1 - \theta_1 c(x, t)) dx.$$

*Proof.* The first-order derivatives of benefit  $\tilde{\pi}_d(i_d(t))$  and penalty  $\tilde{\pi}_d(\gamma_d(t))$  with respect to interest  $i_d(t)$  and growth rate  $\gamma_d(t)$  are derived as follows, respectively:

$$\begin{aligned} \frac{\partial \tilde{\pi}_d(i_d(t))}{\partial i_d(t)} &= \frac{\partial (C_l PWF'(i_d(t), t_2) L(t) + C_s PWF'(i_d(t), t_2) L(t) / D_s(t) + C_{om} PWF'(i_d(t), t_2) PWF(i_d(t), T))}{365 \cdot \partial i_d(t)} \\ &= \left( C_l L(t) + \frac{C_s L(t)}{D_s(t)} \right) \frac{-t_2}{365 \cdot (1 + i_d(t))^{t_2+1}} \\ &\quad + C_{om} \frac{Ti_d(t) (1 + i_d(t))^T + ((1 + i_d(t))^T - 1) ((1 + i_d(t) + Ti_d(t)) - t_2 i_d(t))}{(365) i_d^2 (1 + i_d(t))^{T+t_s+1}} \end{aligned} \tag{36}$$

$$\frac{\partial \tilde{\pi}_d(\gamma_d(t))}{\partial \gamma_d(t)} = \frac{1}{\theta_1} \sum_{t=t_2}^{T+t_2} \int_0^L ((t-1)(1+\gamma^*(t))^{t-2} P^0 p(x,t)) (1-\theta_1 c(x,t)) dx.$$

By solving the equation of  $(\partial \tilde{\pi}_d(\gamma_d(t)) / \partial \gamma_d(t)) (\gamma^*(t) - \gamma_d(t)) = (\partial \tilde{\pi}_d(i_d(t)) / \partial i_d(t)) (i_d(t) - i^*(t))$ , with given values of  $i^*(t), \gamma^*(t), D_s(t)$ , rail length  $L^*(t)$ , and growing rate  $\gamma_d(t)$  or interest rate  $i_d(t)$ , the other threshold value of  $\gamma_d(t)$  or  $i_d(t)$  can be determined. With given values of  $i^*(t), \gamma^*(t), D_s(t), \gamma_d(t)$ , and  $i_d(t)$ , the year-by-year rail length  $L_d(t)$  can also be updated for the alternative of deferring this project as Corollary 2. The interrelationship between interest

rate and growing rate of the total population distribution is explored, shown as the following Proposition 4.  $\square$

**Proposition 4.** Given rail length  $L(t)$  and average station spacing  $D_s(t)$ , IRR  $i^*(t)$  is a strictly increasing function of the break-even growth rate of the total population  $\gamma^*(t)$  in the linear monocentric city.

*Proof.* Let  $\tilde{\pi}(\gamma(t)) = \tilde{\pi}(i(t))$ , in terms of (18) and (19); we have

$$\begin{aligned} \tilde{\pi}(\gamma(t)) &= \sum_{t=0}^T \int_0^L \frac{q(x,t)}{\theta_1} dx \\ &= \frac{1}{\theta_1} \sum_{t=0}^T (1+\gamma(t))^t P^0 \int_0^L \frac{\exp[-\theta_2 (U(x,t) + \alpha_0 \ln(U(x,t) / \int_0^L \sqrt{U(w,t)} dw))] }{\int_0^L \exp[-\theta_2 (U(w,t) + \alpha_0 \ln(U(x,t) / \int_0^L \sqrt{U(w,t)} dw))] dw} (1-\theta_1 c(x,t)) dx \\ &= \frac{1}{\theta_1} \sum_{t=0}^T (1+\gamma(t))^t P^0 \int_0^L \frac{\exp(-\theta_2 U(x,t)) ((2/L) \sqrt{U(x,t)})^{-\theta_2 \alpha_0}}{\int_0^L \exp(-\theta_2 U(w,t)) ((2/L) \sqrt{U(x,t)})^{-\theta_2 \alpha_0} dw} (1-\theta_1 c(x,t)) dx \\ &= \frac{1}{\theta_1} \sum_{t=0}^T (1+\gamma(t))^t P^0 \int_0^L \frac{\exp(-\theta_2 U(x,t)) U(x,t)^{-\theta_2 \alpha_0/2}}{(1/f_1) \int_{U(0,t)}^{U(L,t)} \exp(-\theta_2 U(w,t)) U(w,t)^{-\theta_2 \alpha_0/2} dU(w,t)} (1-\theta_1 c(x,t)) dx \\ &= \frac{1}{\theta_1} \sum_{t=0}^T (1+\gamma(t))^t P^0 \int_0^L \frac{\exp(-\theta_2 U(x,t)) U(x,t)^{-\theta_2 \alpha_0/2}}{(1/\theta_2 f_1) (\sum_{i=0}^n U(w,t)^{-\theta_2 \alpha_0/2-i}) \exp(-\theta_2 U(w,t)) \Big|_0^L} (1-\theta_1 c(x,t)) dx \\ &= \frac{\theta_2 f_1}{\theta_1} \sum_{t=0}^T (1+\gamma(t))^t P^0 \int_0^L \sum_{i=0}^n U(w,t)^i (1-\theta_1 c(w,t)) dw \end{aligned} \tag{37}$$

with

$$\begin{aligned} & \frac{1}{f_1} \int_{U(0,t)}^{U(L,t)} \exp(-\theta_2 U(w,t)) U(w,t)^{-\theta_2 \alpha_0/2} dU(w,t) \\ &= \frac{1}{\theta_2 f_1} \left( \sum_{i=0}^n U(w,t)^{-\theta_2 \alpha_0/2-i} \right) \exp(-\theta_2 U(w,t)) \Big|_0^L \quad (38) \\ &+ \frac{1}{f_1} \int_{U(0,t)}^{U(L,t)} \exp(-\theta_2 U(w,t)) \\ &\cdot U(x,t)^{-\theta_2 \alpha_0/2-n} dU(x,t), \end{aligned}$$

where  $n$  is a positive integer and big enough to confirm that

$$\begin{aligned} & \frac{1}{f_1} \int_{U(0,t)}^{U(L,t)} \exp(-\theta_2 U(w,t)) U(x,t)^{-\theta_2 \alpha_0/2-n} dU(x,t) \\ &= 0 \end{aligned} \quad (39)$$

holds. Then, we have

$$\begin{aligned} & \sum_{t=0}^T (1 + \gamma(t))^t \int_0^L \sum_{i=0}^n U(w,t)^i (1 - \theta_1 c(w,t)) dw = \frac{\theta_1 C_l L(t) + C_s L(t) / D_s + C_{om} PWF(i(t), T)}{\theta_2 365 f_1 P^0}. \\ & \frac{di(t)}{d\gamma(t)} = \frac{\left( \sum_{t=0}^T t (1 + \gamma(t))^{t-1} \int_0^L \sum_{i=0}^n U(w,t)^i (1 - \theta_1 c(w,t)) dw \right)}{\left( \theta_1 C_{om} / \theta_2 f_1 P^0 \right) \left( (Ti(t)(1+i(t))^T - [(1+i(t))^T - 1] [1+i(t) + Ti(t)]) / i^2(t)(1+i(t))^{T+1} \right)} > 0 \end{aligned} \quad (40)$$

In other words, IRR  $i^*(t)$  is a strictly increasing function of the break-even growth rate of the total population  $\gamma^*(t)$ .  $\square$

Proposition 4 shows that rail project is more worthy of investment with a high growth rate of total population.

#### 4. Numerical Examples

To facilitate the presentation of the essential ideas and contributions of this study, two illustrative examples are employed. Example 1 presents the year-by-year design of a rail line with the proposed conditions of three alternatives, in terms of rail length. Sensitivity analyses are conducted on the spatial and temporal covariance of population distribution, to investigate its effect on the standard deviation of rail length. Example 2 presents the year-by-year conditions for fast-tracking or deferring a rail project in a given toy network, in terms of average break-even growth rate of the total population and average IRR.

*4.1. Example 1.* The rail configuration is shown in Figure 1. For simplicity without loss of generality, the average station spacing is 1.1km [30]. The mean and standard values of year-by-year break-even rail length are determined in this numerical example.

The stochastic residential/household distribution function is defined as

$$\varphi(P(x,t)) = \frac{P(x,t)}{100}. \quad (41)$$

Other input notation parameters are summarized in Table 4.

Two scenarios with different spatial and temporal covariance of population distribution are investigated to explore the effects of spatial and temporal covariance on the values of

rail length. In this example, it is assumed that one has the following:

*Scenario a:*

$$\begin{aligned} \sigma_p(x_1, t_1; x_1, t_2) &= 0.2, \\ \sigma_p(x_1, t_1; x_2, t_1) &= 0.3, \\ \sigma_p(x_1, t_1; x_2, t_2) &= 0.1 \end{aligned} \quad (42)$$

*Scenario b:*

$$\begin{aligned} \sigma_p(x_1, t_1; x_1, t_2) &= 0.3, \\ \sigma_p(x_1, t_1; x_2, t_1) &= 0.4, \\ \sigma_p(x_1, t_1; x_2, t_2) &= 0.1 \end{aligned} \quad (43)$$

In each scenario, two alternatives are considered, shown as follows.

*Alternative 5.* Fast-tracking rail transit line project two years: the average IRR  $i^*$  is set as 10%, average break-even growing rate  $\gamma^*$  is 0.2, average total population growth rate for fast-tracking this rail transit line project  $\gamma_f$  by two year is 0.3, the average IRR for fast-tracking the rail project two years is 2%, and planning and operation horizon  $T$  are 3.

*Alternative 6.* Deferring rail transit line project three years: the average IRR  $i^*$  is set as 10%, average break-even growth rate  $\gamma^*$  is 0.2, average IRR for deferring this rail transit line project three years  $i_d$  is 31%, average growth rate of the total population for deferring this rail project three years  $\gamma_d$  is 0.1, and planning and operation horizon  $T$  are 3.

Figure 2 summarizes the results of year-by-year rail length under alternatives of fast-tracking it two years and deferring it three years. Both mean values and standard deviation

TABLE 4: Notations.

Symbol	Definition	Value
$T$	Planning and operation horizon (Years)	–
$i(t)$	Interest rate	–
$\gamma(t)$	Growing factor of total population density along the transportation corridor	–
$x$	Distance between residential location and CBD	–
$D_s(t)$	Average station spacing (HK\$)	10
$f_0$	Fixed component of fare for using the rail service (HK\$)	2.5
$f_r$	Variable component of fare per unit distance for using the rail service (HK\$/km)	0.4
$h(q(x,t))$	Headway of train operation during morning peak-hour in year $t$ (Minutes)	–
$L(t)$	Rail length in year $t$ (Km)	20
$\varepsilon$	Random term of perturbed population distribution function	–
$\varepsilon_1$	Random term of perturbed travel demand function	–
$\varepsilon_2(x,t)$	Random term of perturbed disutility function	–
$\theta_1$	Sensitivity parameter in travel demand function	0.002
$\theta_2$	Sensitivity parameter in probability function of choosing to live at residential locations	0.03
$\alpha$	Parameter of waiting cost	0.5
$\alpha_1$	Parameter of rent function	8
$\alpha_2$	Commonality factor	0.5
$\beta_1$	Parameter of rent function	2
$\rho$	Average annual number of trips to the CBD per household	365
$\eta$	Average daily number of trips to the CBD per household	1.0
$\mu_c/\mu_w/\mu_i$	Parameters for travel cost function	80/100/50
$C_l$	construction cost for rail line per kilometer (million HK\$)	8.76
$C_s$	construction cost of each rail station (million HK\$)	13.14
$C_{om}$	Annual operation and maintenance cost of the rail transit line project (million HK\$)	137.97

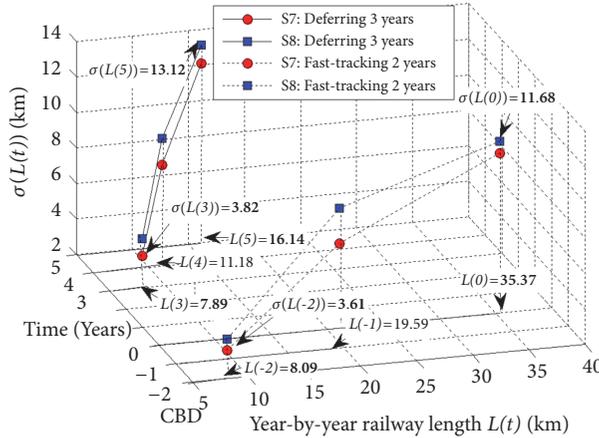


FIGURE 2: Year-by-year rail length with respect to alternatives of fast-tracking this rail transit line project two years and deferring it three years under Scenarios a and b.

values of rail length can be determined year by year with the proposed model.

It can be found that the standard rail length deviation values  $\sigma(L(t))$  increase, as the spatial and temporal covariance of population distribution increases from Scenario a to Scenario b. This implies that while correlationship between the year-by-year population densities distributed along the

candidate rail transit line strengthens, the standard deviation values of rail length increases.

The mean value of rail length under the alternative of fast-tracking two years is larger than that under the alternative of deferring three years. However, the corresponding standard deviation of rail length  $\sigma(L(t))$  under the alternative of fast-tracking two years is smaller than that under alternative of

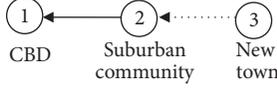


FIGURE 3: The proposed toy network.

deferring three years. For instance,  $L(-2) = 8.09$  km is larger than  $L(3) = 7.89$  km, while  $\sigma(L(-2)) = 3.61$  km is smaller than  $\sigma(L(3)) = 3.82$  km.  $L(0) = 35.37$  km is larger than  $L(5) = 16.14$  km, while  $\sigma(L(0)) = 11.68$  km is smaller than  $\sigma(L(5)) = 13.12$  km.

**4.2. Example 2.** A toy network with three locations, Central Business District (CBD), suburban community, and new town, is used to demonstrate the application of the proposed propositions, as shown in Figure 3. In the base year, a rail exists between locations CBD and suburban community, and a candidate rail extension project is to be implemented between locations suburban community and new town. The distances between locations CBD and suburban community and CBD and new town are 6km and 20km, respectively.

The mean value of population distribution at each location in the base year and the travel demand by rail from suburban community (denoted as location 2) and new town (denoted as location 3) to CBD (denoted as location 1) are set as

$$\begin{aligned} P(1, 0) &= 5000 \text{ (persons)}, \\ P(2, 0) &= 15000 \text{ (persons)}, \\ P(3, 0) &= 15000 \text{ (persons)}. \end{aligned} \quad (44)$$

Other input notation parameters are same with example 1.

Six scenarios, presented below, are used to examine the conditions for fast-tracking or deferring the candidate rail transit line project shown in the Figure 3.

**Scenario 1.** The year-by-year average break-even growth rate  $\gamma^*(t)$  with respect to different average interest rates  $i(t)$  are given in Table 5. It can be seen that the year-by-year average break-even population growth rate  $\gamma^*(t)$  increases, while the given average interest rate  $i(t)$  increases. This result is in accordance with Proposition 4.

**Scenario 2.** The year-by-year average IRR  $i^*(t)$ , with respect to different population growth rates  $\gamma(t)$ , are given in Table 6. For instance, with an average population growth rate  $\gamma$  of 0.3, this rail transit line project is not worthy of investment in year 1, during which period the average interest rate is larger than 2.6%. As time goes by, investment in this rail transit line project in years 2 and 3 is worthwhile as the respective average interest rate is smaller than 7.9% and 13.8%.

**Scenario 3.** The year-by-year conditions for fast-tracking the rail project, in terms of break-even growth rate of the total population  $\gamma_f(t)$ , are given in Table 7. The average IRR  $i^*(t)$  is set as 10%, the average break-even growth rate  $\gamma^*$  is 0.3, the average interest rate for fast-tracking this rail project  $i_f$  is 5%, and planning and operation horizon  $T$  are 3.

TABLE 5: Year-by-year average break-even growing factor of the total population with respect to different average interest rates.

Average interest rate $i$	Average break-even growth factor $\gamma^*$ in each year		
	Year = 1	Year = 2	Year = 3
3%	<b>0.12</b>	<b>0.17</b>	0.26
6%	<b>0.14</b>	<b>0.27</b>	0.30
9%	<b>0.25</b>	<b>0.29</b>	0.41

TABLE 6: Average IRR of the rail transit line project with respect to different growth factor of the total population over years.

Average growth factor $\gamma$	Average IRR $i^*$ in each year (%)		
	Year = 1	Year = 2	Year = 3
0.2	2.4%	7.3%	8.5%
<b>0.3</b>	<b>2.6%</b>	<b>7.9%</b>	<b>13.8%</b>
0.4	2.8%	12.8%	14.9%

TABLE 7: Conditions for fast-tracking this rail transit line project in Scenario 3.

Fast-tracking time (Year)	Conditions $\gamma_f(t)$ in each year $t$ during planning and operation horizon $T$		
	$t = 1$	$t = 2$	$t = 3$
1	<b>0.2684</b>	<b>0.2878</b>	<b>0.2938</b>
2	0.3327	0.3168	0.3097
3	0.3821	0.3421	0.3243

TABLE 8: Conditions for fast-tracking the rail transit line project in Scenario 4.

Fast-tracking time (Year)	$t$	Conditions $i_f(t)$ in each year during planning and operation horizon $T$		
		$t = 1$	$t = 2$	$t = 3$
1		16.12%	9.98%	<b>2.10%</b>
2		3.48%	<b>-10.34%</b>	<b>-20.76%</b>

Table 7 shows the results of these conditions. It can be seen that the year-by-year average population growth rate  $\gamma_f(t)$  for fast-tracking this rail project 1 year is below 0.3 and increases year by year from 0.2684 to 0.2938 during the planning and operation horizon  $T$ . The break-even population growth rates  $\gamma_f(t)$  for fast-tracking this rail project 2 or 3 years are above 0.3 and decreases year by year during the planning and operation horizon  $T$ .

**Scenario 4.** The year-by-year conditions for fast-tracking the rail transit line project, in terms of average IRR  $i_f(t)$ , are given in Table 8. The average IRR  $i^*(t)$  is set as 10%, the average break-even growth rate  $\gamma^*(t)$  is 0.2, the average population growth rate for fast-tracking this rail project  $\gamma_f$  is 0.3, and the planning and operation horizon  $T$  are 3.

TABLE 9: Conditions for deferring this railway project in Scenario 5.

Deferring time (Year)	Conditions $\gamma_d(t)$ in each year $t$ during planning and operation horizon $T$		
	$t = 1$	$t = 2$	$t = 3$
1	<b>-0.2144</b>	-0.03023	0.05611
2	-0.0519	0.0601	0.1125
3	0.1121	0.1512	0.1695

TABLE 10: Conditions for deferring this railway project in Scenario 6.

Deferring time (Year)	Conditions $i_d(t)$ in each year $t$ during planning and operation horizon $T$		
	$t = 1$	$t = 2$	$t = 3$
1	<b>68.74%</b>	72.98%	78.40%
2	<b>37.86%</b>	40.43%	43.64%
3	<b>26.22%</b>	28.22%	30.62%

It can be seen that year-by-year average IRR(s) for fast-tracking the rail transit line project 1 year in this scenario are all positive and decrease year by year from 16.12% to 2.10% during planning and operation horizon  $T$ . For instance, this rail transit line project is not worth fast-tracking 1 year, if the interest rate in year 3 is greater than 2.10%. Table 8 shows that fast-tracking this rail transit line project 2 years is not worthwhile, because the average IRR(s) in years 2 and 3 are negative.

*Scenario 5.* The year-by-year conditions for deferring the rail transit line project, in terms of average population growth rate  $\gamma_d(t)$ , are given in Table 9. The average IRR  $i^*$  is set as a 10%, average break-even growth rate  $\gamma^*$  is 0.3, average interest rate suffered for deferring this rail transit line project  $i_d$  is 20%, and planning and operation horizon  $T$  are 3.

Table 9 presents the results of Scenario 5. It shows that the break-even growth rate for deferring this rail project  $\gamma_d(t)$  increases with the increase of deferring time. For instance, in year  $t = 1$ ,  $\gamma_d(t)$  increases from -0.2144 to 0.1121 as deferring time increases from 1 to 3.

*Scenario 6.* The year-by-year conditions for deferring this rail project, in terms of average IRR  $i_d$ , are given based on Proposition 4 shown in Table 10. The average IRR  $i^*$  is set as 10%, the average break-even growth rate  $\gamma^*$  is 0.2, the average population growth rate for deferring this rail project  $\gamma_d$  is 0.3, and planning and operation horizon  $T$  are 3.

Table 10 shows that the year-by-year average IRR(s) for fast-tracking this rail transit line project  $i_d(t)$  are all larger than the average IRR  $i^*(t)$  10% and decrease year by year during the planning and operation horizon  $T$ . For instance, IRR(s) decreases from 68.74% to 26.22% in the first year of the planning and operation horizon.

### 5. Conclusions and Future Work

This paper proposes models for multiperiod flexible rail line design in a linear monocentric city. The proposed models

consider the effects of year-by-year variation of the total population and spatial-temporal correlation of population distribution explicitly. In contrast with the traditional single-period NDP models for rail transit line, the proposed models have the following merits: the candidate rail project can be fast-tracked or deferred; conditions for fast-tracking or deferring the implemented rail project are analytically investigated.

The proposed models offer several insights. For example, the mean values of rail design variables are closely concerned with interest rates and the total population growth rate. The spatial and temporal covariance of population distribution only affects the standard deviation values of the rail length. The rail transit line project can be fast-tracked or deferred when actual interest rates or actual population growth rate match break-even IRR and break-even growth rate.

This paper provides a new avenue for the modeling and analysis of flexible rail transit line design in a linear monocentric city. Further research is needed in the following directions:

(1) The monocentric city is assumed in this paper, namely, only with one CBD and several other residential locations. The city boundary is not explicitly considered. Therefore, it is necessary to elaborate the city boundary so as to extend the polycentric CBD model in a further study.

(2) In this paper, population are assumed to be homogeneous with trips commuting only from residences to CBD. However, previous studies have shown that income levels dominate residential location choices (Hartwick et al., 1976; Kwon, 2003). Therefore, the proposed model could be extended to incorporate household income levels over the years for determining residential location choices and population distribution.

(3) In this paper, the investigation of investment risk is based on interest rates and temporal and spatial population covariance and variations of population distribution year by year. However, there are many other investment risk sources, for instance, client investment risk related to private operators for allocated government projects. A detailed model should be developed to take into account projects with such contents, since the performance of related projects are influenced considerably.

(4) Only rail mode is considered in this paper. This assumption can be extended to a multimodal situation in further studies. With more travel modes being considered, travelers' travel mode choice behaviour can be incorporated into the extended models (Chowdhury and Chien, 2002; Li et al., 2006) [37].

(5) The decision to extend a rail line involves consideration of technological, social, and economic factors. The prime reason could be social or in other words a desire to make life more convenient as regards manoeuvrability for a specific set of people, namely, those living in the vicinity of the line and new stations to be constructed. However, only pressing economic factor is considered in this paper. More detailed social factors can be taken into account in further studies, for instance, appreciation of land value along the rail line.

## 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 work described in this paper was jointly supported by the National Natural Science Foundation of China (Grant no. 71473060). The authors would like to thank Professor W.H.K. Lam for his comments and suggestions and Mrs. Elaine Anson for her proofreading of this manuscript.

## References

- [1] T. Liu, H. Huang, H. Yang, and X. Zhang, "Continuum modeling of park-and-ride services in a linear monocentric city with deterministic mode choice," *Transportation Research Part B: Methodological*, vol. 43, no. 6, pp. 692–707, 2009.
- [2] Z.-C. Li, W. H. K. Lam, S. C. Wong, and K. Choi, "Modeling the effects of integrated rail and property development on the design of rail line services in a linear monocentric city," *Transportation Research Part B: Methodological*, vol. 46, no. 6, pp. 710–728, 2012.
- [3] Z. Li, W. H. K. Lam, and S. C. Wong, "Modeling intermodal equilibrium for bimodal transportation system design problems in a linear monocentric city," *Transportation Research Part B: Methodological*, vol. 46, no. 1, pp. 30–49, 2012.
- [4] C. Tang, A. A. Ceder, and Y.-E. Ge, "Integrated optimization of bus line fare and operational strategies using elastic demand," *Journal of Advanced Transportation*, vol. 2017, 15 pages, 2017.
- [5] S. T. Waller, J. L. Schofer, and A. K. Ziliaskopoulos, "Evaluation with traffic assignment under demand uncertainty," *Transportation Research Record*, no. 1771, pp. 69–74, 2001.
- [6] Y. Zhao and K. M. Kockelman, "The propagation of uncertainty through travel demand models: an exploratory analysis," *Annals of Regional Science*, vol. 36, no. 1, pp. 145–163, 2002.
- [7] T. Yip, E. Hui, and R. Ching, "The impact of railway on home-moving: Evidence from Hong Kong data," Working paper, 2012.
- [8] D. Liu, "Modelling the effects of spatial and temporal correlation of population densities in a rail transportation corridor," *Journal of Transport & Infrastructure Research*, vol. 15, no. 2, pp. 243–261, 2015.
- [9] H. Shao, W. H. K. Lam, A. Sumalee, and M. L. Hazelton, "Estimation of mean and covariance of stochastic multi-class OD demands from classified traffic counts," *Transportation Research Part C: Emerging Technologies*, vol. 59, pp. 92–110, 2015.
- [10] X. Hu, Y.-C. Chiu, J. A. Villalobos, and E. Nava, "A Sequential Decomposition Framework and Method for Calibrating Dynamic Origin-Destination Demand in a Congested Network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 10, pp. 2790–2797, 2017.
- [11] Y. Yin, S. M. Madanat, and X. Lu, "Robust improvement schemes for road networks under demand uncertainty," *European Journal of Operational Research*, vol. 198, no. 2, pp. 470–479, 2009.
- [12] L. Gao, Y. Ji, Y. Liu et al., "Research on Modeling Intrahousehold Interactions from the Perspective of Space-Time Constraints," *Discrete Dynamics in Nature and Society*, vol. 2017, 11 pages, 2017.
- [13] X. Fu, W. H. K. Lam, and Y. Xiong, "Modelling intra-household interactions in household's activity-travel scheduling behaviour," *Transportmetrica A: Transport Science*, vol. 12, no. 7, pp. 612–628, 2016.
- [14] D. Liu, "Analytical forecasting of population distribution over years in a new rail transportation corridor," *Journal of Urban Planning and Development*, vol. 142, no. 4, 2016.
- [15] H. Lo and W. Y. Szeto, "Planning transport network improvements over time," in *Urban and Regional Transportation modeling: Essays in Honor of David Boyce*, H. D. Lee, Ed., pp. 157–176, Edward Elgar, 2004.
- [16] W. Y. Szeto and H. K. Lo, "Strategies for road network design over time: Robustness under uncertainty," *Transportmetrica*, vol. 1, no. 1, pp. 47–63, 2005.
- [17] W. Y. Szeto and H. K. Lo, "Time-dependent transport network improvement and tolling strategies," *Transportation Research Part A: Policy and Practice*, vol. 42, no. 2, pp. 376–391, 2008.
- [18] W. Y. Szeto and H. K. Lo, "Transportation network improvement and tolling strategies: the issue of intergeneration equity," *Transportation Research Part A: Policy and Practice*, vol. 40, no. 3, pp. 227–243, 2006.
- [19] H. K. Lo and W. Y. Szeto, "Time-dependent transport network design under cost-recovery," *Transportation Research Part B: Methodological*, vol. 43, no. 1, pp. 142–158, 2009.
- [20] 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.
- [21] E. K. Morlok and D. J. Chang, "Measuring capacity flexibility of a transportation system," *Transportation Research Part A: Policy and Practice*, vol. 38, no. 6, pp. 405–420, 2004.
- [22] Y.-T. Peng, Z.-C. Li, and K. Choi, "Transit-oriented development in an urban rail transportation corridor," *Transportation Research Part B: Methodological*, vol. 103, pp. 269–290, 2017.
- [23] H.-S. J. Tsao, W. Wei, and A. Pratama, "Operational feasibility of one-dedicated-lane bus rapid transit/light rail systems," *Transportation Planning and Technology*, vol. 32, no. 3, pp. 239–260, 2009.
- [24] Z. Qian and Z. H. Michael, "Modeling multi-modal morning commute in a one-to-one corridor network," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 2, pp. 254–269, 2011.
- [25] Z.-C. Li, W. H. K. Lam, S. C. Wong, and A. Sumalee, "Environmentally sustainable toll design for congested road networks with uncertain demand," *International Journal of Sustainable Transportation*, vol. 6, no. 3, pp. 127–155, 2012.
- [26] S. Clark and D. Watling, "Modelling network travel time reliability under stochastic demand," *Transportation Research Part B: Methodological*, vol. 39, no. 2, pp. 119–140, 2005.
- [27] H. Shao, W. H. K. Lam, A. Sumalee, A. Chen, and M. L. Hazelton, "Estimation of mean and covariance of peak hour origin-destination demands from day-to-day traffic counts," *Transportation Research Part B: Methodological*, vol. 68, pp. 52–75, 2014.
- [28] M. L. Hazelton, "Some comments on origin-destination matrix estimation," *Transportation Research Part A: Policy and Practice*, vol. 37, no. 10, pp. 811–822, 2003.
- [29] M. L. Hazelton, "Statistical inference for time varying origin-destination matrices," *Transportation Research Part B: Methodological*, vol. 42, no. 6, pp. 542–552, 2008.

- [30] Z. Li, W. H. K. Lam, S. C. Wong, and A. Sumalee, "Design of a rail transit line for profit maximization in a linear transportation corridor," *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 1, pp. 50–70, 2012.
- [31] E. Cascetta, A. Nuzzolo, F. Russo, and A. Vitetta, "A modified logit route choice model overcoming path overlapping problem: specification and some calibration results for interurban networks," in *Proceedings of 13th International Symposium on Transportation and Traffic Theory*, J. B. Lesort, Ed., pp. 697–711, Pergamon, Lyon, France, 1996.
- [32] Z. Zhou, A. Chen, and S. Bekhor, "C-logit stochastic user equilibrium model: formulations and solution algorithm," *Transportmetrica*, vol. 8, no. 1, pp. 17–41, 2012.
- [33] J. N. Prashker and S. Bekhor, "Route choice models used in the stochastic user equilibrium problem: a review," *Transport Reviews*, vol. 24, no. 4, pp. 437–463, 2004.
- [34] S. Guo, L. Yu, X. Chen, and Y. Zhang, "Modelling waiting time for passengers transferring from rail to buses," *Transportation Planning and Technology*, vol. 34, no. 8, pp. 795–809, 2011.
- [35] H. W. Ho and S. C. Wong, "Housing allocation problem in a continuum transportation system," *Transportmetrica*, vol. 3, no. 1, pp. 21–39, 2007.
- [36] B. Van Wee, "Rail Infrastructure: Challenges for Cost - Benefit Analysis and Other ex ante Evaluations," *Transportation Planning and Technology*, vol. 30, no. 1, pp. 31–48, 2007.
- [37] D. Liu and W. H. K. Lam, "Modeling the effects of population density on prospect theory-based travel mode-choice equilibrium," *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, vol. 18, no. 4, pp. 379–392, 2014.

## Research Article

# Optimization of Vehicle Routing with Pickup Based on Multibatch Production

Hongtao Hu,<sup>1</sup> Jiao Mo ,<sup>1</sup> and Chengle Ma <sup>2</sup>

<sup>1</sup>Logistics Engineering College, Shanghai Maritime University, 201306, Shanghai, China

<sup>2</sup>School of Management, Shanghai University, 200444, Shanghai, China

Correspondence should be addressed to Chengle Ma; [ma.chengle@foxmail.com](mailto:ma.chengle@foxmail.com)

Received 1 August 2018; Accepted 11 October 2018; Published 1 November 2018

Guest Editor: Xinchang Wang

Copyright © 2018 Hongtao 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.

To reduce the inventory cost and ensure product quality while meeting the diverse demands of customers, manufacturers yield products in batches. However, the raw materials required for manufacturing need to be obtained from suppliers in advance, making it necessary to understand beforehand how to best structure the pickup routes so as to reduce the cost of picking up and stocking while also ensuring the supply of raw materials required for each batch of production. To reduce the transportation and inventory costs, therefore, this paper establishes a mixed integer programming model for the joint optimization of multibatch production and vehicle routing problems involving a pickup. Following this, a two-stage hybrid heuristic algorithm is proposed to solve this model. In the first stage, an integrated algorithm, combining the Clarke-Wright (CW) algorithm and the Record to Record (RTR) travel algorithm, was used to solve vehicle routing problem. In the second stage, the Particle Swarm Optimization (PSO) algorithm was used to allocate vehicles to each production batch. Multiple sets of numerical experiments were then performed to validate the effectiveness of the proposed model and the performance efficiency of the two-stage hybrid heuristic algorithm.

## 1. Introduction

The “truck dispatching problem” was first introduced by Dantzig and Ramser [1]. Then Clarke and Wright [2] generalized the “vehicle routing problem” (VRP) to the domain of logistics and transport; that is, to how to serve a set of customers who are geographically dispersed around the central depot, using a fleet of trucks with varying capacities. Taking into account real-life complexities, this paper expands the traditional VRP by combining VRP with production, in order to consider the vehicle routing problem with raw material pickup under multibatch production (VRPPMP). The problem at hand is defined as follows. The manufacturer divides the production in each cycle into multiple batches and uses vehicles to undertake the pickup from their suppliers, in order to ensure sufficient raw materials prior to the production of each batch. Since the inseparable pickup demands of dispersed suppliers must be met, the manufacturer thus needs to design both the vehicle pickup routes and the batch allocation of the routes in such a way that they satisfy the demand for the raw materials needed for production, while also minimizing the total inventory and transportation costs.

By designing the vehicle pickup routes in the context of multibatch production, the manufacturer can reduce transportation and inventory costs in the production process and ensure product quality. In the automobile assembly industry, for example, automobile producers need to retrieve the parts needed from the auto parts manufacturer, separately. If the producers retrieve all the parts needed for production during each period, this not only will greatly increase the inventory cost, but also cannot guarantee the quality of the parts. Therefore, multibatch production is needed for automobile assembly production, with the VRPPMP of automobile assembly production being investigated in order to plan the vehicle pickup path for each batch, so as to ensure the quality of the raw material and the products and reduce the costs of inventory and transportation.

Prior to initiating periodic production, the manufacturer must pick up raw materials from geographically dispersed locations (suppliers), with the raw materials being picked up periodically at production batches. In single-cycle production planning, there are often multiple batches of production operations. Before each batch of production and processing

begins, the manufacturer needs sufficient raw materials in order to complete the set production volume. If the purchase quantity of raw materials is too large, it will lead to a certain degree of inventory accumulation. This will, in turn, generate the problems of a high inventory management cost and the low quality of raw materials. Thus, according to the production demands of each batch, the problem lies in there being a reasonable arrangement of vehicle pickup routes, and also in matching these routes with production batches in order to minimize the total costs of transportation and inventory.

In line with the above description, this paper proposes an optimization of vehicle routing involving pickup based on multibatch production. The contributions of this paper are as follows:

(1) Combining production with the vehicle pickup routing problem, this paper studies the optimization of vehicle routing involving pickup in the context of multibatch production. We design vehicle pickup routes based on the raw material demands of each batch in such a way as to minimize transportation costs and inventory costs while meeting production demands.

(2) We put forward a two-stage hybrid heuristic algorithm to solve the proposed model. In the first stage of the algorithm, the Clarke-Wright (CW) algorithm is proposed to generate the initial vehicle routes, following which the Record to Record (RTR) travel algorithm is used to further optimize the generated routes. In the second stage, a Particle Swarm Optimization (PSO) algorithm is developed to allocate the vehicles to each production batch.

The remainder of this paper is organized as follows. Section 2 reviews the related literature. Section 3 details the vehicle routing involving pickup based on multibatch production and develops the mixed integer programming model. Section 4 introduces a two-stage hybrid heuristic algorithm to solve this model. In Section 5, multiple sets of numerical experiments are performed to validate the effectiveness of the proposed model and the performance efficiency of the two-stage hybrid, heuristic algorithm. Section 6 presents the conclusions of this study and recommendations for future work.

## 2. Literature Review

The vehicle path problem (VRP) is a key aspect of logistics and transportation research. Selecting the appropriate vehicle path method can speed up the response to customer demands and reduce operation costs. Since first being proposed by Dantzig and Ramser [1], research on VRP has extended into numerous avenues after decades of exploration and research, such as the capacitated vehicle routing problem (CVRP) [3], the vehicle routing problem with time windows (VRPTW) [4], and vehicle routing problems with pickup and delivery (VRPPD) [5, 6], among others. This paper focuses on the angle of VRP with pickup and vehicle capacity limitation.

Many scholars have conducted in-depth research on the issues of VRPPD. Savelsbergh and Sol [7] discussed several characteristics distinguishing the latter from standard VRP and presented a survey of the problem types and solution

methods. A VRPPD with multiple time windows and vehicle types was considered by Xu et al. [8], where computational results showed that the computational times required by the proposed column-generation-based solution were acceptable. Gábor Nagya [9] proposed a heuristic algorithm that takes the pickup and delivery stages as a whole, and which was used to solve the VRPPD in the context of single and multiple car parks. Kalina and Vokrinek [10] presented a parallel solver for VRPTW and the pickup and delivery problem with time windows, which was based on the parallel competition of particular solvers solving the given problem.

The problem of combining VRPPD with production, which is a type of VRPPD, has been studied extensively. Such problems involve coordinating production, inventory, and delivery operations to meet customer needs and minimize costs. Zheng et al. [11] investigated the vendor-managed, cyclic inventory routing problem under constant customer demand rates. To minimize inventory costs without causing any lack of stock at the customers, a heuristic solution approach was proposed. The latter proved well capable of finding the appropriate cost trade-off under varying circumstances. Shiguemoto and Armentano [12] addressed the problem of optimally coordinating a production-distribution system with a fleet of homogeneous vehicles over a multiperiod finite horizon, proposing a Tabu search procedure for solving the problem. Adulyasak et al. [13] introduced multivehicle production routing problem and inventory routing problem formulations, with and without a vehicle index, to solve these problems under both the maximum level and the order-up-to level inventory replenishment policies. To minimize the total distribution cost, Sainathuni et al. [14] introduced the warehouse inventory transportation problem of determining an optimal distribution plan from vendors to customers via one or more warehouses.

Although the research on combining VRPPD with production has made some progress, there is still some room for development. To date, existing studies that have considered combining multibatch production with vehicle pickup routing issues have been rather limited. To enrich the research in this area, based on this combination, the current paper considers two aspects from the perspective of the manufacturer: (i) VRP – the vehicle routing problem from an upstream supplier to the manufacturer's procurement link and (ii) production and processing problems, considering the single-cycle and multibatch production and processing problems of a single manufacturer, how the allocation of vehicles can best be undertaken so as to correspond to the production batches. Under the premise of completing each batch of production and processing operations, the inventory cost should be as low as possible.

## 3. Problem Description and Modeling

*3.1. Problem Description.* The VRPPMP can be defined as follows. Let  $G = (S, E)$  be a complete oriented graph with a set of vertices  $S' = \{0, 1, \dots, |S'|\}$ , where the vertex 0 represents the manufacturer and the remaining ones represent the suppliers. Each edge  $\{i, j\} \in E$  for  $i, j \in S'$ , has a nonnegative cost,  $C_{ij}$ , and each supplier  $i \in S = S' - \{0\}$  has known

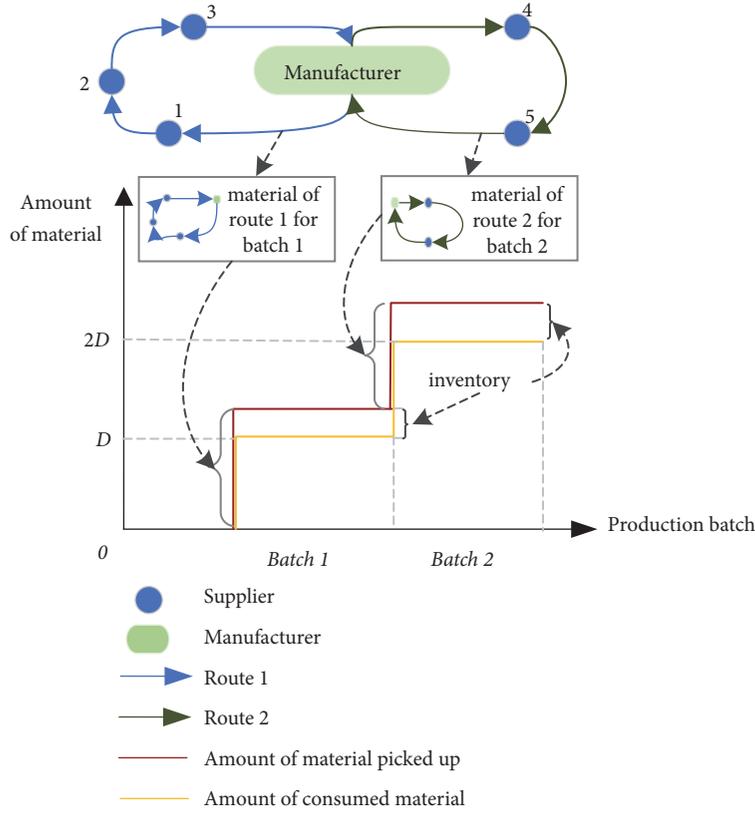


FIGURE 1: A schematic diagram of the VRPPMP.

nonnegative demands,  $p_i$ , with regard to the pickup. Let  $T = (1, \dots, t, \dots, |T|)$  be a set of production batches.  $D$  represents the manufacturer's fixed output for each production batch. Let  $V = \{1, \dots, m, \dots, |V|\}$  be a set of homogeneous vehicles with capacity  $Q$ . The VRPPMP is depicted in Figure 1. The vehicles visit the corresponding suppliers for each production batch and then return to the manufacturer. The retrieved raw material is put into production. If there is any unfinished raw material, this will be used as raw material inventory for the next batch of production. For each batch of access routes, the VRPPMP consists of constructing a structure for the routes in such a way that (i) every route starts and ends at the manufacturer; (ii) all pickup demands are met; (iii) a supplier is visited only by a single vehicle; and (iv) the sum of inventory costs and vehicle transportation costs is minimized.

In addition, this paper makes the following assumptions underlying the VRPPMP:

(1) The manufacturer conducts multiple batches of production and processing operations, with the total amount of materials supplied by the suppliers equaling the total amount acquired for production.

(2) The remaining raw material inventory of each batch can be used for the next batch of production tasks; the inventory capacity of the manufacturer is not considered.

(3) The raw materials for each batch can be delivered to the manufacturer in time for the batch to be produced.

(4) The problem considers vehicle capacity limitations.

In the model, we also assumed that each vehicle can only perform one trip for each production batch. It should be

noted that a vehicle's multiple trips for a production batch (the multitrip vehicle routing problem, VRPMT), and the consistency of a vehicle in terms of pickup from material suppliers for different batches (the consistent vehicle routing problem, CVRP), are not considered. In real production activities, the production plan is often based on a day, a week, or a longer time interval, while the vehicles can complete a trip within a relatively shorter time span. Therefore, if a vehicle performs more than one trip for different batches, it can be considered several dummy vehicles.

### 3.2. Mathematical Model

#### Parameters

$S$ : Set of suppliers,  $S = \{1, 2, \dots, i, \dots, |S|\}$

$S^+$ : 0 represents trucks departing from the manufacturer,  $S^+ = S \cup \{0\}$

$S^-$ :  $|S|+1$  represents vehicles arriving to the manufacturer,  $S^- = S \cup \{|S| + 1\}$

$T$ : Set of production batches,  $T = \{1, 2, \dots, t, \dots, |T|\}$

$V$ : Set of vehicles,  $V = \{1, 2, \dots, m, \dots, |V|\}$

$p_i$ : The amount of raw material for pickup from supplier  $i$

$c_{ij}$ : The cost of transportation from supplier  $i$  to supplier  $j$

$c^{unpro}$ : The inventory cost for each unit of unprocessed raw material during unit production batch

$Q$ : The capacity of homogeneous vehicles

$D$ : The amount of planned product in each production batch

$M$ : A sufficient large positive number.

#### Decision Variables

$x_{im}$ : 1 if vehicle  $m$  provides supplier  $i$  with pickup service; 0 otherwise

$z_{mt}$ : 1 if vehicle  $m$  is assigned to the production batch  $t$ ; 0 otherwise

$y_{ijm}$ : 1 if suppliers  $i$  and  $j$  are successively serviced by vehicle  $m$ ; 0 otherwise

$g_{mt}$ : The amount of raw material transported by vehicle  $m$  for production batch  $t$

$q_{im}$ : The total amount of raw material transported by vehicle  $m$  after servicing supplier  $i$

$\varphi_t$ : The remaining inventory of the raw materials after the  $t$ th production batch

Based on the above definition, the VRPPMP is established as follows:

$$\min \left( \sum_{i \in S^+} \sum_{j \in S^-} \sum_{m \in V} c_{ij} \cdot y_{ijm} + c^{unpro} \cdot \sum_{t \in T} \varphi_t \right) \quad (1)$$

$$s.t. \quad x_{0m} = 1 \quad \forall m \in V \quad (2)$$

$$\sum_{m \in V} x_{im} = 1 \quad \forall i \in S \quad (3)$$

$$\sum_{i \in S^+, i \neq h} y_{ihm} = \sum_{j \in S^-, j \neq h} y_{hjm} = x_{hm} \quad (4)$$

$$\forall h \in S, \forall m \in V$$

$$\sum_{j \in S} y_{0jm} - \sum_{t \in T} z_{mt} = 0 \quad \forall m \in V \quad (5)$$

$$\sum_{t \in T} z_{mt} \leq 1 \quad \forall m \in V \quad (6)$$

$$g_{mt} \leq Q \cdot z_{mt} \quad \forall t \in T, \forall m \in V \quad (7)$$

$$g_{mt} \leq \sum_{j \in S} p_j \cdot x_{jm} + Q \cdot (1 - z_{mt}) \quad (8)$$

$$\forall t \in T, \forall m \in V$$

$$g_{mt} \geq \sum_{j \in S} p_j \cdot x_{jm} - Q \cdot (1 - z_{mt}) \quad (9)$$

$$\forall t \in T, \forall m \in V$$

$$\sum_{m \in V} g_{mt} \geq D \quad t = 1 \quad (10)$$

$$\sum_{m \in V} g_{mt} + \varphi_{t-1} \geq D \quad \forall t \in T / \{1\} \quad (11)$$

$$q_{jm} \geq p_j - Q \cdot (1 - y_{0jm}) \quad \forall j \in S, \forall m \in V \quad (12)$$

$$q_{jm} \leq p_j + Q \cdot (1 - y_{0jm}) \quad \forall j \in S, \forall m \in V \quad (13)$$

$$q_{jm} \geq q_{im} + p_j - Q \cdot (1 - y_{ijm}) \quad (14)$$

$$\forall i \in S, \forall j \in S^-, i \neq j, \forall m \in V$$

$$q_{jm} \leq q_{im} + p_j + Q \cdot (1 - y_{ijm}) \quad (15)$$

$$\forall i \in S, \forall j \in S^-, i \neq j, \forall m \in V$$

$$\varphi_t = \sum_{m \in V} g_{mt} - D \quad t = 1 \quad (16)$$

$$\varphi_t = \varphi_{t-1} + \sum_{m \in V} g_{mt} - D \quad \forall t \in T / \{1\} \quad (17)$$

Objective function (1) maximizes the total costs of the manufacturer, which are equal to the sum of the inventory costs and shipping costs incurred. Constraints (2) state that each vehicle starts at the manufacturer and ends at the manufacturer. Constraints (3) ensure that each supplier will be accessed once. Flow conservation is ensured by Constraints (4). Constraints (5) indicate that each vehicle that provides suppliers with a pickup service must serve in exactly one production batch. Constraints (6) state that each vehicle can, at most, serve at most one production batch. Constraints (7) ensure that the raw material transported by each vehicle does not exceed the capacity of vehicle. Constraints (8) and (9) represent the total amount of raw material delivery by each vehicle in each production batch. Constraints (10) and (11) indicate the constraint of amount of raw materials that can be put into production in each batch. Constraints (12) to (15) state the relationship between the vehicle load and the quantity supplied. Constraints (16) and (17) represent the raw material inventory after each batch of production and processing.

#### 4. Solution Algorithm

The above model can be directly solved by using commercial optimization packages, such as CPLEX, when the scale is small. However, when the scale increases by a certain extent, using CPLEX is time consuming or unable to complete the solution. Therefore, the algorithm needs to be added to solve large-scale problems.

Taking the two aspects into account when making a decision, that is, vehicle routing as well as the matching between the vehicle and the production batch, the VRPPMP mentioned above cannot be solved directly. In light of this, the VRPPMP was divided into two stages. In the first stage, the problem of vehicle routing in the process of picking up cargoes was solved using the Clarke-Wright and the Record-to-Record travel algorithms. In the second stage, we used the PSO algorithm to allocate each vehicle that needs to be picking up to a certain production batch so as to satisfy the demands of the manufacturer in each production batch.

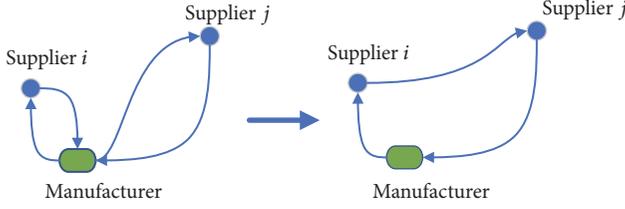


FIGURE 2: Changes in vehicle pickup routes.

**4.1. Routes Initialization (CW).** For the purposes of this paper, the CW algorithm was used to generate the initial vehicle path for the first stage problem. The algorithm was first proposed by Clarke and Wright [2] and is applicable to CVRP. It forms a route in the following ways. First, it yields the shortest route between every two suppliers in the system. The desired aim here is to allocate loads to vehicles in such a manner that all the raw materials are assigned and the total mileage covered is at a minimum.

Consider a general savings value,  $S$ , given by linking two suppliers  $i$  and  $j$  as shown in Figure 2:

$$s_{ij} = d_{0i} + d_{0j} - \lambda d_{ij} \quad (18)$$

Where

$s$  = savings value;

$d_{0i}$  = distance from supplier  $i$  to the manufacturer;

$d_{ij}$  = distance between points  $i$  and  $j$ ;

$\lambda$  = route shape parameter. Special cases:  $\lambda = 1$ , the savings approach, and  $\lambda = 2$ , Gaskell's  $\pi$  method [15]. With an increase in  $\lambda$ , greater emphasis is placed on the distance between the suppliers, rather than on the position relative to the manufacturer in selecting an addition to a route.

In line with the savings value  $S$  between every two suppliers, the algorithm yields the CW table. In the case of satisfying the on-vehicle capacity, the manufacturer preferentially links a pair of suppliers with a larger  $S$ -value and uses one vehicle to provide pickup services for these suppliers. Following the linking method until all the suppliers are serviced, the manufacturer then obtains the initial solution for the pickup vehicle routes. The procedure given is simple but effective in producing a near-optimal solution and has been programmed for several digital computers.

**4.2. Routes Improvement (RTR).** After the initial vehicle route was generated, we further optimized it using the RTR travel method, which was first proposed by Dueck [16]. We improved the solutions using the operations specified by Li et al. ([17], namely, a one-point move and two-point move. In the one-point move, we attempted to move each supplier in the existing solution to a new position on the same route or on a different route. In the two-point move, we tried to exchange the positions of two suppliers. These improvement moves are shown in Figure 3.

The routes were improved by repeatedly applying these two local search operators in two phases: diversification and improvement. In the diversification phase, we attempted to explore new areas in the solution space by accepting both improving and deteriorating moves. In the improvement

phase, we attempted to improve the current solution as much as possible by accepting only improving moves until we reached a local minimum.

By combining the CW algorithm and RTR travel method, we were able to obtain a better pickup vehicle routing solution. Based on the routes, we then solved the decision making problems pertaining to vehicle allocation in each production batch by using the PSO algorithm, in order to satisfy the manufacturer's demands in each production batch and lower the inventory cost as far as possible.

**4.3. Batch Allocation (PSO).** The first stage enabled us to obtain all the routes that the vehicles needed to cover, then making it necessary to assign the routes to each production batch in the second stage. Consequently, the PSO algorithm developed by Eberhart and Kennedy [18] was employed to solve the route allocation problem and can be seen as an efficient stochastic global optimization technique.

In the PSO algorithm, the position of each particle represents a feasible solution in the search domain, and each particle changes its position and velocity depending on its flying experience. The decision variables  $x_{im}$ ,  $y_{jm}$ , and  $q_{im}$ , which are related to the pickup service were determined in the first stage.  $g_{mt}$  and  $\phi_t$  could be calculated according to constraints (16) and (17) as long as  $z_{mt}$  was determined in the second stage. The core variable  $z_{mt}$ , which determines whether the raw material transported by vehicle  $m$  was assigned to production batch  $t$ ,  $m \in V$ ,  $t \in T$ , was coded as follows. All the utilized vehicles were grouped in set  $V'$ ,  $V' \in V$ . Each particle with  $|V'|$  dimensions was represented by  $F = \{f_1, f_2, \dots, f_m, \dots, f_{|V'|}\}$ . Each dimension  $f_m$  was coded as a positive number. After the vehicle sequence was reordered according to the nondescending order of the  $f_m$ , vehicles were assigned to the production batches. The main idea behind the distribution rule is as follows: vehicles at the front of the sorted sequence have priority in terms of assignment to the previous production batch.

For instance, if the amount of materials acquired by a production batch is  $\{10, 10, 10\}$ , the utilized vehicles would be  $\{1, 2, 3, 4, 5\}$ , and the quantity of materials corresponding to the vehicles would be  $\{4, 6, 8, 7, 5\}$ . If the sorted vehicle sequence is  $\{2, 3, 1, 5, 4\}$ , the amount of materials would be  $\{6, 8, 5, 4, 7\}$ , accordingly. Vehicle  $\{2\}$  was first assigned to production batch  $\{1\}$ , given that the amount of material acquired in this batch was 10, which is greater than the amount that vehicle  $\{2\}$  carried. The next vehicle in the sorted sequence, that is, vehicle  $\{3\}$ , was assigned to batch  $\{1\}$ , thus giving the storage material of production batch  $\{1\}$  a value of 4. Accordingly, vehicles  $\{4$  and  $5\}$  were assigned to batch  $\{2\}$  with the storage value of 3, and vehicle  $\{7\}$  was assigned to batch  $\{3\}$  with no storage material in the last.

In applying the PSO algorithm, we assumed that a swarm has  $k$  particles.  $V_k^n = \{v_{kmt}^n\}$  was the velocity of particle  $k$  at iteration  $n$ , and  $P_k^n = \{p_{kmt}^n\}$  the position of particle  $k$  at iteration  $n$ . Each particle had its personal best position,  $PLBest_{kmt}^n$ , at iteration  $n$  on dimensions  $m$  and  $t$ .  $PGBest_{kmt}^n$  is the global best position of the entire swarm on dimensions  $m$  and  $t$ , up until iteration  $n$ . The swarm flies through hyperspace according to the experience of its neighbors. The ordinary

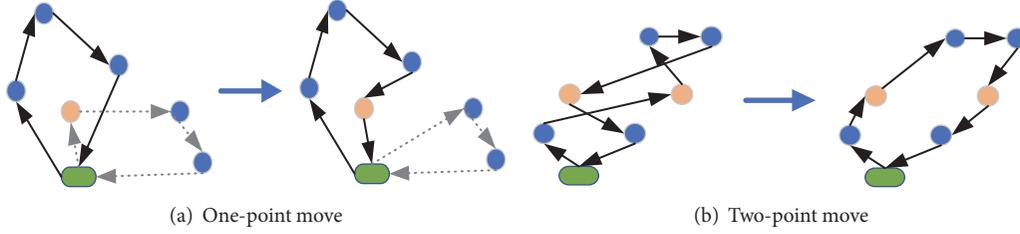


FIGURE 3: Improvement operators used in VRPPMP.

updating formulates used to calculate speed and position were as follows:

$$v_{kmt}^{n+1} = w \cdot v_{kmt}^n + c_1 r_1 (PLBest_{kmt}^n - p_{kmt}^n) + c_2 r_2 (PGBest_{kmt}^n - p_{kmt}^n) \quad (19)$$

$$p_{kmt}^{n+1} = p_{kmt}^n + v_{kmt}^{n+1} \quad (20)$$

Where  $w$  is an inertia weight parameter, which affects the convergence procedure of the optimal solution;  $c_1$  and  $c_2$  are the cognitive and social learning parameters, respectively;  $r_1$  and  $r_2$  are random numbers generated between  $[0, 1]$ .

Based on the above description, we now present the PSO algorithm process for solving the route allocation problem.

*Step 1.* Set the iteration number  $n = 1$ . Initialize  $k$  particles whose positions and velocities are randomly generated. The particles' positions represent the result of the route allocation and determine their qualities.

*Step 2.* Evaluate the fitness value of each particle using the commercial software CPLEX. More specifically, we used the route allocation result obtained by the particles to fix the variables  $z_{mt}$  and then used CPLEX to solve the original model.

*Step 3.* For each particle, compare its fitness value and the global best optimal solution  $PGBest_{kmt}^n$ ; if  $PLBest_{kmt}^n > PGBest_{kmt}^n$ , replace the value of  $PGBest_{kmt}^n$  using the value of  $PLBest_{kmt}^n$ .

*Step 4.* According to formulas (19) and (20), update the particle velocity and position.

*Step 5.* If  $n$  reaches the preset maximum iteration or  $PGBest_{kmt}^n$  has not been improved during a preset number of iterations, end the procedure; otherwise, set  $n = n + 1$  and then go on to *Step 2*.

## 5. Numerical Experiments

To validate the performance of the proposed two-stage heuristic, numerical experiments were carried out on the VRPPMP. For the PSO algorithm, the maximum number of iterations was set to 100, while the population size was 30. According to the previous tests, we set  $c_1 = c_2 = 1$ . From preliminary testing of  $w = 0, 0.5, 1, 1.5,$  and  $2$ , it emerged that  $w = 1$  performed the best; thus, we adopted  $w = 1$  for

the next experiments. Based on the same decision obtained in the first stage, another nature inspired algorithm, the Tabu search (TS) algorithm, was applied for comparison purposes. We listed the optimal solutions obtained by CPLEX in small-scale instances.

All the experiments were performed on a computer with 3.60 gigahertz Intel (R) Core (TM) i7-4790 CPU and 3.60 gigabyte RAM. The model and solution methods, including the CW algorithm, RTR travel algorithm, and the PSO algorithm, were implemented using CPLEX 12.6, with C# (VS2015) concert technology.

*5.1. Generation of Test Instance.* In this section, we demonstrate the implementation of numerical experiments to validate the effectiveness of the proposed model and the efficiency of the proposed algorithm. For small-scale instances, the results of the proposed two-stage heuristic algorithm were compared with the optimal solutions obtained by CPLEX and the solutions obtained using the TS algorithm. Based on the same decision obtained in the first stage, the results obtained using the PSO heuristic were compared with those of the TS algorithm in medium- and large-scale instances, given that CPLEX would have been time consuming or unable to solve the problem.

For the numerical experiments on the VRPPMP, instances were randomly generated in the three scales. We simulated the cases of 10 suppliers and two production batches, and 15 suppliers and three production batches for the small-scale; 30 suppliers and six production batches, and 50 suppliers and 10 production batches for the medium scale; 70 suppliers and 14 production batches, 90 suppliers and 18 production batches, 100 suppliers and 18 production batches for the large scale. Information pertaining to the coordinate positions and pickup demands of the 10 suppliers of the first small-scale instance (i.e., 10-2-1) are shown in Table 1. The coordinate position of the manufacturer was (5, 5), and the cost of unit distance traveled by vehicle was \$2. The manufacturer had sufficient pickup vehicles with the capacity of 10 tons to provide a service to suppliers. The volume of raw materials required for each production batch was 10 tons, and the unit inventory cost of unprocessed raw materials in each batch was \$20.

*5.2. Analysis of Calculation Results.* We tested examples at different scales and used CPLEX, the PSO algorithm, and the TS algorithm, respectively, to solve the VRPPMP. Table 2 yields the following conclusions. In the small-scale examples (10-2), the exact solutions (73.45, 173.17, and 184.65) were

TABLE 1: Pickup demands (tons) and coordinate positions (km) of instances 10-2-1.

Supplier	Pickup demands	Coordinate axis X	Coordinate axis Y
1	1.5577	6.2804	8.3439
2	1.4597	5.6732	5.9806
3	2.2747	2.0603	5.5888
4	1.9595	9.0603	4.4218
5	1.5754	9.7755	2.7370
6	2.5402	2.9191	4.6731
7	2.0115	6.3266	4.6951
8	1.5807	9.8215	0.3037
9	2.3763	8.6237	9.9535
10	2.7253	6.7718	3.1459

TABLE 2: Computational results of small-scale instances.

Instance ID	CPLEX		PSO algorithm			TS algorithm		
	Result (\$)	Time (sec)	Result (\$)	Time (sec)	Gap (%)	Result (\$)	Time (sec)	Gap (%)
10-2-1	73.45	5.45	73.45	4.47	0	73.45	7.55	0
10-2-2	173.17	3.66	176.57	3.07	1.96	176.57	7.08	1.96
10-2-3	184.65	5.10	189.19	3.55	2.46	189.19	7.60	2.46
15-3-1	-	>12hours	248.54	14.24	-	248.54	17.30	-
15-3-2	-	>12hours	435.27	13.69	-	435.27	17.47	-
15-3-3	-	>12hours	236.67	13.73	-	236.67	17.39	-

TABLE 3: Computational results of medium-scale instances.

Instance ID	PSO algorithm		TS algorithm		Relative deviation (TS - PSO)/PSO	
	Result (\$)	Time (sec)	Result (\$)	Time (sec)	Result (%)	Time (%)
30-6-1	572.88	83.46	572.88	510.69	0	511.90
30-6-2	567.30	134.09	551.66	519.58	-2.76	287.49
30-6-3	672.89	37.92	672.89	593.52	0	1465.19
50-10-1	1243.89	237.73	1198.75	2246.67	-3.63	845.05
50-10-2	1642.45	611.63	1577.61	2291.87	-3.95	274.72
50-10-3	2501.16	252.34	2420.29	2295.83	-3.23	809.82

obtained by CPLEX, and the accuracy of the model is verified. Compared with the TS algorithm, the PSO algorithm can obtain the approximate solutions (73.45, 176.57, and 189.19) in a shorter time. In the small-scale examples (15-3), CPLEX had difficulty in getting the results, but the PSO algorithm and TS algorithm can still obtain the results in a short time.

However, after the scale of the study increases, as shown in Tables 3 and 4, the PSO algorithm can take less time to get a relatively appropriate solution than TS algorithm. In the medium-scale examples, the relative value deviations between the PSO algorithm and the TS algorithm are within 4%. However, the relative time deviations between the PSO algorithm and the TS algorithm reach 699% on average. In the large-scale examples, although the relative value deviations between the PSO algorithm and the TS algorithm reached

nearly 7%, the PSO calculation results are still acceptable and it holds great advantages in terms of computing time. This proves that the two-stage heuristic algorithm can help us solve such problems well.

### 6. Conclusion

This paper discusses a novel and practical research issue arising in the manufacturing industry; namely, the problem of vehicle routing involving pickup in the context of multibatch production. In order to reduce transportation costs and inventory costs in the production process, a mixed integer programming model was developed. Following this, a two-stage hybrid, heuristic algorithm was put forward to solve this model. In the first stage, the CW algorithm was proposed in order to generate the initial vehicle routes. The

TABLE 4: Computational results of large-scale instances.

Instance ID	PSO algorithm		TS algorithm		Relative deviation (TS - PSO)/PSO	
	Result (\$)	Time (sec)	Result (\$)	Time (sec)	Result (%)	Time (%)
70-14-1	2266.15	1122.67	2109.41	7069.80	-6.92	529.73
70-14-2	3721.10	561.83	3563.46	6017.07	-4.24	970.98
70-14-3	4534.08	949.29	4314.32	6278.70	-4.85	561.41
90-18-1	4221.80	4564.09	4014.15	21226.73	-4.92	365.08
90-18-2	4274.58	1648.82	4016.14	13713.70	-6.05	731.73
90-18-3	7772.56	3691.00	7405.27	19607.38	-4.73	431.22
100-18-1	5164.45	6982.74	4935.79	28613.04	-4.43	309.77
100-18-2	8047.58	7233.14	7762.98	28562.40	-3.54	294.88
100-18-3	11321.39	6922.66	10866.73	28682.46	-4.02	314.33

RTR travel algorithm was then used further to optimize the generated routes. In the second stage of the algorithm, the PSO algorithm was developed so as to allocate the vehicles to each production batch. This process ensured that the manufacturer first assigned each supplier to the appropriate route and then assigned each route to the appropriate batch.

By testing examples at different scales, we were able to validate the effectiveness of the proposed model and the performance efficiency of the two-stage hybrid, heuristic algorithm. First, in the small-scale examples, the exact solution was obtained using CPLEX, and the accuracy of the model was verified. Second, after the scale was expanded, it was found that CPLEX had difficulty in solving this over a short period of time and that the PSO algorithm took less time to obtain a relatively appropriate result than the TS algorithm. This proves that the two-stage heuristic algorithm is well placed to help us solve such problems.

There are certain limitations to the current study. First, we assumed that the amount of the single raw material required for each production batch would be the same; in practice, however, the amount of multiple raw materials varies for each production batch. Second, this paper only considered the combination of the pickup stage and the production stage, without considering product delivery. The joint optimization of multibatch production and the vehicle routing problem with pickup and delivery, considering inventory, will thus be addressed in future studies.

## Data Availability

The data used to support this study have been deposited in the Baidu Netdisk, readers can access the data at: [https://pan.baidu.com/s/1ZSkGF-syiNTk\\_FN9JNShbA](https://pan.baidu.com/s/1ZSkGF-syiNTk_FN9JNShbA).

## Conflicts of Interest

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

## Acknowledgments

This research is supported by the National Natural Science Foundation of China [no. 71771143 and no. 71501125] and Shanghai Young Eastern Scholar Programme [QD2015041].

## References

- [1] G. B. Dantzig and J. H. Ramser, "The truck dispatching problem," *Management Science*, vol. 6, no. 1, pp. 80–91, 1959.
- [2] G. Clarke and J. W. Wright, "Scheduling of vehicles from a central depot to a number of delivery points," *Operations Research*, vol. 12, no. 4, pp. 568–581, 1964.
- [3] R. Fukasawa, H. Longo, J. Lygaard et al., "Robust branch-and-cut-and-price for the capacitated vehicle routing problem," *Mathematical Programming*, vol. 106, no. 3, pp. 491–511, 2006.
- [4] M. L. Gambardella, D. E. Taillard, and G. Agazzi, "A multiple ant colony system for vehicle routing problems with time windows," 1999.
- [5] P. Sombuntham and V. Kachitvichyanukul, "Multi-depot vehicle routing problem with pickup and delivery requests," in *Proceedings of the International MultiConference of Engineers and Computer Scientists, IMECS 2010*, pp. 71–85, China, March 2010.
- [6] C. K. Y. Lin, "A vehicle routing problem with pickup and delivery time windows, and coordination of transportable resources," *Computers & Operations Research*, vol. 38, no. 11, pp. 1596–1609, 2011.
- [7] M. W. Savelsbergh and M. Sol, "The general pickup and delivery problem," *Transportation Science*, vol. 29, no. 1, pp. 17–29, 1995.
- [8] H. Xu, Z.-L. Chen, S. Rajagopal, and S. Arunapuram, "Solving a practical pickup and delivery problem," *Transportation Science*, vol. 37, no. 3, pp. 347–364, 2003.
- [9] G. Nagy and S. Salhi, "Heuristic algorithms for single and multiple depot vehicle routing problems with pickups and deliveries," *European Journal of Operational Research*, vol. 162, no. 1, pp. 126–141, 2005.
- [10] P. Kalina and J. Vokřínek, "Parallel solver for vehicle routing and pickup and delivery problems with time windows based on agent negotiation," in *Proceedings of the 2012 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2012*, pp. 1558–1563, Republic of Korea, October 2012.

- [11] H.-Z. Zheng, H.-Y. Guo, and X.-D. Zhang, "Modeling and approach for vmi cyclic inventory routing problem," in *Proceedings of the 2009 International Conference on Machine Learning and Cybernetics*, pp. 1393–1398, China, July 2009.
- [12] A. L. Shiguemoto and V. A. Armentano, "A tabu search procedure for coordinating production, inventory and distribution routing problems," *International Transactions in Operational Research*, vol. 17, no. 2, pp. 179–195, 2010.
- [13] Y. Adulyasak, J.-F. Cordeau, and R. Jans, "Formulations and branch-and-cut algorithms for multivehicle production and inventory routing problems," *INFORMS Journal on Computing*, vol. 26, no. 1, pp. 103–120, 2014.
- [14] B. Sainathuni, P. J. Parikh, X. Zhang, and N. Kong, "The warehouse-inventory-transportation problem for supply chains," *European Journal of Operational Research*, vol. 237, no. 2, pp. 690–700, 2014.
- [15] P. C. Yellow, "A Computational Modification to the Savings Method of Vehicle Scheduling," *Journal of the Operational Research Society*, vol. 21, no. 2, pp. 281–283, 2017.
- [16] G. Dueck, "New optimization heuristics: the great deluge algorithm and the record-to-record travel," *Journal of Computational Physics*, vol. 104, no. 1, pp. 86–92, 1993.
- [17] F. Li, B. Golden, and E. Wasil, "Very large-scale vehicle routing: new test problems, algorithms, and results," *Computers & Operations Research*, vol. 32, no. 5, pp. 1165–1179, 2005.
- [18] R. C. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proceedings of the 6th International Symposium on Micromachine and Human Science*, pp. 39–43, Nagoya, Japan, October 1995.

## Research Article

# Integrated Optimization on Assortment Packing and Collaborative Shipping for Fashion Clothing

Qianying Wang, Yiping Jiang , and Yang Liu

College of Engineering, Nanjing Agricultural University, Nanjing 210031, China

Correspondence should be addressed to Yiping Jiang; ypjiang@njau.edu.cn

Received 6 August 2018; Accepted 6 September 2018; Published 21 October 2018

Academic Editor: Lu Zhen

Copyright © 2018 Qianying 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.

With the diversification of customer's demand and the shortage of social resources, meeting diverse requirements of customers and reducing logistics costs have attracted great attention in logistics area. In this paper, we address an integrated optimization problem that combines fashion clothing assortment packing with collaborative shipping simultaneously. We formulate this problem as a mixed integer nonlinear programming model (MINLP) and then convert the proposed model into a simplified model. We use LINGO 11.0 to solve the transformed model. Numerical experiments have been conducted to verify the effectiveness and efficiency of the proposed model, and the numerical results show that the proposed model is beneficial to the fashion clothing assortment packing and collaborative shipping planning.

## 1. Introduction

Within a supply chain for fashion clothing, inefficient operation and irrational resource utilization are key considerations as they lead to high costs and low profits. The inefficient operation is mainly due to the redundant process of sorting and packing items in the supply chain, which is further coursed by the variety of both items and demand. The irrational resource utilization is mainly determined by the process of item distribution in the supply chain, which further depends on the plenty of stores located in different locations. As one of the largest fashion clothing industrial producing countries in the world, Chinese fashion clothing industry support ¥2,000 billion of domestic fashion consumer market [1], and the retail sales of the fashion clothing were at ¥1,035.64 billion in 2017; it is predicted that the retail sales will reach ¥1,095.8 billion in 2018 [2]. Though the fashion clothing industry has achieved certain results, the above two problems are still common in the fashion clothing supply chain. It is reported that most truck loads in the distribution process are usually only 57% of their maximum gross weight [3]. In addition, the proportion of social logistics costs to GDP is almost over two times that of developed countries [4]. Therefore, it is of great significance to improve the operational

efficiency by assorting and packing items in the origin of the supply chain and improve the resource utilization through carrying out collaborative shipping in the end of the supply chain at the same time.

The problem of integration optimization on assortment packing and collaborative shipping has the following characteristics: (1) multiple types of demand but small single item demand and (2) large number of stores but geographically dispersed. For different characteristics, we propose the following countermeasures and possible challenges: for the former, we can introduce the box configurations to meet the diversified demand. However, if there are large number of box configurations, they will be able to meet the demand of each store in an exact manner, but with increased operating costs. If the number of box configurations is limited, the operating costs will be low, but some items will be overloaded or underloaded at the stores. For the latter, we can connect the large number of stores located in different locations through collaborative shipping under the constraints of truck capacity and geographical location.

The integration of assortment packing and collaborative shipping provides an effective way to improve the operational efficiency and reduce the waste of resources in the fashion clothing supply chain. The decision-making problem reveals

the assortment packing ways and the distribution routes according to the demand of stores, the types of box configurations, and shipping cost between different nodes and truck cost. In this paper, we aim to solve two questions: (1) how to determine the type and quantity of each box configuration allocated to each store and (2) how to realize collaborative shipping through vehicle scheduling.

The contributions in this study contain three aspects: (1) addressing an integrated optimization problem that combines fashion clothing assortment packing with collaborative shipping; (2) proposing a mixed integer nonlinear programming model to formulate the fashion clothing assortment packing and collaborative shipping problem; (3) converting the proposed model into a simplified model, applying LINGO 11.0 to solve the converted model, and carrying out numerical experiments and sensitive analysis to verify the efficiency and validity of the converted model.

The remainder of the paper is organized as follows. Section 2 presents the relevant literature; Section 3 formalizes the problem using mathematical models; Section 4 uses a small instance to test the effectiveness of the numerical experiment; Section 5 draws some conclusions and outlines future research directions.

## 2. Literature Review

Our study involves two major streams of research literature, that is, the assortment packing decision and the collaborative shipping optimization.

In the research stream of assortment packing decision, relevant literature basically combines optimization model with heuristic algorithm to study. Depending on research objectives, the existing literature can be divided into two categories: the minimization of entire supply chain costs and the minimization of mismatches between the supply and the demand.

For the minimization of entire supply chain costs, Chettri and Sharma [5] presented the entire decision-making process for the assortment packing problem and they proposed a method to balance the trade-offs between different points in supply chain based on data intensive models. For the minimization of mismatches between the supply and the demand, Sung et al. [6] considered a nonlinear programming model and applied it to Kolon Sport Company by converting the nonlinear programming model into a linear programming model and then they solved the problem with a heuristic algorithm. Based on the data set, the sales increased by approximately 8% in 2015. Other researchers in this field have proposed numbers of similar heuristic algorithms and programming models [7–10], and the optimal assortment packing decision can be found through iteratively solving the problem. Besides, two other papers that combine the above two objectives are Wang [11] and Pratti [12]. Both formulated the assortment packing decision with inner packs and outer packs, and they proposed various heuristic approaches that assume that the available pack configurations are given, which is more comprehensive and practical to solve the problem.

In the research stream on collaborative shipping optimization, many studies focus on the vehicle routing problems with truck capacity-constrained (CVRP). To our knowledge, Clarke and Wright [13] were the first to apply the liner optimization model to the problem, which is ordinarily encountered in the logistics and transportation field under the name of CVRP. They have numerically demonstrated that the utilization of the truck capacity improved by 17%. In addition, researchers [14–18] proposed various approaches to solve CVRP, such as the enhanced version of the artificial bee colony heuristic based on the bee swarming behavior model. On the basis of CVRP, some further studies have been conducted to effectively solve practical problem. Lysgaard [19] formulated an exact branch-and-cut-and-price algorithm for a restricted version of the capacitated vehicle routing problem named pyramidal capacitated vehicle routing problem (PCVRP). Liu et al. [20] established a mix integer programming model and an efficient memetic algorithm to the Close-Open Mixed Vehicle Routing Problem (COMVRP), with the objective of minimizing the variable and fixed costs for operating the private vehicles and the rental vehicles.

This paper was inspired by Sung et al. [6, 10] and Creemers et al. [16]. Sung et al. [6, 10] provided the idea of assortment packing decision with the objective of minimizing the mismatches between the supply and the demand. Creemers et al. [16] introduced the effective algorithms to realize collaborative shipping. They both investigated some challenges faced by supply chain operations management. However, from the perspective of improving the operational efficiency of entire fashion clothing supply chain, their research objects only contain the origins (manufacturers) and the destinations (stores), while some important nodes such as distribution centers in the supply chain are not included. Therefore, this paper incorporates distribution center and combines assortment packing decision with collaborative shipping optimization to optimize the entire supply chain from a systematic perspective.

In summary, assortment packing and collaborative shipping have attracted more and more attention in area of fashion clothing supply chain. The existing researches have studied these two problems for many years but seldom combine the two for research. With the increasing diversity of demand and the continuous reduction of resources in fashion clothing supply chain, the integration of assortment packing and collaborative shipping will be a major trend in the fashion clothing supply chain. Therefore, this study aims to endeavor to bridge a gap in the literature by integrating assortment packing decision and collaborative shipping optimization in the fashion clothing supply chain.

## 3. Model Formulation

We consider a problem of integrated optimization on assortment packing and collaborative shipping for fashion clothing and model this problem as a MINLP. The objective is to minimize the total costs, including the value of overload and underload items, shipping cost, and truck cost.

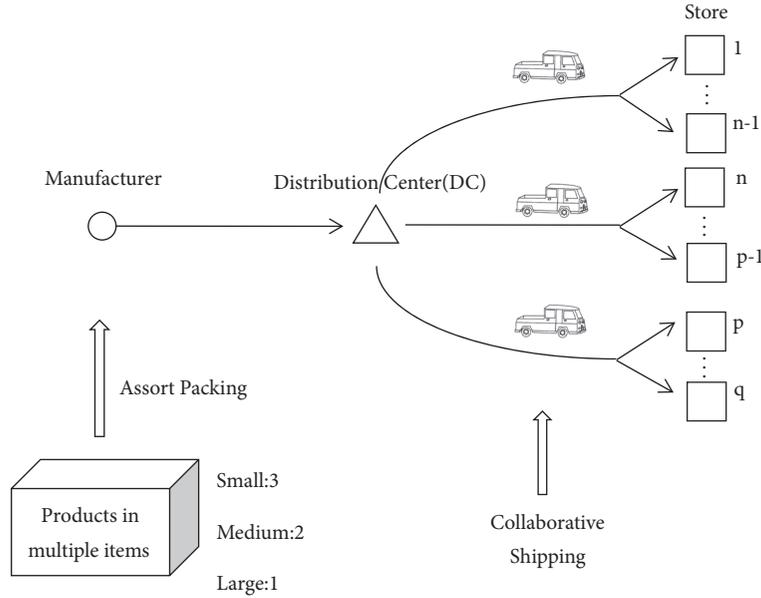


FIGURE 1: Fashion clothing supply chain process.

**3.1. Problem Description.** The integrated optimization problem on assortment packing and collaborative shipping for fashion clothing in this paper is depicted in Figure 1. It contains one manufacturer, one distribution center, various stores located in different locations, and multiple homogeneous trucks. Manufacturer sorts and packs items with multiple colors and sizes into box configurations according to the requirement of each store, and then the box configurations will be delivered to the distribution center. At the distribution center, instead of simply sending a fixed car to deliver fixed box configurations to a fixed store, the stuff will identify the collaborative shipping opportunities before delivering them and the box configurations will be delivered to the corresponding stores through collaborative shipping.

The decisions to be taken in this model are as follows: (1) how many of each item each box configuration should include, (2) how many of each box configuration should be delivered to each store, and (3) which box configuration is delivered by which truck to which store and the optimal served sequence, with the objective of minimizing the total costs including the value of overload and underload items, shipping cost, and truck cost.

Some assumptions are proposed to the integrated assortment packing and collaborative shipping problem. For assortment packing (1) type and quantity of items contained in a box configuration are fixed; (2) every store should receive at least one box and then the total number of boxes should be greater than or equal to the total number of stores; (3) configurations of box is limited; (4) number of each box configuration that can be used is limited; (5) box configurations must be full of items before delivery. Due to the assumptions of (1), (3), and (5), the number of items allocated to each store may not be exactly equal to the store's demand. If the number of items delivered to a store is more (less) than the

store's demand, overload (underload) will occur. Therefore, we should minimize overload and underload simultaneously.

For collaborative shipping (1) we only consider the delivery of boxes from a single distribution center to multiple stores in this paper; (2) collaborative shipping can only be carried out when the geographical location of the two places is similar and the capacity of the truck is sufficient enough; (3) the cost and the load of all trucks are the same. Hence, the decision is to minimize shipping cost and truck cost.

### 3.2. Symbol Description

#### 3.2.1. Index Sets

$I$  = Items indexed with  $i$ .

$S$  = Stores indexed with  $s$ .

$DCS$  = Distribution center and stores.  $DCS = S \cup \{0\}$ , 0 indicates the distribution center.

$K$  = Trucks indexed with  $k$ .

$B$  = Box configurations indexed with  $b$ .

$N$  = Maximum number of different box configurations that can be used for assortment packing.

#### 3.2.2. Parameters

$d_{is}$  = Demand quantity for item  $i$  in store  $s$ .

$c_{mn}$  = Delivery cost from node  $m$  to node  $n$ .

$Q$  = Set of possible box capacities. A capacity value in this set is expressed as  $q \in Q$ .

$M_b$  = Maximum number of box  $b$ .

$p_i$  = Production number of item  $i$ .

$w$  = Capacity of truck.

$g_i$  = Weight of item  $i$ .

$c_i$  = Production cost of item  $i$ .

$F$  = The cost of using a truck.

### 3.2.3. Decision Variables

$t_{bs}$  = Number of box configuration  $b$  distributed to store  $s$ .

$r_{ib}$  = Number of item  $i$  in box configuration  $b$ .

$o_{is}$  = Overload of item  $i$  to store  $s$ .

$u_{is}$  = Underload of item  $i$  to store  $s$ .

$n_b$  = Number of box configuration  $b$  distributed to distribution center.

$x_{mnk} = \{1, \text{Truck } k \text{ arrives at the store } n \text{ from node } m (m \in \text{DCS}, n \in S); 0, \text{If not}\}.$

$y_{ks} = \{1, \text{Truck } k \text{ delivers to store } s; 0, \text{If not}\}.$

$f_{bq} = \{1, \text{Capacity of box configuration } b \text{ is } q; 0, \text{If not}\}.$

$z_k = \{1, \text{Truck } k \text{ is used}; 0, \text{If not}\}.$

**3.3. Objective Function.** We consider three different objectives in this study, i.e., the value of overload and underload items, shipping cost, and truck cost.

**3.3.1. The Value of Overload and Underload Items.** In assortment packing process, please note that we have assumed that each box configuration should be full of items before delivery; hence the number of items allocated to each store may not be exactly equal to store's demand, which will bring loss to stores. Therefore, our first objective is to minimize the value of overload and underload items. The calculation is given in

$$\text{obj1} = \sum_{i \in I} \sum_{s \in S} c_i (u_{is} + o_{is}) \quad (1)$$

**3.3.2. Shipping Cost.** Let  $c_{mn}$  denote the shipping cost from node (distribution center or stores)  $m$  to node (stores)  $n$ .  $x_{mnk}$  is a binary decision variable, taking a value 1 if the truck  $k$  arrives at store  $n$  from node  $m$  and 0 otherwise. Shipping cost can be calculated in

$$\text{obj2} = \sum_{m \in \text{DCS}} \sum_{n \in S} \sum_{k \in K} (c_{mn} x_{mnk}) \quad (2)$$

**3.3.3. Truck Cost.** Let  $F$  denote the usage cost of a single truck, and  $z_k$  is a binary decision variable, taking a value 1 if the truck  $k$  is used and 0 otherwise. Truck cost can be calculated in

$$\text{obj3} = \sum_{k \in K} F z_k \quad (3)$$

**3.4. Constraints.** We formulate the model as a MINLP problem. The constraints are given as follows (note that some constraints are modified according to Sung et al. [6]).

$$\sum_{b \in B} r_{ib} t_{bs} - o_{is} + u_{is} = d_{is}, \quad i \in I, s \in S \quad (4)$$

$$o_{is} u_{is} = 0, \quad i \in I, s \in S \quad (5)$$

$$\sum_{i \in I} r_{ib} = \sum_{q \in Q} q f_{bq}, \quad b \in B \quad (6)$$

$$\sum_{q \in Q} f_{bq} = 1, \quad b \in B \quad (7)$$

$$\sum_{b \in B} \sum_{s \in S} t_{bs} \leq \sum_{b \in B} M_b \quad (8)$$

$$\sum_{b \in B} t_{bs} \geq 1, \quad s \in S \quad (9)$$

$$\sum_{s \in S} \sum_{b \in B} r_{ib} t_{bs} \leq p_i, \quad i \in I \quad (10)$$

$$\sum_{s \in S} \left[ \sum_{i \in I} \left( \sum_{b \in B} r_{ib} t_{bs} \right) g_i \right] y_{ks} \leq w z_k, \quad k \in K \quad (11)$$

$$\sum_{k \in K} y_{ks} = 1, \quad s \in S \quad (12)$$

$$\sum_{m \in \text{DCS}} x_{mnk} = y_{kn}, \quad n \in S, k \in K \quad (13)$$

$$\sum_{n \in \text{DCS}} x_{mnk} = y_{km}, \quad m \in S, k \in K \quad (14)$$

$$\sum_{b \in B} r_{ib} n_b \geq \sum_{j \in I} d_{ij}, \quad i \in I \quad (15)$$

$$n_b = \sum_{s \in S} t_{bs}, \quad b \in B \quad (16)$$

$$\sum_{n \in S} x_{onk} = z_k, \quad k \in K \quad (17)$$

$$t_{bs} \geq 0,$$

$$r_{ib} \geq 0,$$

$$o_{is} \geq 0, \quad (18)$$

$$u_{is} \geq 0,$$

$$n_b \geq 0$$

$$f_{bq} \in \{0, 1\},$$

$$x_{mnk} \in \{0, 1\}, \quad (19)$$

$$y_{ks} \in \{0, 1\},$$

$$z_k \in \{0, 1\}$$

Equation (4) calculates the amount of overload and underload based on the distribution quantity and stores' demand. If the number of the items distributed to a store is smaller than the store's demand, we calculate the underload as the demand minus the distribution quantity. If the number of the items distributed to a store is greater than the

store's demand, we calculate the overload as the distribution quantity minus the demand. Equation (5) limits at least one of the overloads or underloads of allocating item  $i$  to store  $s$  to be 0. That is, if overload occurs, underload must equal to 0. Conversely, if underload occurs, overload must equal to 0. If the number of item  $i$  allocated to store  $s$  is exactly equal to the store's demand, the overload and underload are both 0. Equation (6) indicates that every box configuration has a certain capacity. Equation (7) denotes that there is only one possible capacity for each box configuration. Equation (8) limits the total number of box configurations. Equation (9) limits each store to receive at least one box. Equation (10) limits the total number of each item that can be used for distribution so that it cannot exceed its production quantity. Equation (11) points out that the load of each truck cannot exceed its rated capacity. Equation (12) limits each store can be served by one truck. Equation (13) limits the following: if store  $n$  is delivered by truck  $k$ , then the truck  $k$  must arrive at the store  $n$  from a node  $m$ . Equation (14) limits that if store  $m$  is delivered by truck  $k$ , then the truck  $k$  will reach another node  $n$  after the delivery. Equation (15) limits that the number of item  $i$  received by the distribution center must be greater than or equal to all stores' demand for item  $i$ . Equation (16) connects the manufacturer, the distribution center, and the stores; that is, the number of received box configurations by the distribution center must be equal to the number of delivered box configurations, which indicates that the box must pass through the distribution center before reaching stores. Equation (17) denotes that if the truck is used, the starting point must be the distribution center. Equations (18)–(19) define the conditions on the decision variables.

As a result of the product of two decision variables in (4), (5), (10), (11), and (15), the proposed model becomes nonlinear and is hard to solve. One possible method to linearize the model is to decompose one of the decision variables into its binary expansion [8]. However, with this approach, the model will introduce more variables and constraints, which will make the model too complex to solve [10]. Taking the actual production of boxes into account, if the production of box configurations is based on the real demand of stores for items, the production cost of raw materials will be greatly increased. To control the raw material costs within a certain level, the type of box configurations produced by the box manufacturers should be limited and fixed. Therefore, we convert the proposed model into a simplified model, which can obtain a feasible and timely solution for the industry-size problem.

**3.5. Model Conversion.** According to the practical operations, we use a set of feasible box configurations to simplify the proposed model (Sung et al. [6]). Therefore, the variable  $r_{ib}$  in the original model becomes a parameter; we only determine the type and quantity of each box configuration allocated to each store, without deciding the type and quantity of items to be included in each box configuration.

Let  $e_{ib}$  denote the number of item  $i$  that can be accommodated in box configuration  $b$ , which is given in advance.  $h_b$  is a binary decision variable, taking a value 1 if the box configuration  $b$  is used and 0 otherwise. The converted model is given as follows:

$$\text{Minimize } obj1 + obj2 + obj3 \quad (20)$$

$$\sum_{b \in B} e_{ib} t_{bs} - o_{is} + u_{is} = d_{is}, \quad i \in I, s \in S \quad (21)$$

$$o_{is} u_{is} = 0, \quad i \in I, s \in S \quad (22)$$

$$\sum_{b \in B} h_b \leq N \quad (23)$$

$$\sum_{b \in B} \sum_{s \in S} t_{bs} \leq \sum_{b \in B} M_b \quad (24)$$

$$\sum_{s \in S} t_{bs} \geq h_b, \quad b \in B \quad (25)$$

$$\sum_{b \in B} t_{bs} \geq 1, \quad s \in S \quad (26)$$

$$\sum_{s \in S} \sum_{b \in B} e_{ib} t_{bs} \leq p_i, \quad i \in I \quad (27)$$

$$\sum_{s \in S} \sum_{i \in I} \sum_{b \in B} e_{ib} t_{bs} g_i \gamma_{ks} \leq w z_k, \quad k \in K \quad (28)$$

$$\sum_{k \in K} \gamma_{ks} = 1, \quad s \in S \quad (29)$$

$$\sum_{m \in DCS} x_{mkn} = \gamma_{kn}, \quad n \in S, k \in K \quad (30)$$

$$\sum_{n \in DCS} x_{imnk} = \gamma_{km}, \quad m \in S, k \in K \quad (31)$$

$$\sum_{b \in B} e_{ib} n_b \geq \sum_{j \in J} d_{ij}, \quad i \in I \quad (32)$$

$$n_b = \sum_{s \in S} t_{bs}, \quad b \in B \quad (33)$$

$$\sum_{n \in S} x_{0nk} = z_k, \quad k \in K \quad (34)$$

$$\begin{aligned} t_{bs} &\geq 0, \\ o_{is} &\geq 0, \\ u_{is} &\geq 0, \end{aligned} \quad (35)$$

$$n_b \geq 0$$

$$x_{mkn} \in \{0, 1\},$$

$$\gamma_{ks} \in \{0, 1\},$$

$$z_k \in \{0, 1\},$$

$$h_b \in \{0, 1\} \quad (36)$$

## 4. Numerical Experiments

The numerical experiments are implemented on a DELL laptop with 2.30 GHz Inter(R) Core(TM) and 4GB RAM. Please note that the converted model given in Section 3.5 is a nonlinear problem and extremely difficult to obtain the

TABLE 1: Numbering the downjacket based on different size and color.

Size	Color		
	Black	White	Blue
Small	1	2	3
Medium	4	5	6
Large	7	8	9

Notes. Downjackets with different sizes and colors have different numbers. For example, the number 5 shows that the downjacket is white and of medium size.

TABLE 2: Production, weight, and unit price of each downjackets.

Item	Production (unit)	Weight (kg)	Unit Price (CNY)
1	188	10	32
2	163	6	32
3	175	9	30
4	185	9	26
5	195	7	27
6	198	8	34
7	177	5	27
8	157	5	33
9	157	8	30

Note. Item 1 shows that this kind of downjacket has a total production of 188 pieces, each weighing 10 kg, and the unit price is 32 CNY.

TABLE 3: Demand on different items for different stores.

Item	Store							
	1	2	3	4	5	6	7	8
1	8	4	11	10	4	12	4	6
2	5	5	6	8	9	9	7	5
3	7	6	9	5	3	10	8	7
4	9	9	5	5	4	7	7	6
5	9	4	3	11	8	7	9	12
6	7	9	10	3	4	10	9	11
7	6	4	8	7	10	4	6	3
8	12	9	7	5	10	4	7	10
9	3	7	11	12	10	5	3	5

Note. The demand for nine items in store 1 is 8, 5, 7, 9, 9, 7, 6, 12, and 3, respectively.

global optimal solution. In this section, we adopt LINGO 11.0 as the MINLP solver to solve the converted model and to approximate the optimal solution.

**4.1. Example Setting.** In this example, the fashion clothing supply chain consists of one distribution center, 8 stores, and 9 different items: downjackets available in 3 different sizes (small, medium, and large) and 3 different colors (black, white, and blue). The distribution center has 5 available box configurations (the available number of each box configuration is 5, 7, 6, 9, and 7, respectively) and 5 available trucks (the capacity of truck is 5 tons and the usage cost for each truck is 50 CNY). Table 1 gives the serial number of downjackets according to different sizes and colors. Table 2 gives the production, weight, and unit price of each downjacket. Table 3 gives demand on different items for different stores. Table 4 gives the shipping cost between each node located in different geographical locations. Table 5 presents the number of each

downjacket that can be included in each box configuration and the available number of each box configuration.

**4.2. Numerical Results.** The numerical results are reported in Table 6 and Figure 2. It can be seen that the number of each box configuration delivered to the distribution center is 5, 0, 1, 7, and 7, respectively. The optimum route of truck  $k_1$  starts from distribution center and serves the stores with the sequence  $s_7 \rightarrow s_5$ . Similarly, the optimal route of truck  $k_2$  is  $s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_6 \rightarrow s_8 \rightarrow s_4$ .

Table 7 presents the most detailed results. It is obvious that 8 stores are served by 2 trucks. The driving route for truck  $k_1$  is as follows. First, truck  $k_1$  collects the shipments of store  $s_7$  and  $s_5$  at distribution center and then departs for store  $s_7$ ; upon arrival at store  $s_7$ ,  $k_1$  drops off the shipments for store  $s_7$  and then continues to store  $s_5$ . Finally, truck  $k_1$  backs to distribution center. Similarly, for truck  $k_2$ , it collects the shipments of stores  $s_1, s_2, s_3, s_6, s_8$ , and  $s_4$ , then delivers

TABLE 4: Shipping costs for the node pairs (CNY).

	Store 1	Store 2	Store 3	Store 4	Store 5	Store 6	Store 7	Store 8
D/C	35	42	44	41	41	44	38	34
Store 1	0	30	43	37	39	31	48	44
Store 2	50	0	31	32	48	35	30	39
Store 3	45	36	0	42	40	34	39	33
Store 4	37	33	36	0	49	43	38	37
Store 5	42	34	41	38	0	47	39	42
Store 6	32	38	43	32	49	0	45	34
Store 7	48	32	38	35	35	46	0	45
Store 8	48	42	46	33	44	44	46	0

Note. If the shipping is reversed, the shipping cost between two same nodes may vary.

TABLE 5: Number of each downjacket that can be included in each box configuration and the available number of each box configuration.

Box configuration	Item									No. of corresponding boxes
	1	2	3	4	5	6	7	8	9	
1	3	5	4	5	5	3	4	4	4	5
2	2	6	5	4	5	4	5	4	5	7
3	2	2	2	4	4	3	5	5	5	6
4	3	5	2	4	4	5	4	4	5	9
5	3	4	3	2	5	3	4	5	3	7

Note. Box configuration 1 containing the number of nine items is 3, 5, 4, 5, 5, 3, 4, 4, and 4, respectively. For this configuration, 5 boxes are available.

TABLE 6: Distribution routings of each trucks.

Truck	Distribution Routing
$k_1$	$DC \rightarrow s_7 \rightarrow s_5 \rightarrow DC$
$k_2$	$DC \rightarrow s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_6 \rightarrow s_8 \rightarrow s_4 \rightarrow DC$

the shipments to those stores in sequence. After delivering all stores, it heads back to the distribution center finally.

4.3. Sensitivity Analysis. We further conduct sensitivity analysis in terms of different capacity of truck and whether to carry out collaborative shipping. The input data is generated randomly. To make the approximate local optimal result more accurate, we implement 10 experiments and take their average value as the approximate optimal result for each numerical case in this subsection. The numerical results show that the error of optimal objective for each case is within [-7%, +7%].

4.3.1. Sensitivity Analysis on Capacity of Truck. We vary the capacity of truck from 3.5 to 7.0. The numerical results are reported in Table 8, and Figure 3 depicts the relationship between the capacity of truck and the optimal objective value. It can be seen that the optimal objective value is decreasing with the growing of the capacity of truck. In other words, if the truck has a low capacity, the distribution center will adopt more trucks to distribute items to stores, so as to meet the diversified demand of customers and improve the service satisfaction.

4.3.2. Sensitivity Analysis on Whether to Carry Out Collaborative Shipping. Whether to carry out collaborative shipping

will bring different costs to the entire supply chain. If collaborative shipping is implemented, the total costs of the supply chain are as described above, i.e., the value of overload and underload items, shipping cost, and truck cost. If collaborative shipping is not implemented, then the total costs will be different. First, for the value of overload and underload, since the existence of box configurations is to assort and pack different items for a single store rather than for multiple stores, therefore, whether or not to implement collaborative shipping will not influence the value of overload and underload and the cost of *obj1* is unchanged. Second, for shipping cost, since every store needs one truck to serve, the shipping cost converts to the round-trip cost between distribution center and stores, which will increase the cost of *obj2*. Third, for truck cost, since every store needs one truck, the number of trucks must be greater than or equal to the number of stores, which will definitely increase the cost of *obj3*. Obviously the total costs of noncollaborative shipping are higher than that of collaborative shipping.

We change the number of stores from 3 to 8 and calculate the optimal objective value under two cases, denoted as collaborative shipping and noncollaborative shipping. Figure 4 shows that, with the increasing number of stores, the objective value is increased, and the cost of combined assortment packing and collaborative shipping is always lower than the cost of noncombined one. In other words, the collaborative

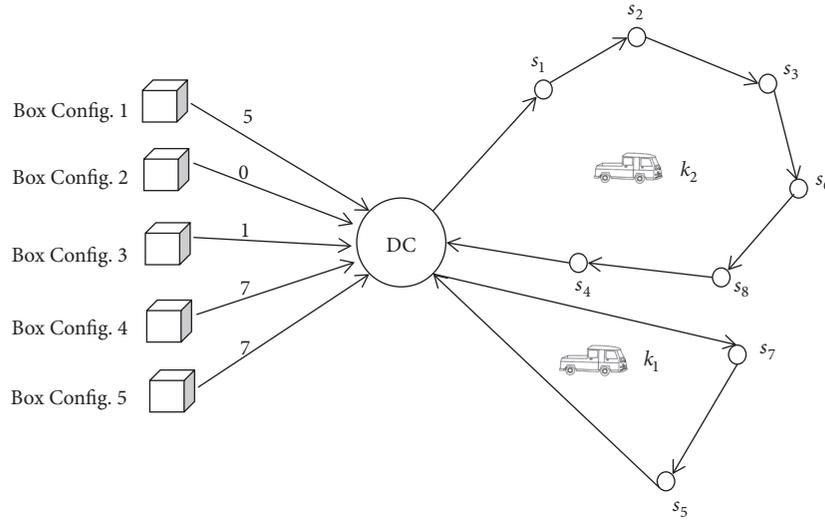


FIGURE 2: Delivery number for different box configurations and optimal route of two trucks.

TABLE 7: Routing of collaborative shipping and changes of different box configurations in each truck.

Truck	Departure location	Box configurations in truck	Arrival location	Box configurations dropped off upon arrival	Box configurations in truck after arrival
$k_1$	$DC$	Box config.1: 1 Box config.4: 3	$s_7$	Box config.1: 1 Box config.4: 1	Box config.4: 2
	$s_7$	Box config.4: 2	$s_5$	Box config.4: 2	/
	$s_5$	/	$DC$	/	/
$k_2$	$DC$	Box config.1: 4 Box config.3: 1 Box config.4: 4 Box config.5: 7	$s_1$	Box config.1: 1 Box config.5: 2	Box config.1: 3 Box config.3: 1 Box config.4: 4 Box config.5: 5
	$s_1$	Box config.1: 3 Box config.3: 1 Box config.4: 4 Box config.5: 5	$s_2$	Box config.1: 1 Box config.4: 1	Box config.1: 2 Box config.3: 1 Box config.4: 3 Box config.5: 5
	$s_2$	Box config.1: 2 Box config.3: 1 Box config.4: 3 Box config.5: 5	$s_3$	Box config.4: 1 Box config.5: 2	Box config.1: 2 Box config.3: 1 Box config.4: 2 Box config.5: 3
	$s_3$	Box config.1: 2 Box config.3: 1 Box config.4: 2 Box config.5: 3	$s_6$	Box config.1: 1 Box config.4: 1 Box config.5: 2	Box config.1: 1 Box config.3: 1 Box config.4: 1 Box config.5: 1
	$s_6$	Box config.1: 1 Box config.3: 1 Box config.4: 1 Box config.5: 1	$s_8$	Box config.3: 1 Box config.5: 1	Box config.1: 1 Box config.4: 1
	$s_8$	Box config.1: 1 Box config.4: 1	$s_4$	Box config.1: 1 Box config.4: 1	/
	$s_4$	/	$DC$	/	/

TABLE 8: Optimal objective value under different truck capacity.

Truck capacity (t)	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
Optimal objective value (CNY)	7243.8	7227.7	7231.1	7229.2	7194.1	7193.2	7193.9	7187.5

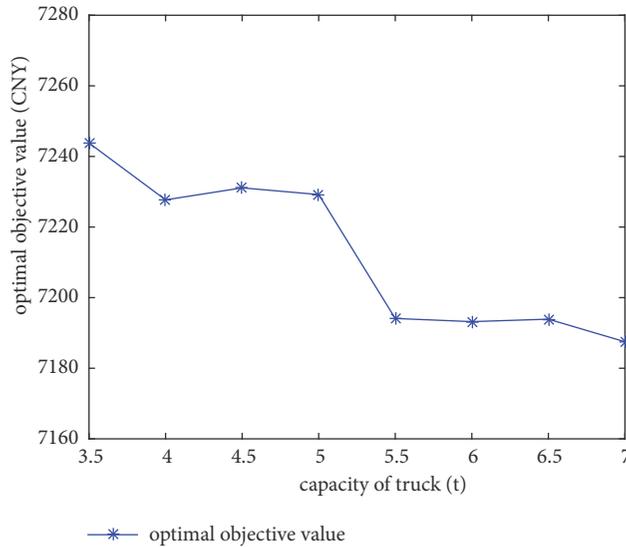


FIGURE 3: Relationship between the capacity of truck and the optimal objective value.

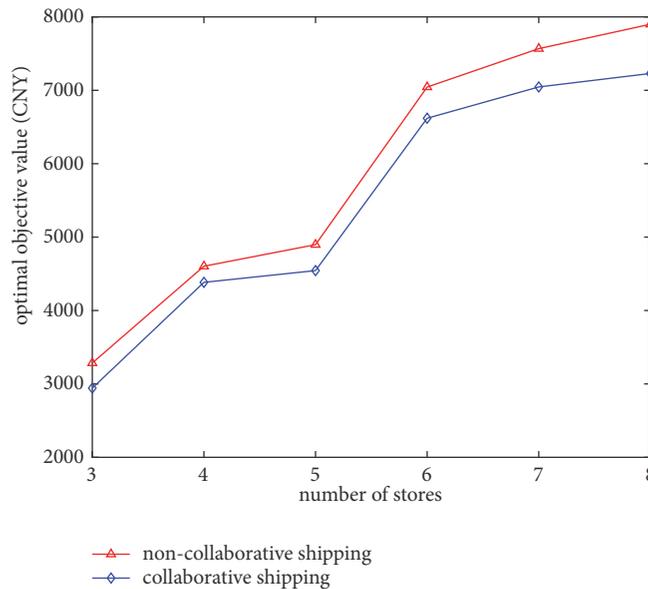


FIGURE 4: Optimal objective value under different number of stores and whether to carry out collaborative shipping.

shipping is always effective throughout the operation of the supply chain, since it can not only reduce the total cost, but also can improve the operational efficiency.

### 5. Conclusions

In this study, we have introduced the integrated optimization problem on assortment packing and collaborative shipping. A MINLP model is presented for the problem, with the

objective of minimizing the value of overload and underload items, shipping cost, and truck cost. The main decisions are as follows: how many of each box configuration should be delivered to each store, and which box is delivered by which truck to which store as well as the optimal served sequence. In the proposed model, taking the actual production of boxes into account, we transform one of the decision variables into parameter and convert the initial model into a simplified model. In addition, numerical experiments have verified the effectiveness and efficiency of the converted model, and the

numerical results and sensitivity analysis have also been reported. Our study contributes to endeavoring to bridge a gap in the integration of assortment packing decision and collaborative shipping optimization of the fashion clothing supply chain.

There are specific areas in this study that can be considered as useful extensions. First of all, this research only considered a single distribution center served multiple stores; hence multiple distribution centers serving multiple stores could be considered as a prospect. Next, collaborative shipping can be extended between distribution centers, between manufacturers, between multiple distribution centers and stores, etc. Finally, items can fully meet the customer's demand to improve the customer satisfaction, which will result in the underload of 0 and we only need to consider the overload. We leave these directions for further study.

### Data Availability

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

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This research was supported by the Natural Science Foundation of Jiangsu Province, China (BK20160742), Students Research Training Program of Nanjing Agricultural University (1730C390), Fundamental Research Funds for the Central Universities (NAU: KYZ201663; NAU: SKYC2017007; NAU: SKTS2016038; NAU: SKYZ2017025), Project of Philosophy and Social Science Research in Colleges and Universities in Jiangsu (2017SJB0030), and Talent Startup Fund of College of Engineering in Nanjing Agricultural University (RCQD16-01).

### References

- [1] H. Gao, L. Liang, R. Liang, and F. Niu, "Chinese fashion: set sail again in the new era," *China Textile*, vol. no. 12, pp. 40–41, 2017.
- [2] "Forecast of Market Size of China's Garment Industry in 2018: Retail Value Nearly RMB 1100 Billion," 2018, [http://www.askci.com/news/chanye/20180424/163605122110\\_2.shtml](http://www.askci.com/news/chanye/20180424/163605122110_2.shtml).
- [3] S. Doherty and S. Hoyle, "Supply chain decarbonization: The role of logistics and transport in reducing supply chain carbon emissions," Report, World Economic Forum, Geneva, Switzerland, 2009.
- [4] W. Jigang, "Why Logistics Costs are High in China?" 2018, [http://www.sz.gov.cn/szsfzyjzx/jczx/zjgd\\_1/201707/t20170714\\_7876798.htm](http://www.sz.gov.cn/szsfzyjzx/jczx/zjgd_1/201707/t20170714_7876798.htm).
- [5] IS. Chettri and D. Sharma, "Pre-pack optimization: Increasing supply chain efficiency," 2018, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.94.6116&rep=rep1&type=pdf>.
- [6] S. W. Sung, Y. J. Jang, J. H. Kim, and J. Lee, "Business analytics for streamlined assort packing and distribution of fashion goods at kolon sport," *Interfaces*, vol. 47, no. 6, pp. 555–573, 2017.
- [7] A. Vakhutinsky, S. Subramanian, Y. Popkov, and A. Kushkuley, *Retail pre-pack optimizer*, U.S. Patent, 2012.
- [8] M. Hoskins, R. Masson, G. Gauthier Melançon, J. E. Mendoza, C. Meyer, and L. Rousseau, "The PrePack Optimization Problem," in *Integration of AI and OR Techniques in Constraint Programming*, vol. 8451 of *Lecture Notes in Computer Science*, pp. 136–143, Springer International Publishing, Cham, 2014.
- [9] M. Fischetti, M. Monaci, and D. Salvagnin, "Mixed-integer linear programming heuristics for the prepack optimization problem," *Discrete Optimization*, vol. 22, no. part A, pp. 195–205, 2016.
- [10] S. W. Sung and Y. J. Jang, "Heuristic for the assort-packing and distribution problem in the fashion apparel industry," *International Journal of Production Research*, pp. 1–18, 2017.
- [11] P. Wang, *Pack optimization problem: Models and solution methods*, 2010, Dissertations & Theses.
- [12] R. W. Pratt, "Computer-implemented systems and methods for pack optimization," *SAS Institute, US20090271241 A1*, 2009.
- [13] G. Clarke and J. W. Wright, "Scheduling of vehicles from a central depot to a number of delivery points," *Operations Research*, vol. 12, no. 4, pp. 568–581, 1964.
- [14] W. Y. Szeto, Y. Wu, and S. C. Ho, "An artificial bee colony algorithm for the capacitated vehicle routing problem," *European Journal of Operational Research*, vol. 215, no. 1, pp. 126–135, 2011.
- [15] J. Jin, T. G. Crainic, and A. Lokketangen, "A parallel multi-neighborhood cooperative tabu search for capacitated vehicle routing problems," *European Journal of Operational Research*, vol. 222, no. 3, pp. 441–451, 2012.
- [16] S. Creemers, G. Woumans, R. Boute, and J. Beliën, "Tri-vizor uses an efficient algorithm to identify collaborative shipping opportunities," *Interfaces*, vol. 47, no. 3, 2017.
- [17] L. Zhen, "Tactical berth allocation under uncertainty," *European Journal of Operational Research*, vol. 247, no. 3, pp. 928–944, 2015.
- [18] 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.
- [19] J. Lysgaard, "The pyramidal capacitated vehicle routing problem," *European Journal of Operational Research*, vol. 205, no. 1, pp. 59–64, 2010.
- [20] R. Liu, Z. Jiang, H. Hu, and S. Yao, "A memetic algorithm for the close-open mixed vehicle routing problem," in *Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management, IEEM2010*, pp. 728–732, China, December 2010.

## Research Article

# Optimal Strategies for Manufacturers with the Reference Effect under Carbon Emissions-Sensitive Random Demand

Baojun Zhang <sup>1</sup>, Shaojian Qu <sup>1</sup>, Panpan Li,<sup>1</sup> and Ripeng Huang<sup>1,2</sup>

<sup>1</sup>Business School, University of Shanghai for Science and Technology, Shanghai, China

<sup>2</sup>School of Mathematics and Finance, Chuzhou University, Chuzhou, China

Correspondence should be addressed to Shaojian Qu; [qushaojian@usst.edu.cn](mailto:qushaojian@usst.edu.cn)

Received 21 August 2018; Accepted 17 September 2018; Published 10 October 2018

Academic Editor: Lu Zhen

Copyright © 2018 Baojun 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.

In this paper, we study the optimal strategies for a newsvendor system with joint reference effect, carbon emissions-sensitive random demand, and strategic customers' behavior. The newsvendor's decisions are to determine the selling price, production quantity, and carbon emissions under exogenous and endogenous price cases, respectively. We also explore how the loss aversions affect the newsvendor's decisions. It is shown that the newsvendor has a uniquely optimal policy. The influence of the reference effect makes the final decisions deviate from the optimal solutions of the classical model. Furthermore, the sensitivity analysis indicates that the loss aversions have a great impact on the newsvendor's decisions. Finally, more managerial implications are derived by numerical simulations.

## 1. Introduction

Environment and resource issues are the hot spot and new focus of games in the 21st century. Efforts of all countries in the world on energy conservation and emission reduction have become one of the most efficient methods to win the competition for economic strength and the dominance of the rules governing international trade. In particular, the issue of global warming has received increasing attention. Since the main cause of global warming is carbon dioxide emissions, controlling carbon emissions has become an effective way to curb global warming (<http://www.nrdc.org/globalwarming/gsteps.asp>). With the popularity of carbon footprint and the improvement of consumer's awareness of environmental protection, consumers are concerned about the product's carbon emissions besides the price [1]. Researchers recognize that consumers are passionate about the product's carbon footprint labeling and willing to pay higher prices than regular products [2, 3]. In the face of carbon sensitive consumers, enterprises will invest in green technology to improve their own profits [4].

In business practice and academic research, a growing number of people find that optimal solutions of the classic newsvendor model deviate from actual operation

and management decisions [5–7]. We know that the classic newsvendor model is based on a strong assumption that the newsvendor is completely rational. But in reality the newsvendor is not completely rational, often showing a lot of psychological behaviors. Due to these behavioral factors, final decisions of the newsvendor deviate from the optimal solution of the classic model. At present, researches on the behavior factors of the newsvendor mainly include bounded rationality [8–10], decision bias [11, 12], the reference effect [13, 14], overconfidence [15, 16], and the fairness [17, 18]. The reference effect has been extensively studied in recent years. Prior studies find that the reference effect of prospect theory plays an important role in the analysis of newsvendor's decision-making behavior [13].

In order to avoid the loss of the surplus stock, enterprises usually clean up excess products at a discount price. With the help of smartphones, computers, and various price trackers, consumers find it easier than ever to get pricing information, to decipher sales patterns, and to pick the best time to buy. The behavior that consumers choose to wait for lower prices to buy to maximize consumer surplus is defined as strategic consumer behavior [19–22]. Du and Chen [23] point out that strategic consumer behavior will diminish firms' opportunity to adopt price skimming strategy. Past studies empirically

TABLE 1: Comparisons with other recent researches.

Research papers	Carbon emissions	Reference effect	Strategic customer behavior	Stockout loss
Wu et al. [20]		√	√	
Jiang and Chen [25]	√		√	
Cao et al. [26]		√		√
Liu et al. [14]		√	√	
Our paper	√	√	√	√

confirm that the ignorance of strategic customer behavior by retailers will result in about 20 percent losses of their total profit [24].

At present, there are few studies to consider joint reference effect, strategic customer behavior, and the reference effect in supply chain operation management. Previous literatures show that it is very meaningful to study the newsvendor's reference effect and consumers' strategic behavior, which are the two-way behavioral factors that influence the operation management [14]. However, in the literatures integrating strategic customer behavior and the reference effect, they only consider one psychological loss (i.e., clearance loss). We choose an appropriate reference point of newsvendor's profit so that the model can be combined with both clearance loss and stockout loss. In fact, we find that if the newsvendor's sensitivities to these two kinds of losses caused by the reference effect are different, optimal decisions are quite different. All in all, our work is different from the previous studies, as we innovatively integrate the reference effect, two kinds of psychological losses caused by reference effect and strategic customer behavior.

As shown in Table 1, we compare the relevant literatures closely related to our work.

Wu et al. [20] examine the reference effect on a retailer's dynamic pricing and inventory strategies with strategic consumers. They find that consumers can learn about the history of the retailer's discounted prices, form their reference prices, and pick the best time to buy. Different from their research, we study the enterprise's reference effect. Firstly, the enterprise will set a profit reference point as his expectation. Then he would compare the difference between actual profits and the reference point. Finally, the enterprise will make decisions based on the expected total utility that includes actual profits and psychological satisfaction.

Jiang and Chen [25] study newsvendor decision problem with strategic consumer behavior under carbon emissions-sensitive demand. They conclude that if demand changes from price-sensitive to carbon emissions-sensitive, the manufacturer's optimal prices are the same but optimal production, optimal unit carbon emissions, and maximum expected profit go down. Although they consider the behavioral factor of strategic consumers, they assume that the manufacturer is perfect rationality. Thus we introduce the manufacturer's reference effect into the model and analyze the influence of this behavioral factor on his decisions.

Cao et al. [26] investigate the newsvendor model with the reference effect and derive joint inventory, pricing, and advertising decision. They show that the loss-averse solutions are different from the loss-neutral solutions since loss aversion

behaviors of a newsvendor affect the order quantity, pricing, and advertising decisions; specifically, the order quantity increases with the stockout aversion parameter and decreases with the surplus aversion parameter. When the price is endogenous, they set some variables that are fixed. Their work is not in a context of energy-saving and emission-reduction. We think an extended study based on their model with carbon emissions-sensitive demand is interesting. Besides that, our study differs from their study in that we introduce strategic consumer behavior to study newsvendor's decisions under endogenous price case.

Liu et al. [14] utilize the reference effect to model newsvendor's decision-making behavior with strategic customers. They find that the ordering quantity and the pricing strategy are influenced by newsvendor's reference effect, loss aversion, product cost, and salvage price. We choose a different reference point from them and analyze the influences of two kinds of losses caused by the reference point on the newsvendor's decisions. In fact, as the newsvendor's sensitivities to these two kinds of losses are different, his decisions will be quite different.

The reference effect in the newsvendor system has been widely studied. However, to our knowledge, few studies investigate influences of the reference effect on low-carbon manufacturing with strategic customers behavior. Our study aims to address the above knowledge gaps in the operations management literature. The newsvendor usually chooses one reference profit point according to the market environment and his management strategy. So the total utility of newsvendor consists of the realized profit and a psychological component that captures how the realized payoff compares with the reference profit point. Only when the production quantity equals the demand can the newsvendor achieve his maximum profit. If the newsvendor regards the theoretical maximum profit as the reference point and the production quantity is not equal to the demand, he may feel extra losses in addition to the actual economic loss. Due to the reference effect, we consider two kinds of losses, namely, (i) the clearance loss (caused by overproduction) and (ii) the stockout loss (caused by underproduction). We investigate the newsvendor's joint production, pricing, and carbon emissions decision problem in two cases of exogenous and endogenous price, respectively. Furthermore, we want to answer the following questions.

(1) Is there an optimal solution in this newsvendor system?

(2) Does the reference effect have an impact on newsvendor's decisions? If there is an impact, how does it affect newsvendor's decisions?

(3) What advices should we give to this newsvendor?

TABLE 2: Notation.

Notation	Explanation
$p$	Unit price
$c$	Unit product cost
$v$	Unit salvage price of the product which is an exogenous variable
$q$	The production quantity of the manufacturer
$s$	Unit shortage cost
$D(e, \varepsilon)$	The random demand and $D = a - bp - ke + \varepsilon$
$f(\cdot)$	Probability density function of $\varepsilon$
$F(\cdot)$	Distribution function of $\varepsilon$
$u$	The customers' utility from consuming the product
$e_0$	Unit carbon emissions without green technology investment
$e$	Unit carbon emissions with green technology investment
$h$	The coefficient of green technology investment
$I(e)$	Green technology investment

We find that there is always a unique optimal solution. The optimal production quantity of the manufacturer is different from that of the classic model. The manufacturer's optimal production quantity decreases with the clearance loss aversion  $\alpha$  but increases with the stockout loss aversion  $\beta$ . When the price is exogenous, the optimal unit carbon emission is independent of the reference effect. Nevertheless, the optimal unit carbon emission decreases with  $\alpha$  but increases with  $\beta$  when the price is endogenous. And, the optimal unit price increases with  $\alpha$  but decreases with  $\beta$ . We also find that when  $\alpha = \beta$ , the manufacturer's optimal production is a fixed value and is equal to the value without the reference effect. If the newsvendor's sensitivities to these two kinds of losses are different, there is a big divergence in his decisions.

The rest of this paper is structured as follows. Section 2 specifies the model. Section 3 analyzes the model under the exogenous price. Section 4 solves the newsvendor system under the endogenous price. Section 5 conducts numerical simulations. Section 6 concludes the work.

## 2. The Model

In this paper, the notations are included in Table 2. We set up a newsvendor model with the reference effect, which includes a manufacturer who sells products to consumers directly under carbon emissions-sensitive random demand. In the context of the above, we investigate production, pricing, and carbon emissions strategies for the manufacturer. Because of the influence of psychological behavior, the manufacturer is driven by the economic payoff and loss aversion. Referring to previous studies [13, 27, 28], we characterize the total utility of the manufacturer as follows:

$$Utility = economic\ payoff + psychological\ satisfaction. \quad (1)$$

Utility consists of the realized profit and a psychological component that captures how the realized payoff compares with the reference payoff. We will further present the profit and psychological satisfaction of the manufacturer in detail in the following sections.

**2.1. Economic Payoff.** In this section, we assume that each customer purchases one product at most. In the joint production, pricing, and carbon emissions, the manufacturer sets the production quantity  $q$  at a unit product cost  $c$  and sells directly to consumers at a unit price  $p$ . We assume that  $q$  cannot be observed by the consumers. As the selling season is very short, the inventory cannot be replenished during the sales period. We suppose that the consumers' demand is sensitive to carbon emissions, which is affected not only by the price  $p$  but also by unit carbon emission  $e$ . Referring to previous studies [25, 29], we suppose the demand function of the manufacturer as follows:

$$D(p, e, \varepsilon) = a - bp - ke + \varepsilon, \quad (2)$$

where  $a > 0$  denotes potential market size and  $b \geq 0$  and  $k > 0$  represent the demand sensitivity of unit price and unit carbon emissions, respectively.  $\varepsilon$  is a random variable and represents the stochastic part of the market,  $\varepsilon \in [A, B]$ . The probability density function and cumulative distribution function (CDF) of  $\varepsilon$  are  $f(x)$  and  $F(x)$ , respectively, where  $f(x) > 0$  on  $[A, B]$ . The unique complementary and inverse functions of CDF are  $\bar{F}(x) = 1 - F(x)$  and  $F^{-1}(x)$ , respectively. Further,  $F(x)$  is twice differentiable, strictly increasing, and absolutely continuous.  $F(x)$  satisfies an increasing failure rate (IFR), i.e.,  $\partial g(x)/\partial x > 0$ , where  $g(x) = f(x)/\bar{F}(x)$  [30]. In order to ensure positive demand within certain range of  $p$ , we require that  $A > -a$ .

In response to the change of customers demand, the manufacturer can invest in green technology to reduce carbon emissions per unit product. Initial unit carbon emission is  $e_0$ , and the unit carbon emissions after green technology investments of the manufacturer is  $e$ ,  $e_0 > e$ . Referring to previous studies [25, 29], we suppose that the manufacturer is struggling to reduce unit carbon emissions  $e_0$  to a level  $e$  with cost of effort given by

$$I(e) = h(e_0 - e)^2, \quad (3)$$

where,  $h$  is a positive coefficient and captures the manufacturer's efficiency in carbon emissions reduction. Obviously, if

the manufacturer wants to reduce unit carbon emissions to a relatively low level, he needs to increase investment in green technology. So the manufacturer needs to weigh the benefits of rising demand and technology investments.

Moreover, if there is a consumer demand that is not met at the end of the selling season, a shortage cost  $s$  will arise. On the contrary, if there is surplus stock at the end of the selling season, the manufacturer will clean up products with the salvage price  $v$ . Because the parameters must meet some conditions to make sense, we assume:  $p > c > v > 0, s > c$ .

Therefore, the profit function of the manufacturer, denoted as  $\pi$ , can be written as

$$\pi = \begin{cases} (p-c)D - (c-v)(q-D) - h(e_0-e)^2, & D < q, \\ (p-c)q - s(D-q) - h(e_0-e)^2, & D \geq q. \end{cases} \quad (4)$$

For convenient calculation and clear results, we introduce the inventory factor  $z = q - (a - bp - ke)$  into the model. It represents the risk-free inventory level [25, 26, 31]. So the profit function of the manufacturer can be rewritten as

$$\pi = \begin{cases} (p-v)(a-bp-ke+\varepsilon) - (c-v)(a-bp-ke+z) - h(e_0-e)^2, & \varepsilon < z, \\ (p-c+s)(a-bp-ke+z) - s(a-bp-ke+\varepsilon) - h(e_0-e)^2, & \varepsilon \geq z. \end{cases} \quad (5)$$

When the stochastic part of the market  $\varepsilon$  is lower than the risk-free inventory level  $z$ , the profit function of the manufacturer is  $(p-v)(a-bp-ke+\varepsilon) - (c-v)(a-bp-ke+z) - h(e_0-e)^2$ . When  $\varepsilon$  is greater than  $z$ , the profit function of the manufacturer can be rewritten as  $(p-c+s)(a-bp-ke+z) - s(a-bp-ke+\varepsilon) - h(e_0-e)^2$ .

**2.2. Psychological Satisfaction.** Instead of being completely rational, decision makers in real world often show a lot of psychological behaviors. Due to the influence of these psychological behaviors, decision maker's final decisions deviate from optimal solutions of the classical model. The reference effect is an important psychological concept which receives increasing attention in the operation management, especially in behavioral operation management in recent years. Long and Nasiry [13] prove that the newsvendor's reference effect on profits has a significant impact on newsvendor's decision-making behavior. Referring to previous studies [13, 14], we characterize the psychological satisfaction by a piecewise-linear value function as

$$v(y) = \begin{cases} \eta y, & \text{if } y \geq 0, \\ \lambda \eta y, & \text{if } y < 0, \end{cases} \quad (6)$$

where  $\eta \in [0, 1]$  characterizes the strength of the reference effect, and  $\lambda \in [1, \infty)$  is the coefficient of loss aversion. A higher value of  $\eta$  denotes being more sensitive to the deviation of the reference point; the larger  $\lambda$  is, the more the loss aversion of policymakers is.

The manufacturer usually subjectively chooses the profit reference point based on factors such as market environment and his business strategy. Under the above conditions, the theoretical maximum profit of the manufacturer is

$$\pi_{max} = (p-c)D - h(e_0-e)^2. \quad (7)$$

The manufacturer does not choose a negative reference point because the manufacturer seeks to maximize profits. So the manufacturer will choose one point as a reference point for the profit in  $[0, \pi_{max}]$ .

We regard the theoretical maximum profit as the reference point for the manufacturer, i.e.,  $\pi_r = \pi_{max} = (p-c)D - h(e_0-e)^2$  [26]. Because the realized profit of the manufacturer is less than his expectation, the manufacturer may feel extra loss in addition to the actual economic loss, i.e.,  $\pi \leq \pi_r$ . From (6),  $y = \Delta\pi = \pi - \pi_r \leq 0$ , so  $v(y) = \lambda\eta y$ . The manufacturer's underproduction and overproduction will make the actual profit lower than the reference point, and the manufacturer may have a different response to these two kinds of losses.

When  $q > D$ , the loss caused by overproduction is defined as the clearance loss  $\Delta\pi_1$ . We define the clearance loss aversion utility as  $v(\Delta\pi_1) = \lambda_1\eta\Delta\pi_1 \triangleq \alpha\Delta\pi_1$ , where  $\alpha$  shows the degree of decision-makers' aversion on the clearance loss with the reference effect and  $\alpha \geq 0$ . So, the clearance loss aversion utility can be denoted by

$$v(\Delta\pi_1) = \alpha\Delta\pi_1 = -\alpha(c-v)(q-D). \quad (8)$$

Similarly, when  $q \leq D$ , the loss caused by underproduction is defined as the stockout loss  $\Delta\pi_2$ . We define the stockout loss aversion utility as  $v(\Delta\pi_2) = \lambda_2\eta\Delta\pi_2 \triangleq \beta\Delta\pi_2$ , where  $\beta$  shows the degree of decision-makers' aversion on the stockout loss with the reference effect and  $\beta \geq 0$ . So, the stockout loss aversion utility can be denoted by

$$v(\Delta\pi_2) = \beta\Delta\pi_2 = -\beta(p-c+s)(D-q). \quad (9)$$

So, in this paper,  $v(y)$  can be written as

$$v(y) = \begin{cases} -\alpha(c-v)(q-D), & \text{if } q > D, \\ -\beta(p-c+s)(D-q), & \text{if } q \leq D. \end{cases} \quad (10)$$

If  $\alpha = \beta$ , the manufacturer has the same sensitivity to the clearance loss and the stockout loss. When  $\alpha > \beta$ , the manufacturer is more sensitive to the clearance loss; that is, he pays more attention to the actual loss. On the contrary, when  $\alpha < \beta$ , the manufacturer is more sensitive to the stockout loss; that is, he pays more attention to the potential loss.

2.3. *The Total Utility.* Thus the total utility function of the manufacturer is

$$U = \pi + v(\Delta\pi_i), \quad i = 1, 2. \quad (11)$$

Furthermore, if the actual demand is lower than the production quantity, that is,  $q > D$ , the total utility function is

$$U_{q>D} = \pi_{q>D} + v(\Delta\pi_1). \quad (12)$$

$$U = \begin{cases} (p-c)D - (1+\alpha)(c-v)(q-D) - h(e_0-e)^2, & D < q, \\ (p-c)q - s(D-q) - \beta(p-c+s)(D-q) - h(e_0-e)^2, & D \geq q. \end{cases} \quad (14)$$

Similar to (5), we rewrite the above function with the inventory factor  $z = q - (a - bp - ke)$  as

$$U = \begin{cases} (p-c)(a - bp - ke + \varepsilon) - (1+\alpha)(c-v)(z - \varepsilon) - h(e_0-e)^2, & \varepsilon < z, \\ (p-c)(a - bp - ke + z) - s(\varepsilon - z) - \beta(p-c+s)(\varepsilon - z) - h(e_0-e)^2, & \varepsilon \geq z. \end{cases} \quad (15)$$

The expected total utility function of the manufacturer is

$$\begin{aligned} E(U) &= (p-c)(a - bp - ke + \mu) - h(e_0-e)^2 \\ &\quad - (1+\alpha)(c-v)\Lambda(z) \\ &\quad - (1+\beta)(p-c+s)\Theta(z), \end{aligned} \quad (16)$$

where,  $\mu = \int_A^B xf(x)dx$ ,  $\Lambda(z) = \int_A^z (z-x)f(x)dx$ ,  $\Theta(z) = \int_z^B (x-z)f(x)dx$ .

Our ultimate goal is to maximize expected total utility of the manufacturer. In the following, we will explore optimal decisions of the manufacturer under price exogenous and price endogenous cases and discuss how the reference effect affects the manufacturer's decisions.

### 3. Analysis under the Exogenous Price

In some industries, the selling price of products is determined by external factors such as competitive market or government, and manufacturers do not have the ability to make the selling price of products. Low-carbon products usually involve the national economy and the people's livelihood. In order to guarantee the rights and interests of the people, in many countries, this kind of resources is priced by the government. Because of the fierce competition in the market today, the exogenous price is also common. In this case, we seek the optimal production quantity and the optimal carbon emissions. As the manufacturer has no pricing power, he can only determine production and carbon emissions by maximizing expected utility. In this case, the strategic consumer behavior has no effect on the decisions of the manufacturer.

If the actual demand is higher than the production quantity, that is,  $D \geq q$ , the total utility function is

$$U_{D \geq q} = \pi_{D \geq q} + v(\Delta\pi_2). \quad (13)$$

Therefore, the total utility function of the manufacturer can be expressed as

Given the price, we explore best decisions of the manufacturer to achieve his maximum expected utility. We can get the following propositions. We note that proofs of these propositions here and after can be founded in the Appendix.

**Proposition 1.** *When the price  $p$  is exogenous, there is a unique optimal inventory factor  $z^*$ , and it meets the following condition:*

$$F(z^*) = \frac{(1+\beta)(p-c+s)}{(1+\alpha)(c-v) + (1+\beta)(p-c+s)}, \quad (17)$$

and given price  $p$ , there is a unique optimal unit carbon emissions  $e^*$ , that is,

$$e^* = e_0 - \frac{k(p-c)}{2h}. \quad (18)$$

From Proposition 1, we can see that when the price is exogenous, the optimal unit carbon emission is independent of the reference effect. However, when the price is exogenous, the optimal inventory factor is related to aversions of the clearing loss and stockout loss. We also find that as the demand sensitivity of unit carbon emissions  $k$  increases, the optimal unit carbon emission  $e^*$  falls, holding all else constant. This provides an economic explanation for government to launch a massive campaign to educate all citizens to improve their awareness of environmental protection.

We can further obtain the inventory factor as follows:

$$z^* = F^{-1} \left( \frac{(1+\beta)(p-c+s)}{(1+\alpha)(c-v) + (1+\beta)(p-c+s)} \right). \quad (19)$$

The production quantity of the manufacturer meets the following condition:  $q = z + (a - bp - ke)$ . With the above

analysis, we can get the optimal production quantity of the manufacturer  $q^*$ ,

$$q^* = a - bp - k \left( e_0 - \frac{k(p-c)}{2h} \right) + F^{-1} \left( \frac{(1+\beta)(p-c+s)}{(1+\alpha)(c-v) + (1+\beta)(p-c+s)} \right). \quad (20)$$

We find that the optimal production quantity of the manufacturer is related to the market environment, the sensitivity of consumers to the price and carbon emissions, the efficiency in carbon emissions reduction, and the manufacturer's loss aversions. Specifically, the optimal production quantity increases in potential market size. The manufacturer adjusts the optimal output according to the distribution of random variables in demand. And, consumers' sensitivities for the price and carbon emissions have negative influences on the optimal production. Then, we will discuss the effect of loss aversions on the optimal production.

**Proposition 2.** When the manufacturer has the same sensitivity of aversion between the clearance loss and the stockout loss ( $\alpha = \beta$ ), the optimal production quantity is not related to the above loss aversion behavior, and it is the same as the optimal production quantity of the manufacturer without the reference effect.

For the above interesting phenomenon, this is due to the difference between the two kinds of loss aversions to the optimal production. When the manufacturer showed the same feeling of aversion for two kinds of losses, the two effects counteract each other. Cao et al. [26] derive a similar result for the loss aversion caused by the newsvendor's reference effect in terms of the exogenous price.

**Proposition 3.** When the price is exogenous and  $\alpha$  is not equal to  $\beta$ , the manufacturer's optimal production quantity decreases with  $\alpha$  but increases with  $\beta$ .

From Proposition 3, we verify the explanation of Proposition 2. We find that the optimal output decision is different from the optimal solution of the classic model on account of the manufacturer's reference effect. Zhang et al. [30] find the newsvendor's optimal order quantity decreases in the loss aversion coefficient in terms of the exogenous price. However, we find that when the manufacturer is more sensitive to the clearance loss, he will reduce the production quantity accordingly. While he will increase the output in order to eliminate the potential loss if he is more sensitive to the stockout loss.

#### 4. Analysis under the Endogenous Price

When the price is endogenous, we assume that the manufacturer faces strategic consumers. The whole sales period is divided into two phases: in the first phase, the manufacturer sells the products at full price, while, in the second phase, the manufacturer empties the products at the salvage price. Frequent discounts make consumers more selective when they buy products. As consumers can foresee that the product

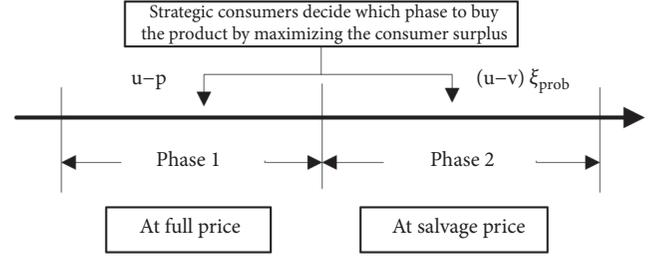


FIGURE 1: The sequence of events.

may be discounted in the future, they choose to wait at the full price and to buy when the product is discounted. The behavior that consumers choose to wait for buying at lower prices to maximize consumer surplus is known as strategic consumer behavior [14, 25].

We assume that consumers are homogeneous; that is, all consumers have the same utility to the product,  $u$ . The utility satisfies the following condition:  $u > p$ . This condition states that there is a positive consumer surplus when a product is sold to customers. Further, for consumers, the utility of the product in the first or second phase will remain unchanged. Strategic consumers can buy discounted products at the second phase with a certain probability. The beliefs of consumers in this probability is  $\xi_{prob}$ . Strategic consumers determine the stage of purchasing a product by comparing the expected consumer surplus over the two phases. Figure 1 shows the sequence of events. The consumers' reservation price is  $r$ , which is the private information of consumers and cannot be observed by the manufacturer. When the consumer reservation price is larger than the retail price, the consumer may buy the product at full price. The beliefs of the manufacturer over the consumers' reservation price is  $\delta_r$ .

Obviously, the manufacturer will set  $p = \delta_r$  to maximize the profit. According to the definition of rational expectation equilibrium [32, 33], the solution of rational expectation equilibrium  $(p, z, r, \delta_r, \xi_{prob})$  must meet the following conditions: (i)  $r = u - (u - v)\xi_{prob}$ ; (ii)  $p = \delta_r$ ; (iii)  $z, e$  are a solution of  $\arg\max_{z,e} E(U)$ ; (iv)  $\xi_{prob} = F(z)$ ; and (v)  $\delta_r = r$ .

Conditions (i), (ii), and (iii) show that the manufacturer and strategic consumers will choose the action to maximize their own utility. Conditions (iv) and (v) ensure that the solution satisfies the rational expectation hypothesis; that is, the actual economic situation is consistent with the people's expectation. Many researchers denote that the nature of the problem is a static game and the solution of rational expectation equilibrium satisfies the definition of Nash equilibrium [33, 34]. From the above analysis we can get

$$p = u - (u - v) F(z). \quad (21)$$

Price  $p$  is a function of  $F(z)$ . We have proved that given  $p$ , the total utility function of the manufacturer  $E[U(z, p, e)]$  is jointly strict concave in  $z$  and  $e$ . That is, there are unique  $z, e$  which maximize  $E[U(z, p, e)]$ . So, we can obtain

$$\begin{aligned}
p &= u - (u - v)F(z), \\
(1 + \beta)(p - c + s)(1 - F(z)) - (1 + \alpha)(c - v)F(z) &= 0,
\end{aligned} \tag{22}$$

$$2h(e_0 - e) - k(p - c) = 0.$$

Solving (22), we can derive the optimal production quantity, the optimal retail price, and the optimal carbon emissions of the manufacturer.

**Proposition 4.** *When the price  $p$  is endogenous and consumers have strategic behaviors, there is a unique optimal inventory factor  $z^*$ , and it meets the following condition:*

$$\begin{aligned}
F(z^*) &= \frac{(1 + \beta)(2u - c - v + s) + (1 + \alpha)(c - v) - \varphi(\alpha, \beta)}{2(1 + \beta)(u - v)},
\end{aligned} \tag{23}$$

where  $\varphi(\alpha, \beta) = (((1 + \beta)(v + s - c) + (1 + \alpha)(c - v))^2 + 4(1 + \alpha)(1 + \beta)(c - v)(u - v))^{1/2}$ .

From Proposition 4, when the price  $p$  is endogenous and consumers have strategic behavior, the loss aversions affect the optimal inventory factor. And manufacturers pay more attention to the product value determination of consumers; that is, he will adjust the risk-free inventory level according to the utility of consumers from the product.

**Proposition 5.** *When the price  $p$  is endogenous with strategic customers behavior, there is a unique optimal unit price  $p^*$ ,*

$$p^* = u - (u - v)F(z^*), \tag{24}$$

and the unique optimal unit carbon emissions  $e^*$ ,

$$e^* = e_0 - \frac{k(p^* - c)}{2h}. \tag{25}$$

From Proposition 5, we can find that the optimal unit price is related to the loss aversions, which is caused by considering the reference effect. Because the optimal unit carbon emission is related to unit price, in this case, the reference effect will affect the optimal carbon emissions of products. We find that as the demand sensitivity of unit carbon emissions  $k$  increases, the optimal unit carbon emission  $e^*$  falls, holding all else constant. Similar to previous literature such as [25], the manufacturer will be keen on green technology investment demonstrably with carbon emissions-sensitive random demand.

For convenient calculation and clear results, we define  $\Psi(\alpha, \beta) \triangleq ((1 + \beta)(2u - c - v + s) + (1 + \alpha)(c - v) - \varphi(\alpha, \beta))/(2(1 + \beta)(u - v))$ . We can further obtain the inventory factor,  $z^* = F^{-1}(\Psi(\alpha, \beta))$ . According to  $q = z + a - bp - ke$ , we can get the optimal production quantity of the manufacturer  $q^*$ ,

$$\begin{aligned}
q^* &= F^{-1}(\Psi(\alpha, \beta)) + a - ke_0 - \frac{k^2c}{2h} \\
&\quad + \left(\frac{k^2}{2h} - b\right)(u - (u - v)F(z^*)).
\end{aligned} \tag{26}$$

When the price  $p$  is endogenous and consumers have strategic behaviors, the consumers' utility from consuming the product affects the optimal production quantity. When  $k^2/2h > b$ , the optimal production quantity increases with the consumers' utility, while the optimal production quantity decreases with the consumers' utility if  $k^2/2h \leq b$ .

We can find that when  $\alpha = \beta$ ,  $F(z^*) = ((2u - c - v + s) + (c - v) - \varphi(\alpha, \beta))/(2(u - v))$ , where  $\varphi(\alpha, \beta) = (((v + s - c) + (c - v))^2 + 4(c - v)(u - v))^{1/2}$ . We find that when the manufacturer has the same sensitivity of aversion between the clearance loss and the stockout loss, optimal decisions are not related to the above loss aversion behavior, and it is the same as decisions of the manufacturer without the reference effect.

When  $\alpha \neq \beta$ , we introduce the Cramer's Rule to determine the influence of the manufacturer's loss aversion on his decisions. By using the Cramer's Rule, we can get  $\partial p/\partial \alpha > 0$ ,  $\partial F(z)/\partial \alpha < 0$ ,  $\partial p/\partial \beta < 0$ , and  $\partial F(z)/\partial \beta > 0$ . So the following proposition is obtained.

**Proposition 6.** *When the price  $p$  is endogenous, consumers have strategic behavior, and  $\alpha \neq \beta$ , the manufacturer's optimal production quantity decreases with  $\alpha$  but increases with  $\beta$ . And the optimal unit price increases with  $\alpha$  but decreases with  $\beta$ .*

Liu et al. [14] derive a result that the newsvendor's optimal order is a decreasing function of the loss aversion caused by the newsvendor's reference effect in terms of the endogenous price and strategic consumers behavior. However, we find that whether the price is exogenous or endogenous, the reference effect maintains a good consistency on manufacturer's optimal production quantity; that is,  $q^*$  decreases with  $\alpha$  and increases with  $\beta$ . In addition, when the price  $p$  is endogenous and consumers have strategic behavior, the formulation of the optimal price is affected by the reference effect. They also find that the newsvendor's optimal price is an increasing function of the loss aversion caused by the reference effect. Obviously, our conclusion is consistent with their conclusion when we do not consider the stockout loss.

From (25), we can get the following proposition.

**Proposition 7.** *When the price  $p$  is endogenous, consumers have strategic behavior, and  $\alpha \neq \beta$ , the optimal unit carbon emission decreases with  $\alpha$  and increases with  $\beta$ .*

If manufacturer's sensitivities for two loss aversions are different, his reaction to the carbon emissions-sensitive demand is opposite. When the manufacturer is more sensitive to the clearance loss, he is more motivated to invest in green technology to reduce unit carbon emissions. When the manufacturer is more sensitive to the stockout loss, he has a negative attitude towards reducing unit carbon emissions. In the next section, we will conduct numerical experiments to verify the impact of loss aversion caused by the manufacturer's reference effect on his decision making.

## 5. Numerical Simulations

In this section, we consider numerical simulations to analyze the feasibility of the analysis models, to illustrate the impact of the reference effect on the manufacturer's decision making, and to verify the propositions presented in the last section.

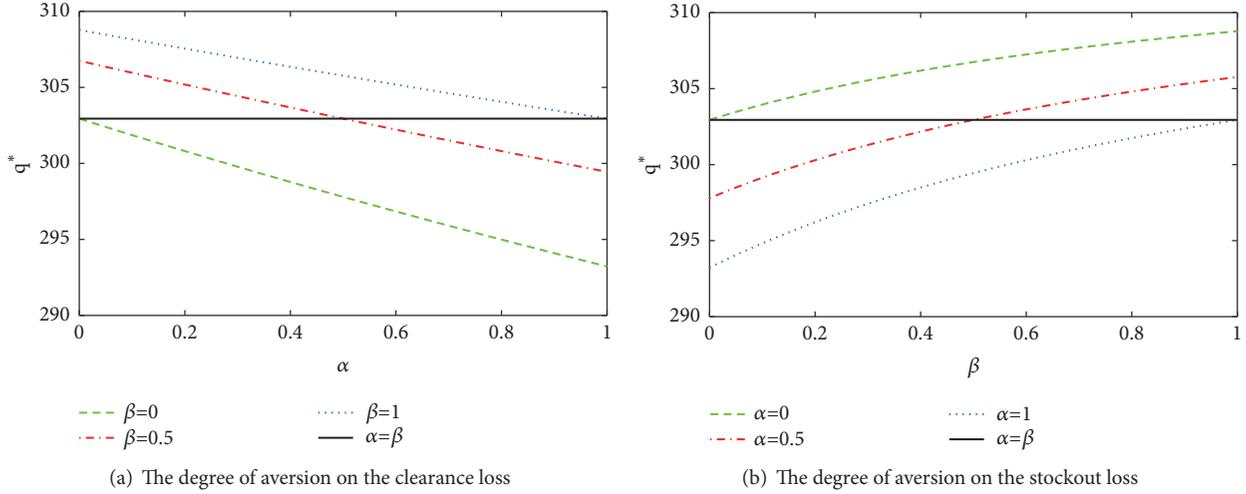


FIGURE 2: The impact of reference effect on the optimal production under the exogenous price.

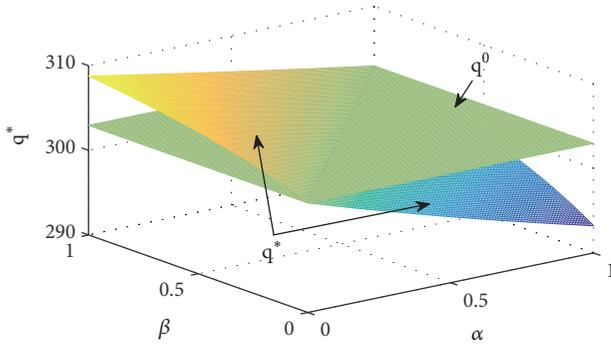


FIGURE 3: The impact of reference effect on the optimal production under the exogenous price.

We refer to the previous research results [25, 26] and specify that  $a = 200$ ,  $b = 1.5$ ,  $c = 10$ ,  $v = 5$ ,  $s = 15$ ,  $u = 50$ ,  $e_0 = 20$ ,  $h = 90$ ,  $k = 2$ . We also assume that  $\varepsilon$  follows the uniform distribution,  $\varepsilon \sim U[100, 200]$ . According to [35, 36], the loss aversion with the reference effect of the newsvendor is usually not greater than 2. Here, we denote that  $\alpha, \beta \in [0, 1]$ . Then we illustrate the optimal production quantity and carbon emissions under the exogenous price and the endogenous price, respectively. When the price is endogenous, we also analyze the optimal price.

When the price is exogenous, we assume the unit price  $p$  is 30. We can get the optimal production quantity and total carbon emissions, which are in Figures 2–5.

From Figures 2 and 3, we have verified some explanations of Propositions 2 and 3. As is illustrated in Figure 2, we find that when  $\alpha = \beta$ , the manufacturer's optimal production is a fixed value and is equal to the value without considering the reference effect. When the price is exogenous and  $\alpha$  is not equal to  $\beta$ , from Figure 2(a) the manufacturer's optimal production quantity decreases with  $\alpha$ , while from Figure 2(b) it increases with  $\beta$ . We also find that  $\alpha$  will amplify the effect of  $\beta$  on the optimal production quantity, while  $\beta$  will

reduce the impact of  $\alpha$  on the optimal production quantity. In Figure 3,  $q^0$  represents the optimal production without the reference effect under the exogenous price. We find that when  $\alpha < \beta$ , that is, when the manufacturer is more sensitive to the stockout loss, the manufacturer's optimal production is higher than the value without the reference effect. Conversely, when the manufacturer pays more attention to the clearance loss, i.e.,  $\alpha > \beta$ , the manufacturer's optimal output is lower than the value without the reference effect. We also find that  $\alpha$  will amplify the effect of  $\beta$  on the optimal production quantity, while  $\beta$  will reduce the impact of  $\alpha$  on the optimal production quantity.

Because the price is exogenous, optimal unit carbon emission is 19.78 when  $p = 30$ . From Figures 4 and 5, we can illustrate the impact of reference effect on total carbon emissions under the exogenous price. When  $\alpha$  is not equal to  $\beta$ , from Figure 4(a) we find that total carbon emission decreases with  $\alpha$ , while from Figure 4(b) it increases with  $\beta$ . When  $\alpha = \beta$ , total carbon emission is a fixed value. In Figure 5,  $q^0 e^0$  represents total carbon emissions without the reference effect under the exogenous price. From Figure 5, when  $\alpha < \beta$ , total carbon emission is higher than the value without the reference effect. If  $\alpha > \beta$ , total carbon emissions will be lower than the value without the reference effect.

When the price  $p$  is endogenous and consumers have strategic behavior, we can get the optimal price, the optimal production quantity, optimal unit carbon emissions, and total carbon emissions, which are shown in Figures 6–13, respectively.

From Figure 6, we can verify the formulation of Proposition 6; that is, the optimal unit price increases with  $\alpha$  (Figure 6(a)) but decreases with  $\beta$  (Figure 6(b)). When  $\alpha = \beta$ , the optimal unit price is a fixed value and is equal to the value without the reference effect. In Figure 7,  $p^0$  represents the optimal price without the reference effect under the endogenous price. From Figure 7, when  $\alpha < \beta$ , the optimal unit price is lower than the value that does not consider the reference effect. If  $\alpha > \beta$ , the optimal unit price will be higher than the value without the reference effect.

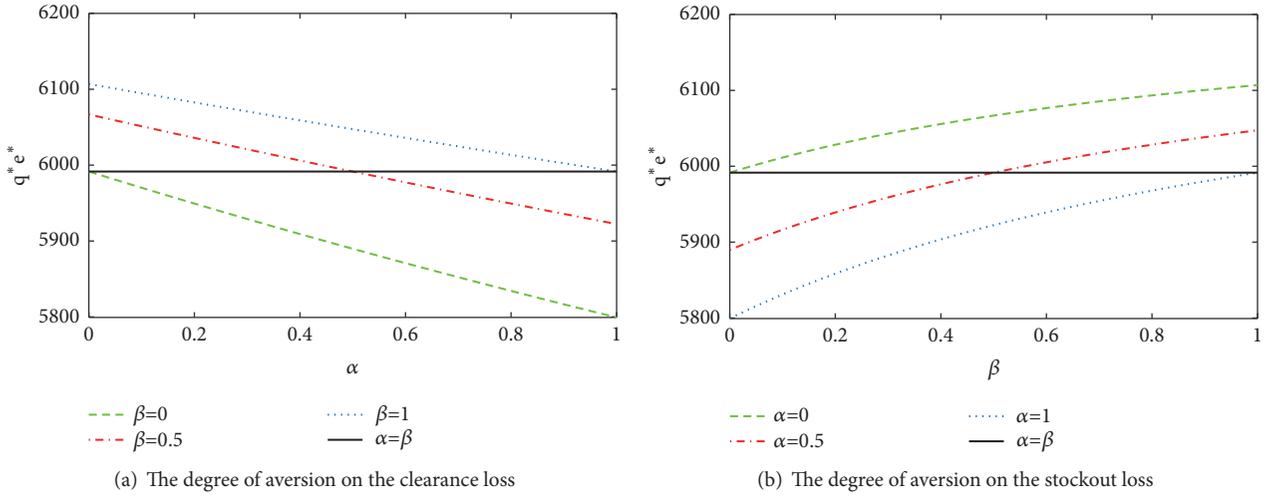


FIGURE 4: The impact of reference effect on total carbon emissions under the exogenous price.

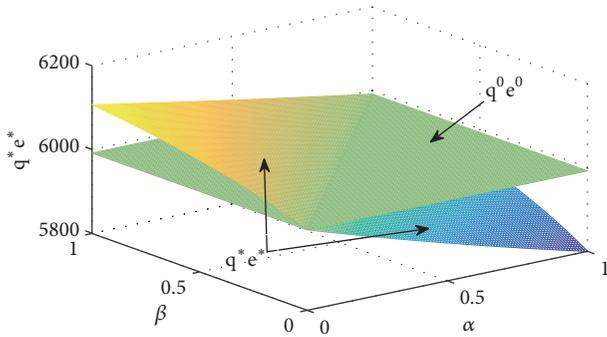


FIGURE 5: The impact of reference effect on total carbon emissions under the exogenous price.

Figures 8 and 9 demonstrate the impact of reference effect on the optimal production quantity under the endogenous price. In Figure 9,  $q^0$  represents the optimal production without the reference effect under the endogenous price. We find the reference effect maintains a good consistency on manufacturer's optimal production quantity  $q^*$  whether the price is exogenous or endogenous; that is, (i) when  $\alpha = \beta$ ,  $q^*$  is a fixed value and is equal to the value without considering the reference effect (Figure 8), (ii) when  $\alpha \neq \beta$ ,  $q^*$  decreases with  $\alpha$  (Figure 8(a)) and increases with  $\beta$  (Figure 8(b)), (iii) when  $\alpha < \beta$ ,  $q^*$  is higher than the value without the reference effect (Figure 9), and (iv) when  $\alpha > \beta$ , the result is opposite (Figure 9).

From Figure 10, we find that the optimal unit carbon emission decreases with  $\alpha$  (Figure 10(a)) and increases with  $\beta$  (Figure 10(b)) under the endogenous price. That is, when a manufacturer has a pricing power, he has an internal drive to adjust carbon emissions. When  $\alpha = \beta$ , the optimal unit carbon emission is a fixed value and is equal to the value that does not consider the reference effect. In Figure 11,  $e^0$  represents optimal unit carbon emissions without the reference effect under the endogenous price. We find that when  $\alpha < \beta$ , that is,

when the manufacturer is more sensitive to the stockout loss, the optimal unit carbon emission is higher than the value that does not consider the reference effect. Conversely, when the manufacturer pays more attention to the clearance loss, i.e.,  $\alpha > \beta$ , the optimal unit carbon emission is lower than the value without the reference effect.

Figures 12 and 13 illustrate the impact of reference effect on total carbon emissions under the endogenous price. In Figure 13,  $q^0 e^0$  represents total carbon emissions without the reference effect under the endogenous price. We find the reference effect also maintains a good consistency on total carbon emissions whether the price is exogenous or endogenous; that is, (i) when  $\alpha \neq \beta$ , total carbon emission  $q^* e^*$  decreases with  $\alpha$  (Figure 12(a)) and increases with  $\beta$  (Figure 12(b)), (ii) when  $\alpha < \beta$ ,  $q^* e^*$  is higher than the value without the reference effect (Figure 13), (iii) when  $\alpha > \beta$ ,  $q^* e^*$  is lower than the value without the reference effect (Figure 13), and (iv) when  $\alpha = \beta$ ,  $q^* e^*$  is a fixed value and is equal to the value without the reference effect (Figure 12).

## 6. Conclusions

This work considers a decision-making problem of the newsvendor system with the reference effect. Different from the classical model, the manufacturer's loss aversions are caused by the reference profit point under carbon emissions-sensitive random demand. Our ultimate goal is to maximize the expected total utility of the manufacturer and find out the changes in carbon emissions under different conditions. We explore the optimal production quantity and carbon emissions under price exogenous and price endogenous cases, respectively. When the price is endogenous, we introduce consumer strategic behavior which is more consistent with economic activities to explore the optimal price. Furthermore, we also discuss the impact of the reference effect on optimal decisions.

We find that there is always a unique optimal solution. Under exogenous price and endogenous price cases, the

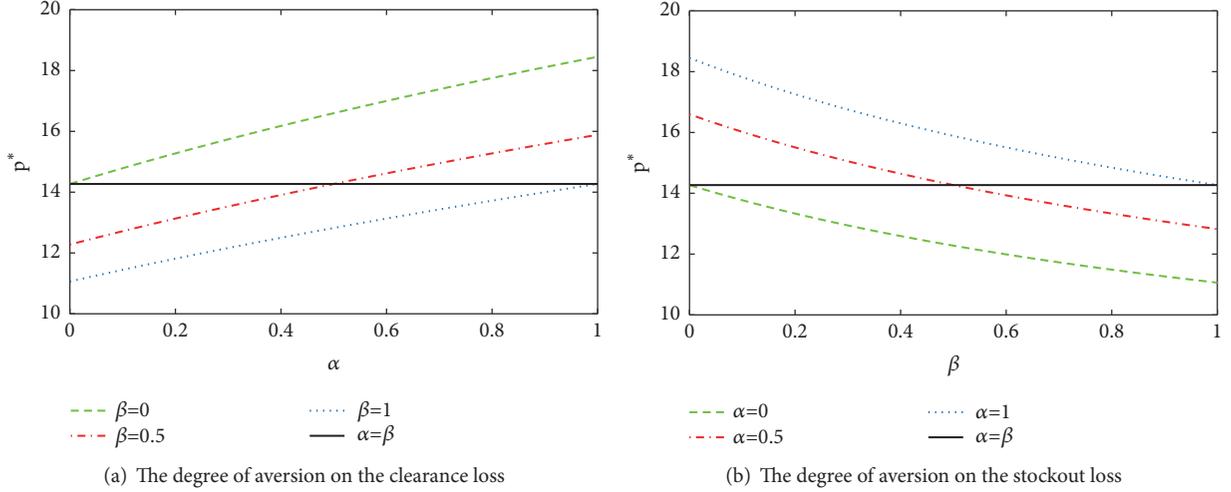


FIGURE 6: The impact of reference effect on the optimal price under the endogenous price.

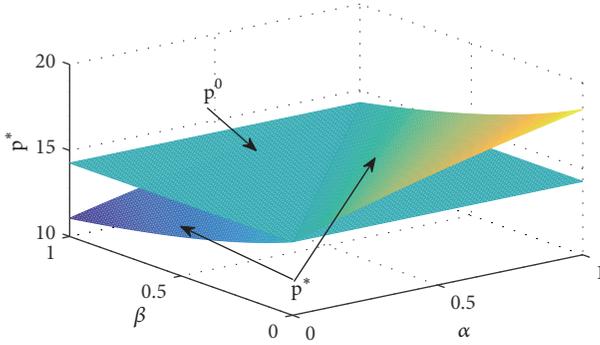


FIGURE 7: The impact of reference effect on the optimal price under the endogenous price.

optimal production quantity of the manufacturer is different from the optimal value of the classic model. Specifically, the manufacturer adjusts the optimal production due to the loss aversion caused by the manufacturer's reference profit point. When the manufacturer has a higher degree of aversion for the clearance loss  $\alpha$ , he should reduce the production. Contrarily, when the manufacturer has a higher degree of aversion for the stockout loss  $\beta$ , he should increase the production. When the price is exogenous, the manufacturer's decision for unit carbon emissions is passive; that is, decision making is affected by the competitive environment. In this case, the government can carry out technical guidance to improve the manufacturer's efficiency in carbon emissions reduction. In the market-driven environment, this measure can effectively achieve the target of emission reduction. When the price is endogenous, the decision for unit carbon emissions is influenced by the manufacturer's reference effect. Specifically, the optimal unit carbon emission decreases with  $\alpha$  but increases with  $\beta$ . The economic explanation for this is as follows: in order to avoid the clearance loss, the manufacturer reduces unit carbon emissions to attract consumers when facing the consumers who are sensitive to carbon emissions. If the manufacturer puts the consumers' loyalty at a high level,

he will lack the enthusiasm to reduce carbon emissions to the stockout loss. The government should give appropriate subsidies to such manufacturers. We also find an economic explanation for government to launch a massive campaign to educate all citizens to improve their awareness of environmental protection; that is, the optimal unit carbon emission  $e^*$  falls as the demand sensitivity of unit carbon emissions  $k$  rises, all else being equal.

We also carry out numerical studies to find more implications for manufacturer's decisions. We find that when  $\alpha = \beta$ , the manufacturer's optimal production is a fixed value and is equal to the value without the reference effect. When the manufacturer is more sensitive to the stockout loss, the manufacturer's optimal production is higher than the value without considering the reference effect. Conversely, when the manufacturer pays more attention to the clearance loss, the manufacturer's optimal output is lower than the value without the reference effect. When the price  $p$  is endogenous, the optimal unit carbon emission is equal to the value without the reference effect, if  $\alpha = \beta$ . When  $\alpha < \beta$ , the optimal unit carbon emission is higher than the value that does not consider the reference effect. Conversely, when  $\alpha > \beta$ , the optimal unit carbon emission is lower than the value without the reference effect.

However, there are still some flaws in the study. The main disadvantage is that our work does not consider the carbon emissions policy. It is far sighted to take the carbon emissions policy into consideration in future work.

## Appendix

### A. Proof of Proposition 1

When the price is exogenous, we can get the first and twice order partial differential equations of  $E(U)$  on the inventory factor  $z$  and the unit carbon emissions  $e$  from (16):

$$\begin{aligned} \frac{\partial E(U)}{\partial z} &= (1 + \beta)(p - c + s)(1 - F(z)) \\ &\quad - (1 + \alpha)(c - v)F(z), \end{aligned} \quad (\text{A.1})$$

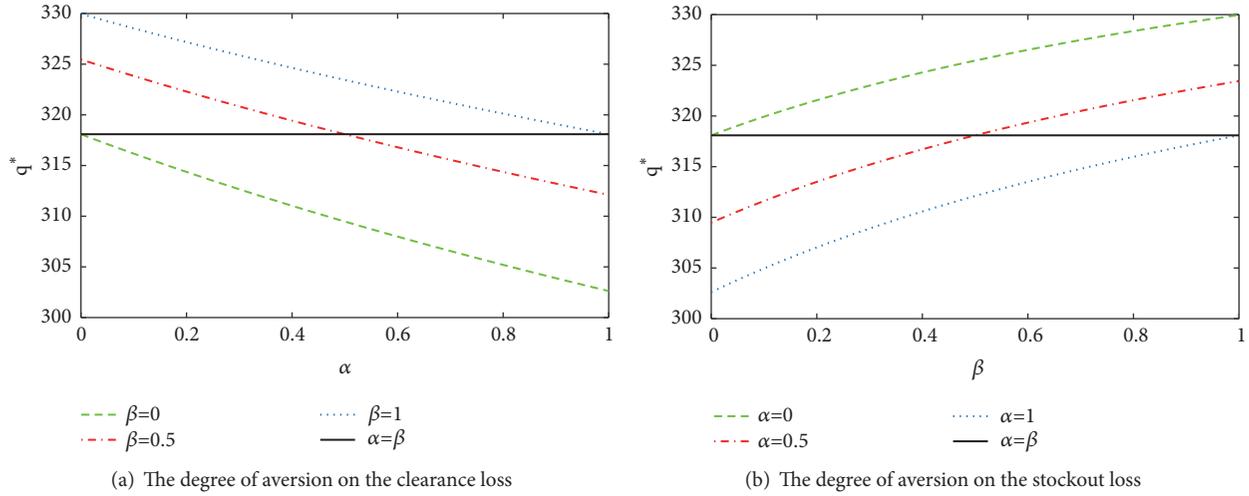


FIGURE 8: The impact of reference effect on the optimal production under the endogenous price.

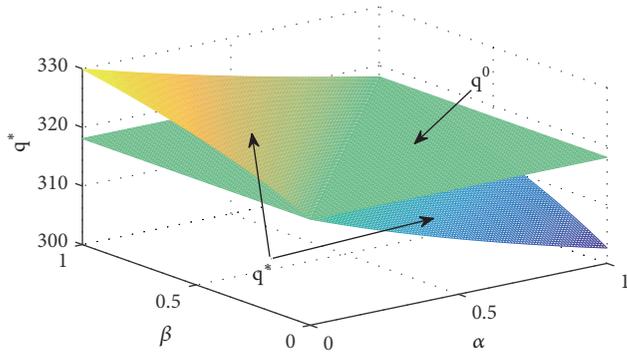


FIGURE 9: The impact of reference effect on the optimal production under the endogenous price.

$$\begin{aligned} \frac{\partial^2 E(U)}{\partial z^2} &= -((1 + \beta)(p - c + s) + (1 + \alpha)(c - v)) f(z) \\ &< 0, \end{aligned} \quad (\text{A.2})$$

$$\frac{\partial E(U)}{\partial e} = 2h(e_0 - e) - k(p - c), \quad (\text{A.3})$$

$$\frac{\partial^2 E(U)}{\partial e^2} = -2h < 0, \quad (\text{A.4})$$

$$\frac{\partial^2 E(U)}{\partial z \partial e} = \frac{\partial^2 E(U)}{\partial e \partial z} = 0. \quad (\text{A.5})$$

Thus, the Hessian matrix is

$$\begin{aligned} &\begin{vmatrix} \frac{\partial^2 E(U)}{\partial z^2} & \frac{\partial^2 E(U)}{\partial z \partial e} \\ \frac{\partial^2 E(U)}{\partial e \partial z} & \frac{\partial^2 E(U)}{\partial e^2} \end{vmatrix} \\ &= 2h((1 + \beta)(p - c + s) + (1 + \alpha)(c - v)) f(z) \\ &> 0. \end{aligned} \quad (\text{A.6})$$

So, the total utility function of the manufacturer  $E(U)$  is jointly strict concave in  $z$  and  $e$ . That is, there are unique  $z$  and  $e$  which maximize  $E(U)$ . Let  $\partial E(U)/\partial z = 0, \partial E(U)/\partial e = 0$ ; we have

$$\begin{aligned} (1 + \beta)(p - c + s)(1 - F(z)) - (1 + \alpha)(c - v)F(z) \\ = 0, \end{aligned} \quad (\text{A.7})$$

$$2h(e_0 - e) - k(p - c) = 0.$$

From (A.7), we can get

$$F(z^*) = \frac{(1 + \beta)(p - c + s)}{(1 + \alpha)(c - v) + (1 + \beta)(p - c + s)}, \quad (\text{A.8})$$

$$e^* = e_0 - \frac{k(p - c)}{2h}. \quad (\text{A.9})$$

## B. Proof of Proposition 2

When  $\alpha = \beta$ , from (20), we can get

$$\begin{aligned} q^* &= a - bp - k \left( e_0 - \frac{k(p - c)}{2h} \right) \\ &\quad + F^{-1} \left( \frac{p - c + s}{p - v + s} \right). \end{aligned} \quad (\text{B.1})$$

So, if  $\alpha = \beta$ , the optimal production quantity is not related to the above loss aversion behavior, and it is the same as the optimal production quantity of the loss-neutral manufacturer.

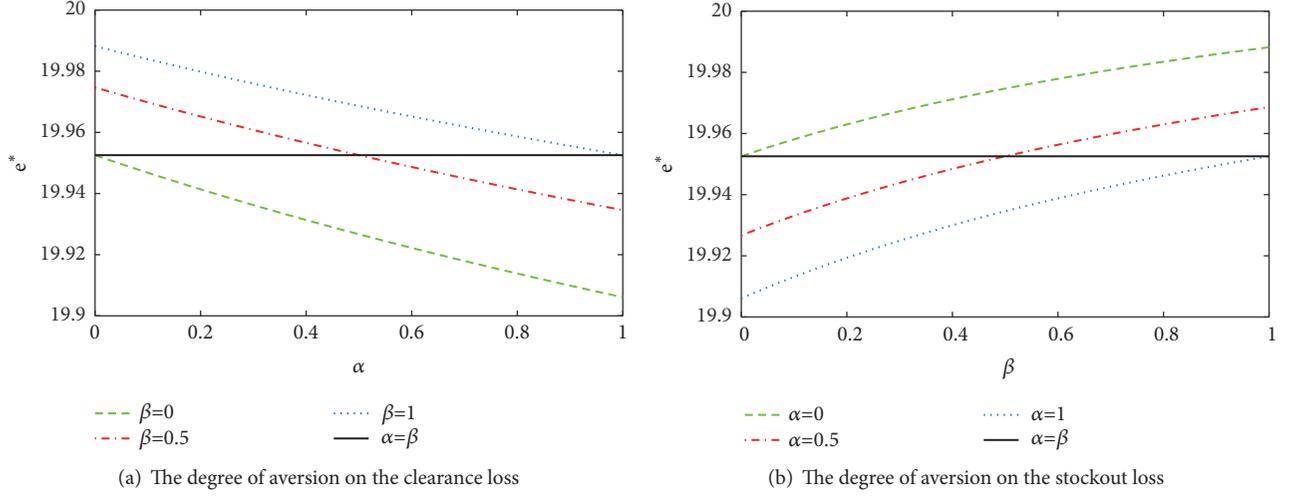


FIGURE 10: The impact of reference effect on optimal unit carbon emissions under the endogenous price.

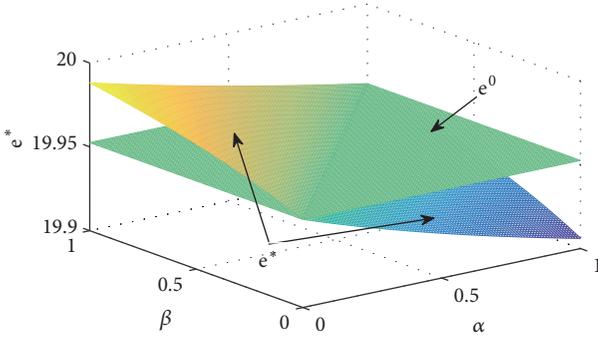


FIGURE 11: The impact of reference effect on optimal unit carbon emissions under the endogenous price.

### C. Proof of Proposition 3

The proof of the positive and negative properties of  $\partial q^*/\partial \alpha$  is equivalent to show the positive and negative properties of  $\partial F(z^*)/\partial \alpha$ .

$$\frac{\partial F(z^*)}{\partial \alpha} = \frac{-(1+\beta)(c-\nu)(p-c+s)}{[(1+\alpha)(c-\nu) + (1+\beta)(p-c+s)]^2} \quad (C.1)$$

$< 0$ .

So,  $\partial q^*/\partial \alpha < 0$ , that is, the manufacturer's optimal production quantity decreases with  $\alpha$ .

The proof of the positive and negative properties of  $\partial q^*/\partial \beta$  is equivalent to show the positive and negative properties of  $\partial F(z^*)/\partial \beta$ .

$$\frac{\partial F(z^*)}{\partial \beta} = \frac{(1+\alpha)(p-c+s)(c-\nu)}{[(1+\alpha)(c-\nu) + (1+\beta)(p-c+s)]^2} \quad (C.2)$$

$> 0$ .

So,  $\partial q^*/\partial \beta > 0$ , that is, the manufacturer's optimal production quantity increases with  $\beta$ .

### D. Proofs of Propositions 4 and 5

According to the definition of rational expectation equilibrium (Desai et al., 2004), the solution of rational expectation equilibrium  $(p, z, r, \delta_r, \xi_{prob})$  must meet the following conditions: (i)  $r = u - (u - \nu)\xi_{prob}$ ; (ii)  $p = \delta_r$ ; (iii)  $z, e$  are a solution of  $\text{argmax}_{z,e} E(U)$ ; (iv)  $\xi_{prob} = F(z)$ ; and (v)  $\delta_r = r$ .

From the above analysis we can get

$$p = u - (u - \nu)F(z). \quad (D.1)$$

Combined with Appendix A, we can get

$$\begin{aligned} p &= u - (u - \nu)F(z), \\ (1 + \beta)(p - c + s)(1 - F(z)) - (1 + \alpha)(c - \nu)F(z) &= 0, \end{aligned} \quad (D.2)$$

$$2h(e_0 - e) - k(p - c) = 0.$$

Given  $p, E(U)$  is jointly strict concave in  $z$  and  $e$ . So there are unique  $z$  and  $e$  which maximize  $E[U(z, e, p)]$ . Solving (D.2), we can derive the optimal condition of the inventory factor  $F(z^*)$ , the optimal retail price  $p^*$ , and the optimal carbon emission  $e^*$  of the manufacturer, respectively:

$$\begin{aligned} F(z^*) &= \frac{(1 + \beta)(2u - c - \nu + s) + (1 + \alpha)(c - \nu) \pm \varphi(\alpha, \beta)}{2(1 + \beta)(u - \nu)}, \end{aligned} \quad (D.3)$$

where  $\varphi(\alpha, \beta) = (((1 + \beta)(\nu + s - c) + (1 + \alpha)(c - \nu))^2 + 4(1 + \alpha)(1 + \beta)(c - \nu)(u - \nu))^{1/2}$ ,

$$p^* = u - (u - \nu)F(z^*), \quad (D.4)$$

$$e^* = e_0 - \frac{k(p^* - c)}{2h}. \quad (D.5)$$

When the price  $p$  is endogenous and consumers have strategic behavior,  $F(z^*)$  decreases in  $\alpha$  and increases with

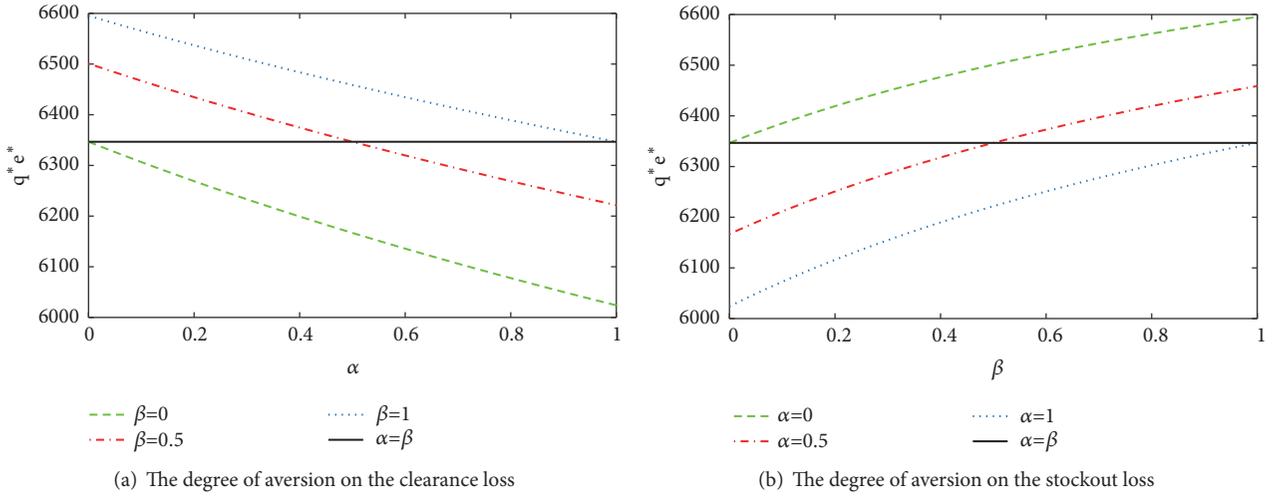


FIGURE 12: The impact of reference effect on total carbon emissions under the endogenous price.

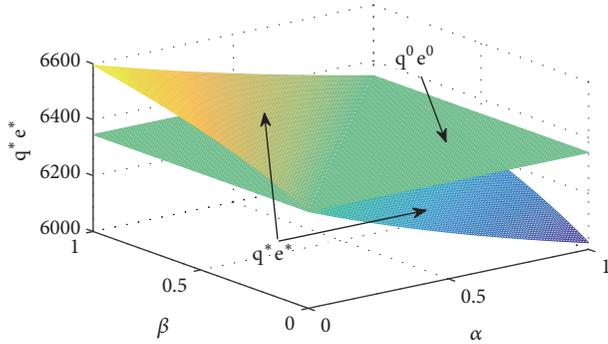


FIGURE 13: The impact of reference effect on total carbon emissions under the endogenous price.

$\beta$  (see Appendix E). When  $\alpha = 1$  and  $\beta = 0$ , we can get  $\varphi(1, 0) = (((v + s - c) + 2(c - v))^2 + 8(c - v)(u - v))^{1/2}$ . Due to  $(c - v)(u - v) > 0$ ,  $\varphi(1, 0) > (v + s - c) + 2(c - v)$ . Therefore, we can obtain  $((2u - c - v + s) + 2(c - v) + ((v + s - c) + 2(c - v)))/(2(u - v)) = ((u - v) + (s + c - v))/(u - v) > 1$ . That is,  $((1 + \beta)(2u - c - v + s) + (1 + \alpha)(c - v) + \varphi(\alpha, \beta))/(2(1 + \beta)(u - v))$  is always greater than 1. Because of  $F(z) \in (0, 1)$ , we remove the solution that is not in line with the requirement.

So, we can obtain

$$F(z^*) = \frac{(1 + \beta)(2u - c - v + s) + (1 + \alpha)(c - v) - \varphi(\alpha, \beta)}{2(u - v)}, \quad (\text{D.6})$$

## E. Proof of Propositions 6 and 7

In order to determine the impact of the manufacturer's aversion on the clearance loss and the shortage loss of his decisions, we introduce the Cramer's Rule. According to (22),

we can get the first order partial derivatives of all variables with respect to  $\alpha$ ; that is,

$$\begin{aligned} \frac{\partial p}{\partial \alpha} + (u - v) \frac{\partial F(z)}{\partial \alpha} &= 0, \\ (1 + \beta)(1 - F(z)) \frac{\partial p}{\partial \alpha} & \\ - ((1 + \beta)(p - c + s) + (1 + \alpha)(c - v)) \frac{\partial F(z)}{\partial \alpha} & \\ = (c - v)F(z). & \end{aligned} \quad (\text{E.1})$$

From (E.1), we can get

$$\frac{\partial p}{\partial \alpha} = \frac{\left| \begin{array}{c} 0 \\ (c-v)F(z) - ((1+\beta)(p-c+s) + (1+\alpha)(c-v)) \end{array} \right|}{\left| \begin{array}{c} 1 \\ (1+\beta)(1-F(z)) - ((1+\beta)(p-c+s) + (1+\alpha)(c-v)) \end{array} \right|} > 0, \quad (\text{E.2})$$

$$\frac{\partial F(z)}{\partial \alpha} = \frac{\left| \begin{array}{c} 1 \\ (1+\beta)(1-F(z)) \end{array} \right| \begin{array}{c} 0 \\ (c-v)F(z) \end{array}}{\left| \begin{array}{c} 1 \\ (1+\beta)(1-F(z)) - ((1+\beta)(p-c+s) + (1+\alpha)(c-v)) \end{array} \right|} < 0. \quad (\text{E.3})$$

Similarly, according to (22), we can get the first order partial derivatives of all variables with respect to  $\beta$ ; that is,

$$\begin{aligned} \frac{\partial p}{\partial \beta} + (u - v) \frac{\partial F(z)}{\partial \beta} &= 0, \\ (1 + \beta)(1 - F(z)) \frac{\partial p}{\partial \beta} & \\ - ((1 + \beta)(p - c + s) + (1 + \alpha)(c - v)) \frac{\partial F(z)}{\partial \beta} & \\ = -(p - c + s)(1 - F(z)). & \end{aligned} \quad (\text{E.4})$$

From (E.4), we can get

$$\frac{\partial p}{\partial \beta} = \frac{\left| \begin{array}{c} 0 \\ -(p-c+s)(1-F(z)) - ((1+\beta)(p-c+s)^{u-v} + (1+\alpha)(c-v)) \end{array} \right|}{\left| \begin{array}{c} 1 \\ (1+\beta)(1-F(z)) - ((1+\beta)(p-c+s)^{u-v} + (1+\alpha)(c-v)) \end{array} \right|} \quad (\text{E.5})$$

$< 0,$

$$\frac{\partial F(z)}{\partial \beta} = \frac{\left| \begin{array}{c} 1 \\ (1+\beta)(1-F(z)) - ((1+\beta)(p-c+s)^{u-v} + (1+\alpha)(c-v)) \end{array} \right|}{\left| \begin{array}{c} 0 \\ -(p-c+s)(1-F(z)) - ((1+\beta)(p-c+s)^{u-v} + (1+\alpha)(c-v)) \end{array} \right|} \quad (\text{E.6})$$

$> 0.$

So, when the price  $p$  is endogenous and consumers have strategic behavior, the manufacturer's optimal production quantity decreases with  $\alpha$  and increases with  $\beta$ . And the optimal unit price increases with  $\alpha$  and decreases with  $\beta$ .

Due to  $e^* = e_0 - k(p^* - c)/2h$ , the optimal unit carbon emission decreases with  $\alpha$  and increases with  $\beta$ .

## Data Availability

The 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 financially supported by the National Social Science Foundation of China (Grant No. 17BGL083).

## References

- [1] S. Du, L. Hu, and L. Wang, "Low-carbon supply policies and supply chain performance with carbon concerned demand," *Annals of Operations Research*, vol. 255, no. 1-2, pp. 569–590, 2017.
- [2] S. Arora and S. Gangopadhyay, "Toward a theoretical model of voluntary overcompliance," *Journal of Economic Behavior and Organization*, vol. 28, no. 3, pp. 289–309, 1995.
- [3] R. Echeverria, V. Hugo Moreira, C. Sepulveda, and C. Wittwer, "Willingness to pay for carbon footprint on foods," *British Food Journal*, vol. 116, no. 2, pp. 186–196, 2014.
- [4] X. Chen and X. Wang, "Effects of carbon emission reduction policies on transportation mode selections with stochastic demand," *Transportation Research Part E: Logistics and Transportation Review*, vol. 90, pp. 196–205, 2016.
- [5] Y. Chen, X. Su, and X. Zhao, "Modeling bounded rationality in capacity allocation games with the quantal response equilibrium," *Management Science*, vol. 58, no. 10, pp. 1952–1962, 2012.
- [6] Y. Chen and X. Zhao, "Decision bias in capacity allocation games with uncertain demand," *Production Engineering Research and Development*, vol. 24, no. 4, pp. 634–646, 2015.
- [7] X. Wu and J. A. Niederhoff, "Fairness in selling to the newsvendor," *Production Engineering Research and Development*, vol. 23, no. 11, pp. 2002–2022, 2014.
- [8] X. Su, "Bounded rationality in newsvendor models," *Manufacturing & Service Operations Management*, vol. 10, no. 4, pp. 566–589, 2008.
- [9] M. Becker-Peth, E. Katok, and U. W. Thonemann, "Designing buyback contracts for irrational but predictable newsvendors," *Management Science*, vol. 59, no. 8, pp. 1800–1816, 2013.
- [10] Y. Song and X. Zhao, "A newsvendor problem with boundedly rational strategic customers," *International Journal of Production Research*, vol. 55, no. 1, pp. 228–243, 2017.
- [11] M. E. Schweitzer and G. P. Cachon, "Decision bias in the newsvendor problem with a known demand distribution: experimental evidence," *Management Science*, vol. 46, no. 3, pp. 404–420, 2000.
- [12] X. D. Ding, H. Zhuang, X. L. Huang, G. F. Jiang, and T. Li, "Formation mechanism of newsvendor decision bias based on CARA utility function," *Control and Decision*, vol. 31, no. 2, pp. 287–296, 2016.
- [13] X. Long and J. Nasiry, "Prospect theory explains newsvendor behavior: the role of reference points," *Management Science*, vol. 61, no. 12, pp. 3009–3012, 2015.
- [14] J. Liu, C. Wu, and T. Su, "The reference effect in newsvendor model with strategic customers," *Management Decision*, vol. 55, no. 5, pp. 1006–1021, 2017.
- [15] Y. Ren and R. Croson, "Overconfidence in newsvendor orders: an experimental study," *Management Science*, vol. 59, no. 11, pp. 2502–2517, 2013.
- [16] J. Tong and D. Feiler, "A behavioral model of forecasting: Naive statistics on mental samples," *Management Science*, vol. 63, no. 11, pp. 3609–3627, 2017.
- [17] T. H. Cui, J. S. Raju, and Z. J. Zhang, "Fairness and channel coordination," *Management Science*, vol. 53, no. 8, pp. 1303–1314, 2007.
- [18] Y. Liu, C. Ding, C. Fan, and X. Chen, "Pricing decision under dual-channel structure considering fairness and free-riding behavior," *Discrete Dynamics in Nature and Society*, vol. 2014, Article ID 536576, 10 pages, 2014.
- [19] G. P. Cachon and R. Swinney, "The value of fast fashion: quick response, enhanced design, and strategic consumer behavior," *Management Science*, vol. 57, no. 4, pp. 778–795, 2011.
- [20] S. Wu, Q. Liu, and R. Q. Zhang, "The reference effects on a retailer's dynamic pricing and inventory strategies with strategic consumers," *Operations Research*, vol. 63, no. 6, pp. 1320–1335, 2015.
- [21] M. Wang, M. Ma, X. Yue, and S. Mukhopadhyay, "A capacitated firms pricing strategies for strategic consumers with different search costs," *Annals of Operations Research*, vol. 240, no. 2, pp. 731–760, 2016.
- [22] T. Li and M. Yu, "Coordinating a supply chain when facing strategic consumers," *Decision Sciences*, vol. 48, no. 2, pp. 336–355, 2017.
- [23] P. Du and Q. Chen, "Skimming or penetration: optimal pricing of new fashion products in the presence of strategic consumers," *Annals of Operations Research*, vol. 257, no. 1-2, pp. 275–295, 2017.
- [24] Y. Aviv and A. Pazgal, "Optimal pricing of seasonal products in the presence of forward-looking consumers," *Manufacturing & Service Operations Management*, vol. 10, no. 3, pp. 339–359, 2008.
- [25] W. Jiang and X. Chen, "Optimal strategies for manufacturer with strategic customer behavior under carbon emissions-sensitive random demand," *Industrial Management Data Systems*, vol. 116, no. 4, pp. 759–776, 2016.

- [26] B.-B. Cao, Z.-P. Fan, H. Li, and T.-H. You, "Joint inventory, pricing, and advertising decisions with surplus and stockout loss aversions," *Discrete Dynamics in Nature and Society*, vol. 2016, Article ID 1907680, 14 pages, 2016.
- [27] D. E. Bell, "Disappointment in decision making under uncertainty," *Operations Research*, vol. 33, no. 1, pp. 1–27, 1985.
- [28] F. Herweg, "The expectation-based loss-averse newsvendor," *Economics Letters*, vol. 120, no. 3, pp. 429–432, 2013.
- [29] B. Yalabik and R. J. Fairchild, "Customer, regulatory, and competitive pressure as drivers of environmental innovation," *International Journal of Production Economics*, vol. 131, no. 2, pp. 519–527, 2011.
- [30] B. Zhang, D. Wu, L. Liang, and D. L. Olson, "Supply chain loss averse newsboy model with capital constraint," *IEEE Transactions on Systems Man Cybernetics-Systems*, vol. 46, no. 5, pp. 646–658, 2016.
- [31] N. C. Petruzzi and M. Dada, "Pricing and the newsvendor problem: a review with extensions," *Operations Research*, vol. 47, no. 2, pp. 183–194, 1999.
- [32] P. Desai, O. Koenigsberg, and D. Purohit, "Strategic decentralization and channel coordination," *Quantitative Marketing and Economics*, vol. 2, no. 1, pp. 5–22, 2004.
- [33] X. Su and F. Zhang, "Strategic customer behavior, commitment, and supply chain performance," *Management Science*, vol. 54, no. 10, pp. 1759–1773, 2008.
- [34] D. Yang, E. Qi, and Y. Li, "Quick response and supply chain structure with strategic consumers," *Omega*, vol. 52, pp. 1–14, 2015.
- [35] C. X. Wang, "The loss-averse newsvendor game," *International Journal of Production Economics*, vol. 124, no. 2, pp. 448–452, 2010.
- [36] C. X. Wang and S. Webster, "The loss-averse newsvendor problem," *Omega*, vol. 37, no. 1, pp. 93–105, 2009.

## Research Article

# A Multiobjective Route Robust Optimization Model and Algorithm for Hazmat Transportation

Changxi Ma <sup>1</sup>, Wei Hao <sup>2,3</sup>, Ruichun He <sup>1</sup> and Bahman Moghimi<sup>4</sup>

<sup>1</sup>School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China

<sup>2</sup>Key Laboratory of Road Traffic Engineering of the Ministry of Education, Changsha University of Science and Technology, Changsha, Hunan 410114, China

<sup>3</sup>Hunan Provincial Key Laboratory of Intelligent Processing of Big Data on Transportation, Changsha University of Science and Technology, Changsha, Hunan 410114, China

<sup>4</sup>Department of Civil Engineering, City College of New York, NY 10031, USA

Correspondence should be addressed to Changxi Ma; machangxi@mail.lzjtu.cn

Received 12 June 2018; Accepted 13 September 2018; Published 9 October 2018

Guest Editor: Xiaobo Qu

Copyright © 2018 Changxi Ma 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.

Aiming at route optimization problem of hazardous materials transportation in uncertain environment, this paper presents a multiobjective robust optimization model by taking robust control parameters into consideration. The objective of the model is to minimize not only transportation risk but also transportation time, and a robust counterpart of the model is introduced through applying the Bertsimas-Sim robust optimization theory. Moreover, a fuzzy C-means clustering-particle swarm optimization (FCMC-PSO) algorithm is designed, and the FCMC algorithm is used to cluster the demand points. In addition the PSO algorithm with the adaptive archives grid is used to calculate the robust optimization route of hazmat transportation. Finally, the computational results show the multiobjective route robust optimization model with 3 centers and 20 demand points' sample studied and FCMC-PSO algorithm for hazmat transportation can obtain different robustness Pareto solution sets. As a result, this study will provide basic theory support for hazmat transportation safeguarding.

## 1. Introduction

Hazardous materials (hazmat) refer to products with flammable, poisonous, and corrosive properties that can cause casualties, damage to properties, and environmental pollution and require special protection in the process of transportation, loading, unloading, and storage. In recent years, the demand for hazmat has increased, its freight volume has increased year by year, and the potential transportation risk is also expanding. Practice has proved that the optimization of the transportation route of hazmat can effectively reduce the transportation risk, and it has significant influence to ensure the safety of people along route and protect the surrounding ecological environment.

Many scholars have studied the transportation route optimization problem for hazmat. Rhyné (1994) conducted a statistical analysis of the hazmat transport accident using the diffusion formula [1]. Gordon et al. (2000) used the

British transport accident to prove the different risk level in different places, which provided a realistic basis for the evacuation of the population and the balance of risks during the accident [2]. Power et al. (2000) established a risk-cost analysis model and conducted a systematic analysis to find out the transport plan [3]. Wu et al. (2002) studied the vehicle routing optimization problem of hazmat transportation with multidistribution center, and the clustering algorithm was used to solve the complex model [4]. Kara et al. (2003) presented a risk model based on the primitive roads to reduce the error of risk estimation, and then they proposed the minimum risk route selection algorithm [5]. Fabiano et al. (2005) used the experimental data to calculate the probability and consequence of the transport accident on the specific route of the experimental area and got the optimal transportation scheme [6]. Erku and Ingolfsson (2005) studied the hazmat transport route by considering the shortest distance, the smallest population exposure, and the smallest accident risk

[7]. Bubbico et al. (2006) analyzed the transportation risk of hazmat and obtained the safety route algorithm by using the experimental data from Italy [8]. Wei et al. (2006) analyzed the route selection model under the influence of time-varying conditions without considering the uncertain risk [9]. Verma (2009) presented a biobjective optimization model considering the cost and the risk and used the boundary algorithm to solve the model [10]. Wang et al. (2009) established a hazmat transportation path optimization model based on a geographic information system [11]. Jassbi et al. (2010) developed a multiobjective optimization framework for the hazmat transportation by minimizing the transport mileage, the number of affected residents, social risks, the probability of accidents, and so on [12]. Pradhananga et al. (2014) created a dual-objective optimization model with time windows for the hazmat transportation and designed a heuristic algorithm to solve this model [13]. Suh-Wen Chiou (2016) proposed a dual-objective and dual-level signal control policy for the hazmat transportation [14]. Assadipour et al. (2016) proposed a toll-based dual-level programming method for the hazmat transportation and designed hybrid speed constraints for multiobjective particle swarm optimization algorithm [15]. Pamucar et al. (2016) proposed a multiobjective route planning method for hazmat transportation and designed a solution algorithm combining neurofuzzy and artificial bee colony approach [16]. Mohammadi et al. (2017) studied the hazmat transportation under uncertain conditions using a mixed integer nonlinear programming model and the metaheuristic algorithm was utilized to solve this model [17]. Kheirkhah et al. (2017) established a bilevel optimization model for the hazmat transportation, two heuristic algorithms were designed to solve the dual-level optimization model, and some randomly generated problems were used to verify the applicability and validity of the model [18]. Bula et al. (2017) focused on the heterogeneous fleet vehicle routing problem and designed a variant of the variable neighborhood search algorithm to solve the problem [19]. Ma et al. (2013) studied the route optimization models and algorithms for hazardous materials transportation under different environments [20]. Ma et al. (2018) proposed a road screening algorithm and created a distribution route multiobjective robust optimization for hazardous materials [21]. Obviously, although the route optimization of hazmat transportation has made many achievements, the robustness of solution is rarely considered in this field.

Ben-Tal and Neimirovski (1998) proposed robust optimization theory based on ellipsoid uncertainty set [22]. Bertsimas and Sim (2003) further put forward adjustable robustness robust optimal theory [23]. Based on the adjustable robust optimization theory, this paper will propose a route robust optimization model for hazmat transportation with multiple distribution centers, and the author also designed a kind of improved PSO algorithm.

The rest of this paper is structured as follows: Section 2 introduces the hazmat transportation route problem and establishes a multiobjective route robust optimization model of hazmat transportation. Section 3 presents the FCMC-PSO algorithm. Section 4 is the case study. At the end of the paper, the conclusion is proposed.

## 2. Multiobjective Route Robust Optimization Model of Hazmat Transportation

**2.1. Problem Definition.** Hazmat transportation route robust optimization for multidistribution center is defined as follows: there are several hazmat distribution centers, and each distribution center owns enough hazmat transport vehicles; meanwhile, multiple need points exist which should be assigned to the relevant hazmat distribution center. Vehicles from distribution center will service the corresponding demand points. Each vehicle can service several customer demand points while each customer demand point only can be serviced by one vehicle. After completing the transport mission, the vehicles must return to the distribution center [24].

Uncertainty of hazmat transport refers to the uncertainty of the transportation time and transportation risk, which may be caused by the traffic accident, weather, and traffic density of the road. Compared to ordinary goods transport, hazmat transport is more complicated, and it needs more security demands. Therefore, it is needed to set the goal of minimizing the total hazmat transportation risk. In the process of hazmat transportation, transport time reduction is also necessary. As a consequence, this paper will target minimizing the total transportation time. In conclusion, the scientific transportation routes should be found to guarantee the hazmat transported safely and quickly.

### 2.2. Model

**2.2.1. Assumption.** There are a few assumptions in this study:

- (1) Multiple hazmat distribution centers are existent
- (2) The supply of hazmat distribution centers is adequate
- (3) Vehicle loading capacity is provided and the demand of each customer is specified
- (4) Multiple vehicles of the distribution center can service the customer
- (5) The transportation risk and transportation time are identified among the customer demand points, but they are uncertain number as interval number

**2.2.2. Symbol Definition.**  $S_0$ : set of the Hazmat distribution centers, where  $S_0 = \{i \mid i = 1, 2, \dots, m\}$  shows that the number distribution center is  $m$  and the sequence number of nodes set is  $1 \dots m$ .

$S_1$ : set of customer demand points, where  $S_1 = \{i \mid i = 1, \dots, n\}$  shows that the number of customer demand points is  $n$  and the sequence number of nodes set is  $1 \dots n$ .

$S$ : all nodes set in the transportation network, where  $S = S_0 \cup S_1$ .

$V_d$ : available transportation vehicle set in the hazmat distribution center, where  $V = \{k \mid k = 1, 2, \dots, K\}$ ,  $d \in S_0$ .

$E$ : road section set among nodes.

$q_i$ : demand of customer demand point  $i$ .

$L_k^d$ : maximum load of transport vehicle  $k$  from hazmat distribution center  $d$ .

$\tilde{r}_{ij}$ : variable transport risk from customer demand points  $i$  and  $j$ , where  $\tilde{r}_{ij} \in [r_{ij}, r_{ij} + \tilde{r}_{ij}]$  ( $\tilde{r}_{ij} \geq 0$ ).

$r_{ij}$ : transportation risk nominal value from customer demand points  $i$  and  $j$ .

$\tilde{r}_{ij}$ : deviation of the variable transport risk to its nominal value from customer demand points  $i$  and  $j$ , where  $\tilde{r}_{ij} \geq 0$ .

$t_{ij}$ : travel time nominal value from customer demand points  $i$  and  $j$ .

$\hat{t}_{ij}$ : deviation of variable travel time to its nominal value from customer demand points  $i$  and  $j$ , where  $\hat{t}_{ij} \geq 0$ .

$\tilde{t}_{ij}$ : variable transport risk from customer demand points  $i$  and  $j$ , where  $\tilde{t}_{ij} \in [t_{ij}, t_{ij} + \hat{t}_{ij}]$  ( $\tilde{t}_{ij} \geq 0$ ).

$J_i^r$ : the set of columns which all uncertain data  $\tilde{r}_{ij}$  belonging to the  $i$ th row of the variable risk matrix are in, here  $|J_i^r| \leq n$ .

$\Gamma_i^r$ : parameter  $\Gamma_i^r \in [0, |J_i^r|]$  to adjust robust risk of robust discrete optimization method and control the risk degree of conservatism, where decimal is permitted.

$\lfloor \Gamma_i^r \rfloor$ : maximum integer less than  $\Gamma_i^r$ .

$\Psi_i^r$ : the set of column subscripts  $j$  of uncertain data  $\tilde{r}_{ij}$  of line  $i$  in the variable risk matrix  $\tilde{r}_{ij}$ .

$J_i^t$ : the set of columns with all uncertain data  $\tilde{t}_{ij}$  belonging to the  $i$ th row of the variable time matrix, where  $|J_i^t| \leq n$ .

$\Gamma_i^t$ : parameter  $\Gamma_i^t \in [0, |J_i^t|]$  to adjust robust time of robust discrete optimization method and control the time degree of conservatism, where decimal is permitted.

$\lfloor \Gamma_i^t \rfloor$ : the maximum integer less than  $\Gamma_i^t$ .

$\Psi_i^t$ : the set of column subscript  $j$  of uncertain data  $\tilde{t}_{ij}$  of line  $i$  in variable time matrix  $\tilde{t}_{ij}$ .

### 2.2.3. Multiobjective Route Robust Optimization Model

$$\begin{aligned} \min \quad & Z_1 \\ & = \sum_{d \in S_0} \sum_{k \in V_d} \sum_{i \in S} \sum_{j \in S_1} r_{ij} x_{ijk}^d \\ & + \max_{\{\Psi_i^r \cup \{m^r\} | \Psi_i^r \subseteq J_i^r, |\Psi_i^r| = \lfloor \Gamma_i^r \rfloor, m^r \in J_i^r \setminus \Psi_i^r\}} \left\{ \sum_{d \in S_0} \sum_{k \in V_d} \sum_{i \in S} \sum_{j \in \Psi_i^r} \tilde{r}_{ij} x_{ijk}^d + \sum_{d \in S_0} \sum_{k \in V_d} \sum_{i \in S} \sum_{m^r \in J_i^r \setminus \Psi_i^r} (\Gamma_i - \lfloor \Gamma_i \rfloor) \tilde{r}_{im^r} x_{im^r k}^d \right\} \\ & - \sum_{d \in S_0} \sum_{k \in V_d} \sum_{i \in S} r_{i0} x_{i0k}^d \end{aligned} \quad (1)$$

$$\begin{aligned} \min \quad & Z_2 \\ & = \sum_{d \in S_0} \sum_{k \in V_d} \sum_{i \in S} \sum_{j \in S} t_{ij} x_{ijk}^d \\ & + \max_{\{\Psi_i^t \cup \{m^t\} | \Psi_i^t \subseteq J_i^t, |\Psi_i^t| = \lfloor \Gamma_i^t \rfloor, m^t \in J_i^t \setminus \Psi_i^t\}} \left\{ \sum_{d \in S_0} \sum_{k \in V_d} \sum_{i \in S} \sum_{j \in \Psi_i^t} \tilde{t}_{ij} x_{ijk}^d + \sum_{d \in S_0} \sum_{k \in V_d} \sum_{i \in S} \sum_{m^t \in J_i^t \setminus \Psi_i^t} (\Gamma_i - \lfloor \Gamma_i \rfloor) \tilde{t}_{im^t} x_{im^t k}^d \right\} \end{aligned} \quad (2)$$

$$\text{s.t.} \quad \sum_{i \in S} \sum_{j \in S_1} \sum_{d \in S_0} x_{ijk}^d q_j \leq L_k^d, \quad \forall k \in V_d \quad (3)$$

$$\sum_{j \in S_1} \sum_{k \in V_d} x_{ijk}^d \leq |V_d|, \quad \forall i \in S_0, \quad \forall d \in S_0 \quad (4)$$

$$\max_{(i,j) \in E} \{r_{ij} x_{ijk}^d\} \leq r_{\max}, \quad \forall i \in S, \quad \forall j \in S, \quad \forall d \in S_0, \quad \forall k \in V_d \quad (5)$$

$$\max_{(i,j) \in E} \left\{ \sum_{i \in S_0 \cup S_1} \sum_{j \in S_0 \cup S_1} r_{ij} x_{ijk}^d - \sum_{i \in S_1} \sum_{j \in S_0} r_{ij} x_{ijk}^d \right\} \leq R_{\max}, \quad \forall i \in S, \quad \forall j \in S, \quad \forall d \in S_0, \quad \forall k \in V_d \quad (6)$$

$$\sum_{j \in S_1} x_{ijk}^d - \sum_{j \in S_1} x_{jik}^d = 0, \quad \forall i \in S_0, \quad \forall d \in S_0, \quad \forall k \in V_d \quad (7)$$

$$\sum_{i \in S} \sum_{d \in S_0} \sum_{k \in V_d} x_{ijk}^d = 1, \quad \forall j \in S_1 \quad (8)$$

$$\sum_{j \in S} \sum_{d \in S_0} \sum_{k \in V_d} x_{ijk}^d = 1, \quad \forall i \in S_1 \quad (9)$$

$$\sum_{j \in S_0} x_{ijk}^d = \sum_{j \in S_0} x_{jik}^d = 0, \quad \forall i \in S_0, \quad \forall d \in S_0, \quad \forall k \in V_d \quad (10)$$

$$x_{ijk}^d = \{1, 0\}, \quad \forall i \in S, \quad \forall j \in S, \quad \forall d \in S_0, \quad \forall k \in V_d \quad (11)$$

where the objective function (1) minimizes the total transportation risk of hazmat [25], and the objective function (2) minimizes the total transportation time of hazmat. Constraint (3) as the load constraint that means any vehicle of any distribution center should satisfy the corresponding load constraint, namely, cannot overload. Constraint (4) means that vehicle number of the distribution center is limited, and vehicles arranged to transport the hazmat should not exceed the distribution center owning. Constraint (5) expresses that the transportation risk of each section must be less than or equal to threshold  $r_{max}$  set by decision makers. Constraint (6) expresses that the transportation risk of each route must be less than or equal to threshold  $R_{max}$  set by decision makers. Constraint (7) indicates every hazmat vehicle departing from distribution center should return back to the original distribution center after finishing the transportation task. Constraint (8) and Constraint (9) guarantee that each demand point is served once and by one vehicle form a distribution center. Constraint (10) indicates hazmat transport vehicles cannot depart from one distribution center but back to another one. Constraint (11) defines the 0-1 integer variable  $x_{ijk}^d$ , if the route of Hazmat transport vehicle  $k$  from distribution center  $d$  containing the segment from node  $i$  to node  $j$ ,  $x_{ijk}^d = 1$ ; else  $x_{ijk}^d = 0$ .

**2.2.4. Robust Counterpart Model.** Each objective function of the above multiobjective robust model corresponds to parameter  $\Gamma$ . The purpose is to control the degree of conservatism of the solution. Objective functions (1) and (2) of the robust optimization model contain "max" extreme value problem. Set feasible solution set  $X_{vrp}$  to satisfy all constraints, and robust discrete optimization criterion is used to transform the multiobjective route robust optimization model, and a new robust counterpart of the model is as follows [26, 27].

Objective function is

$$R(r) = \widehat{\Gamma} \widehat{r}_1 + \min \left( \sum_{m=1}^{n^2} r_m x_m + \sum_{m=1}^l (\widehat{r}_m - \widehat{r}_1) x_m \right) \quad (12)$$

$$T(t) = \widehat{\Gamma} \widehat{t}_1 + \min \left( \sum_{m=1}^{n^2} t_m x_m + \sum_{m=1}^l (\widehat{t}_m - \widehat{t}_1) x_m \right) \quad (13)$$

Constraint condition is

$$x \in X_{vrp} \quad (14)$$

Then, the optimal objective function value can be obtained as  $R^* = \min_{l=1, \dots, n^2+1} R(l)$  and  $T^* = \min_{l=1, \dots, n^2+1} T(l)$ .

### 3. Algorithm

In this section, we propose FCMC-PSO algorithm to solve the multiobjective route optimization problem of hazmat transportation in uncertain environment. The demand points is clustered by the fuzzy C means algorithm, and the transportation route for each demand points is determined based on the adaptive archives grid multiobjective particle swarm optimization [28–30].

**3.1. Fuzzy C-Means Clustering.** Suppose that  $n$  data samples are  $X = \{X_1, X_2, \dots, X_n\}$ ,  $C(2 \leq C \leq n)$  is the number of types into which the data samples are to be divided,  $\{A_1, A_2, \dots, A_c\}$  indicating the corresponding  $C$  categories,  $U$  is its similar classification matrix, the clustering centers of all categories are  $\{V_1, V_2, \dots, V_c\}$ , and  $\mu_k(X_i)$  is the membership degree of sample  $X_i$  to category  $A_k$  (abbreviated as  $\mu_{ik}$ ). Then the objective function can be expressed as follows:

$$J_b(U, V) = \sum_{i=1}^n \sum_{k=1}^c (\mu_{ik})^b (d_{ik})^2 \quad (15)$$

where  $d_{ik} = d(X_i - V_k) = \sqrt{\sum_{j=1}^m (X_{ij} - V_{kj})^2}$ , it is the synthetic weighted value of the transport risk and time after nondimensionalization between the  $i$ -th sample  $X_i$  and the  $k$ -th category center point;  $m$  is the characteristics number of the sample;  $b$  is the weighting parameter, and the value range is  $1 \leq b \leq \infty$ . The fuzzy C-means clustering method is to find an optimal classification, so that the classification can produce the smallest function value  $J_b$ . It requires a sample for the sum of the membership degree of each cluster is 1, which is satisfied:

$$\sum_{j=1}^c \mu_j(X_i) = 1, \quad i = 1, 2, \dots, n \quad (16)$$

Formulas (17) and (18) are used to calculate separately the membership degree  $\mu_{ik}$  of the sample  $X_i$  for the category  $A_k$  and  $C$  clustering centers  $\{V_j\}$ :

$$\mu_{ik} = \frac{1}{\sum_{j=1}^c (d_{ik}/d_{jk})^{2/(b-1)}} \quad (17)$$

Let  $I_k = \{i \mid 2 \leq C < n; d_{ik} = 0\}$ , for all the  $i$  categories,  $i \in I_k, \mu_{ik} = 0$ .

$$V_{ij} = \frac{\sum_{k=1}^n (\mu_{ik})^b X_{kj}}{\sum_{k=1}^n (\mu_{ik})^b} \quad (18)$$

Using formulas (17) and (18) to repeatedly modify the cluster center, data membership, and classification, when the algorithm is convergent in theory, we can get the membership degree of the cluster center and each sample for each pattern class; thus the division of the fuzzy clustering is completed.

**3.2. Multiobject PSO Algorithm.** Particle swarm algorithm is derived from the study of the predatory behavior of birds. It is used to solve the problem of path optimization. Each particle in the algorithm represents a potential solution, and the fitness value for each particle is determined by the fitness function, and the value of fitness determines the pros and cons of the particle. The particle moves in the  $N$ -dimensional solution space and updates the individual position by the individual extremum and the group extremum. In the algorithm, the velocity, position, and fitness value are used to represent the characteristics of the particle. The velocity of the particle determines the direction and distance of the particle

Individuals [9 2 1 4 3 6 7 5 8] → New 1 [9 2 2 4 1 5 9 5 8]
Extremum [8 3 2 4 1 5 9 7 6] → New 2 [8 3 1 4 3 6 7 7 6]

FIGURE 1: Crossover operation.

movement, and the velocity is dynamically adjusted with the moving experience of its own and other particles. Once the position of the particle is updated, the fitness value will be calculated, and the individual extremum and the population extremum are updated by the fitness values of the new particles, the individual extremum, and the population extremum. Multiobjective particle swarm optimization algorithm is a method based on particle swarm optimization algorithm to solve multiobjective problem. At the same time, the best location of multiple populations exists in the population, and the optimal positions of multiple particles themselves are also found in the iterative process. Therefore, *gbest* and *pbest* also need to adopt certain strategies to choose. Aiming at the robust optimization model of hazmat transportation, the key elements of multiobjective particle swarm optimization are as follows.

(1) *Individual Coding*. In this paper, the method of particle encoding adopts integer encoding, and each particle represents the experienced demand point. For example, when the number of required points is 9, the individual coding is [9 2 1 4 3 6 7 5 8], indicating that the requirement point traversal starting from the distribution center, followed by 9 2 1 4 3 6 7 5 8, and ultimately return to the distribution center.

(2) *Fitness Value*. In the hybrid particle swarm algorithm, the fitness value is the criterion of judging the quality of the particle. And the fitness function is to facilitate the search and improve the performance of the algorithm. In the paper, the fitness value of the particle is expressed by the objective function of the built model.

(3) *Crossover Operation*. Crossover operation is the process of replacing the partial structure of the parent individual and reorganizing the new individual. The design of the crossover operation is related to the representation of the coding, the cross-operation design based on the coarranged coding method of the demand point and the distribution center [31, 32]. The method of integer crossing is adopted. Set the two individuals of the parent as [9 2 1 4 3 6 7 5 8] and [8 3 2 4 1 5 9 7 6], Firstly, two crossover positions are selected, and then the individual is crossed; the operation process can be seen in Figure 1.

The new individuals need to be adjusted if there is a duplicate position, and the adjustment method is to replace the repeated demand points by using the absence of demand points in individuals. For the new individual 1, there are mappings about 2 to 3, 9 to 6, and 5 to 7. The specific adjustments process can be seen in Figure 2.

The strategy of retaining outstanding individuals is used for the owned new individuals, and the particles are updated only when the new particle fitness is better than the old particles.

New individuals [9 2 2 4 1 5 9 5 8] → [9 2 3 4 1 5 6 7 8]
New individuals [8 3 1 4 3 6 7 7 6] → [8 3 1 4 2 6 7 5 9]

FIGURE 2: Adjustment operation.

Individuals [9 2 3 4 1 5 6 7 8] → [9 6 3 4 1 5 2 7 8]
Individuals [8 3 1 4 2 6 7 5 9] → [8 3 1 4 6 2 7 5 9]

FIGURE 3: Mutation operation.

(4) *Mutation Operation*. The mutation of the particles is to make some changes in some genes of the particle; the mutation can increase the ability of searching particles and increase the diversity of the populations, to avoid falling into the local optimal situation.

The variation is also related to the way the particle is encoded. Based on the lease point and the dispatch center, there are many methods about the coarranged coding and the variation. In this paper, the variation method adopts the individual internal exchange method. For example, for an individual [9 2 3 4 1 5 6 7 8], at first, the mutated positions *pos1* and *pos2* are selected randomly; and then the positions of two variants are swapped. Assuming that the selected mutation positions are 2 and 6, the mutation operation process can be seen in Figure 3.

The strategy of retaining outstanding individuals is used for the owned new individuals, and the particles are updated only when the new particle fitness is better than the old particles [33].

(5) *Multiobject PSO Algorithm Based on Adaptive Archives Grid*. Multiobjective PSO based on adaptive archives grid is a particle swarm optimization algorithm proposed by Coello and Lechuga to solve multiobjective problem [34]. Its basic idea is to divide the target space into several hypercubes and to judge the number of noninferiority contained in each hypercube to maintain the external files. In each iteration, if the file does not exceed the given size, then a new nondominated solution will be added to the file. If the file has been filled, the file is maintained according to the density of noninferior solution contained in the hypercube, the noninferior solution is removed from the high-density hypercube, and the noninferior solution with low density is added to ensure the diversity of the population. The algorithm steps are as follows.

*Step 1*. Create and initialize a group so that the *ex\_archives* external file is empty.

*Step 2*. Evaluate all particles and add the noninferior solution to the external file.

*Step 3*. Maintain external files according to the adaptive grid method.

*Step 4*. Select *gbest* and *pbest* for each particle.

*Step 5*. Updating the velocity and position of the particles according to the speed formula and the position formula of the particle swarm.

TABLE 1: Customer demands.

Demand point	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Demand (ton)	2	1.5	4.5	3	1.5	4	2.5	3	3	4.5	2	1.5	2.5	3.5	2	2	2.5	3.5	2.5	3

TABLE 2: The nominal transportation risk value of hazmat.

$r_{ij}$	a	b	c	1	2	3	.....	18	19	20
a	0			39	71	59	.....	38	80	52
b		0		30	30	49	.....	74	67	78
c			0	77	57	37	.....	49	34	64
1	39	30	77	0	32	30	.....	35	60	49
2	71	30	57	32	0	67	.....	33	65	55
3	59	49	37	30	67	0	.....	39	39	71
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
18	38	74	49	35	33	39	.....	0	55	34
19	80	67	34	60	65	39	.....	55	0	67
20	52	78	64	49	55	71	.....	34	67	0

TABLE 3: The nominal transportation time value of hazamt.

$r_{ij}$	a	b	c	1	2	3	.....	18	19	20
a	0			92	103	43	.....	107	97	44
b		0		37	36	88	.....	95	57	115
c			0	41	35	101	.....	40	74	64
1	92	37	41	0	75	102	.....	64	101	80
2	103	36	35	75	0	62	.....	108	84	95
3	43	88	101	102	62	0	.....	103	77	85
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
18	107	95	40	64	108	103	.....	0	52	43
19	97	57	74	101	84	77	.....	52	0	86
20	44	115	64	80	95	85	.....	43	86	0

Step 6. Make sure the particles exist in the search space.

Step 7. If the termination condition is satisfied, the output result algorithm is terminated; if it is not satisfied, Step 2 is executed and the execution is continued.

### 4. Case Study

There are 3 distribution centers and 20 demand points, the maximum load for each transport vehicle is 8 ton, and each distribution center has adequate hazmat. The distribution centers are marked as a, b, and c, and the hazmat demand points are marked as 1, 2,..., 20. The demand amount of each demand point is shown in Table 1, and transportation risk and time from each distribution center to demand points and between the demand points are, respectively, shown in Tables 2 and 3. The risk and time nominal values are given in Tables 2 and 3. The transportation risk deviation  $\hat{r}_m(0 \leq \hat{r}_m < 0.5r_m)$  and transportation time deviation  $\hat{t}_m(0 \leq \hat{t}_m < 0.5r_m)$  are provided.

We use FCMC algorithm to calculate the demand points clustering results, and the results can be seen in Table 4.

TABLE 4: Demand points clustering result.

Distribution center	Demand points
a	3,9,14,15,16,20
b	1,2,6,8,10,11,12
c	4,5,7,13,17,18,19

Based on the cluster results, we use the multiobjective PSO to solve the robust optimization problem for each distribution center. The parameters of the algorithm are set as follows: population size is 100, maximum evolution generation is 1000, inertia weight is 0.6, accelerated factor is 1.7, crossover rate is 0.95, and mutation rate is 0.09. The Pareto solution set with different robust control parameters can be obtained by calculation, which are showed in Tables 5–13 and Figures 4–6.

In Table 5, the Pareto optimal solutions sets are obtained by the program running on the condition that uncertain transportation risk and uncertain time are taken the nominal values. It is known from Table 5 that when  $\Gamma = 0$ , the program finds 6 Pareto solutions; in the encoding sequence, the first

TABLE 5: Pareto solution set of robust control parameters  $\Gamma=0$  for distribution center a.

Encoding	Decoding	Total risk	Total time
9-20-14-3-15-16	a-9-20-a-14-3-a-15-16	237	536
9-16-14-3-15-20	a-9-16-a-14-3-a-15-20	261	448
15-20-9-14-3-16	a-15-20-9-a-14-3-a-16	247	520
14-3-16-9-20-15	a-14-3-a-16-9-20-a-15	253	469
15-14-16-9-3-20	a-15-14-16-a-9-3-a-20	291	404
14-3-15-9-16-20	a-14-3-a-15-9-16-a-20	285	425

TABLE 6: Pareto solution set of robust control parameters  $\Gamma=10$  for distribution center a.

Encoding	Decoding	Total risk	Total time
3-14-20-9-16-15	a-3-14-a-20-9-16-a-15	317.817	637.81
9-20-16-14-3-15	a-9-20-16-a-14-3-a-15	319.893	621.749
9-20-14-16-3-15	a-9-20-a-14-16-a-3-15	377.505	520.018
20-9-16-14-3-15	a-20-9-16-a-14-3-a-15	337.903	532.141
9-3-15-14-16-20	a-9-3-a-15-14-16-a-20	385.542	463.803
9-3-20-15-14-16	a-9-3-a-20-15-a-14-16	354.151	527.306
9-20-14-3-15-16	a-9-20-a-14-3-a-15-16	305.49	657.577
20-15-14-3-9-16	a-20-15-a-14-3-a-9-16	359.107	525.349

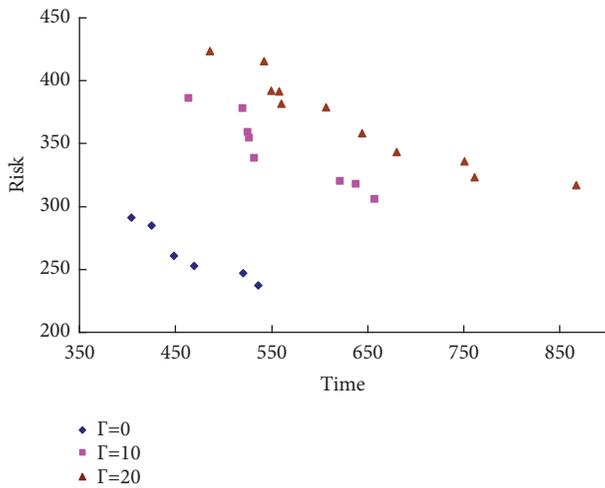


FIGURE 4: Pareto optimal solution distribution of robust control parameters  $\Gamma=0$ ,  $\Gamma=10$ , and  $\Gamma=20$  for distribution center a.

number a represents the distribution center. While, in the decoding sequence, every alphabet represents one vehicle, the figures behind alphabet show the customer demand points and the corresponding service order. Such as in Table 5, the first decoding sequence indicates that 3 vehicles are needed and each vehicle corresponds to a subroute, they are, respectively,  $a \rightarrow 9 \rightarrow 20 \rightarrow a$ ,  $a \rightarrow 14 \rightarrow 3 \rightarrow a$ , and  $a \rightarrow 15 \rightarrow 16 \rightarrow a$ , and it is clearly known that all transport vehicles from distribution center a, after serving the allocated customer demand points, eventually return to distribution center a. When the robustness control parameters  $\Gamma = 10$  and  $\Gamma = 20$ , Pareto optimal solutions sets obtained by the program running are given in Tables 6 and 7, respectively. Tables 8, 9, and 10 present the optimal Pareto solutions sets for distribution center b when the robustness control parameters are taken 0, 10, and 20. Tables 11, 12, and 13 present the

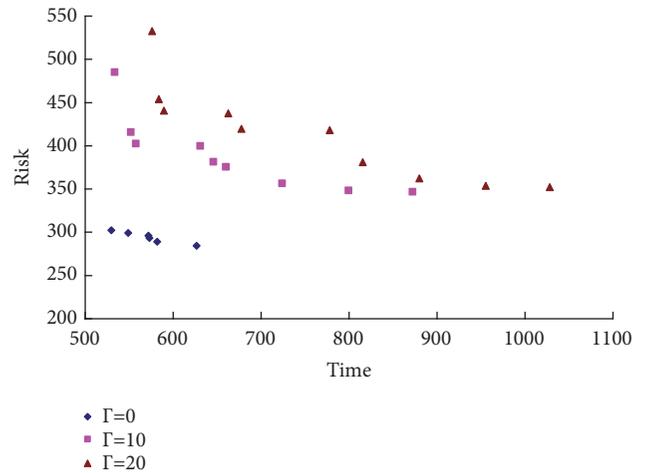


FIGURE 5: Pareto optimal solution distribution of robust control parameters  $\Gamma=0$ ,  $\Gamma=10$ , and  $\Gamma=20$  for distribution center b.

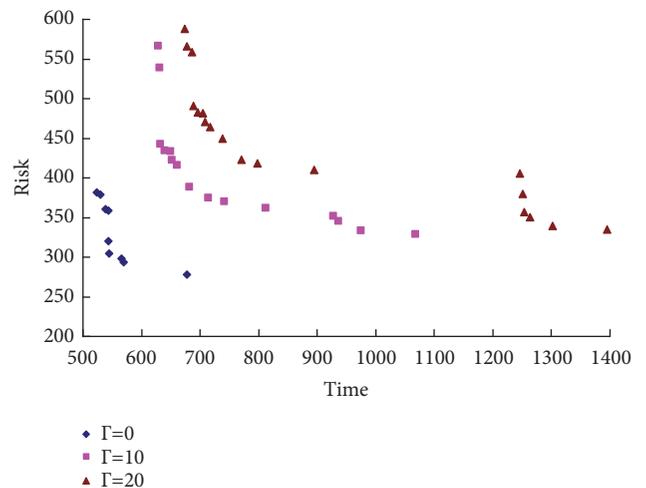


FIGURE 6: Pareto optimal solution distribution of robust control parameters  $\Gamma=0$ ,  $\Gamma=10$ , and  $\Gamma=20$  for distribution center c.

TABLE 7: Pareto solution set of robust control parameters  $\Gamma=20$  for distribution center a.

Encoding	Decoding	Total risk	Total time
3-14-20-9-16-15	a-3-14-a-20-9-16-a-15	323.453	761.423
9-3-15-20-14-16	a-9-3-a-15-20-a-14-16	381.844	560.087
9-20-16-14-3-15	a-9-20-16-a-14-3-a-15	357.98	644
9-16-20-14-3-15	a-9-16-20-a-14-3-a-15	391.378	558.043
9-20-14-16-3-15	a-9-20-a-14-16-a-3-15	415.592	542.269
9-3-15-14-16-20	a-9-3-a-15-14-16-a-20	423.629	486.054
9-20-15-14-3-16	a-9-20-15-a-14-3-a-16	378.69	606.48
3-14-9-20-16-15	a-3-14-a-9-20-16-a-15	316.995	866.785
9-3-20-15-14-16	a-9-3-a-20-15-a-14-16	392.239	549.557
3-14-16-9-20-15	a-3-14-a-16-9-20-a-15	336.135	750.943
9-20-14-3-15-16	a-9-20-a-14-3-a-15-16	343.577	679.828

TABLE 8: Pareto solution set of robust control parameters  $\Gamma=0$  for distribution center b.

Encoding	Decoding	Total risk	Total time
12-10-11-1-2-6-8	b-12-10-11-b-1-2-6-b-8	289	582
11-10-1-12-6-2-8	b-11-10-b-1-12-6-b-2-8	303	471
12-6-10-11-1-2-8	b-12-6-b-10-11-b-1-2-8	302	530
12-10-11-1-6-8-2	b-12-10-11-b-1-6-b-8-2	293	573
1-12-10-2-6-8-11	b-1-12-10-b-2-6-b-8-11	296	572
8-6-12-10-11-1-2	b-8-6-b-12-10-11-b-1-2	299	549
2-1-6-12-10-11-8	b-2-1-6-b-12-10-11-b-8	284	627
8-6-1-12-2-10-11	b-8-6-b-1-12-2-b-10-11	363	458

TABLE 9: Pareto solution set of robust control parameters  $\Gamma=10$  for distribution center b.

Encoding	Decoding	Total risk	Total time
2-6-8-12-1-11-10	b-2-6-b-8-12-1-b-11-10	415.472	552.167
1-6-2-11-10-12-8	b-1-6-2-b-11-10-12-b-8	348.014	799.695
2-8-6-1-12-11-10	b-2-8-b-6-1-12-b-11-10	399.364	631.219
1-12-6-11-10-8-2	b-1-12-6-b-11-10-b-8-2	375.264	660.062
2-12-1-6-8-11-10	b-2-12-1-b-6-8-b-11-10	485.04	533.589
2-8-6-12-1-11-10	b-2-8-b-6-12-1-b-11-10	402.203	557.881
1-6-2-12-10-8-11	b-1-6-2-b-12-10-b-8-11	346.521	872.417
1-6-12-11-10-8-2	b-1-6-12-b-11-10-b-8-2	356.381	724.274
2-6-1-11-10-12-8	b-2-6-1-b-11-10-12-b-8	381.268	645.953

TABLE 10: Pareto solution set of robust control parameters  $\Gamma=20$  for distribution center b.

Encoding	Decoding	Total risk	Total time
2-6-1-12-10-8-11	b-2-6-1-b-12-10-b-8-11	417.863	778.297
1-6-2-12-10-8-11	b-1-6-2-b-12-10-b-8-11	352.157	1028.237
2-6-1-11-10-12-8	b-2-6-1-b-11-10-12-b-8	419.356	677.732
1-6-2-11-10-12-8	b-1-6-2-b-11-10-12-b-8	353.65	955.516
2-8-6-1-12-11-10	b-2-8-b-6-1-12-b-11-10	437.452	662.997
2-8-6-12-1-11-10	b-2-8-b-6-12-1-b-11-10	440.291	589.659
2-6-8-12-1-11-10	b-2-6-b-8-12-1-b-11-10	453.559	583.946
1-6-12-11-10-8-2	b-1-6-12-b-11-10-b-8-2	362.017	880.094
1-12-6-11-10-8-2	b-1-12-6-b-11-10-b-8-2	380.9	815.883
2-12-1-6-8-10-11	b-2-12-1-b-6-8-b-10-11	532.501	576.115

TABLE 11: Pareto solution set of robust control parameters  $\Gamma=0$  for distribution center c.

Encoding	Decoding	Total risk	Total time
17-4-19-7-13-18-5	c-17-4-19-c-7-13-c-18-5	278	678
7-5-18-17-4-19-13	c-7-5-18-c-17-4-19-c-13	294	569
19-4-13-7-5-18-17	c-19-4-13-c-7-5-18-c-17	298	566
7-4-13-17-19-5-18	c-7-4-13-c-17-19-5-c-18	382	524
7-5-18-17-19-4-13	c-7-5-18-c-17-19-4-c-13	305	545
17-4-13-19-5-7-18	c-17-4-13-c-19-5-7-c-18	361	539
17-18-19-4-13-7-5	c-17-18-c-19-4-13-c-7-5	320	544
18-5-19-17-4-13-7	c-18-5-19-c-17-4-13-c-7	359	543
19-5-7-17-18-13-4	c-19-5-7-c-17-18-c-13-4	379	530

TABLE 12: Pareto solution set of robust control parameters  $\Gamma=10$  for distribution center c.

Encoding	Decoding	Total risk	Total time
18-17-13-4-19-7-5	c-18-17-c-13-4-19-c-7-5	433.965	640.855
18-5-7-17-4-19-13	c-18-5-7-c-17-4-19-c-13	374.262	714.495
4-19-13-7-5-18-17	c-4-19-13-c-7-5-18-c-17	328.774	1068.644
18-17-19-4-13-7-5	c-18-17-c-19-4-13-c-7-5	432.616	649.645
18-17-13-4-19-5-7	c-18-17-c-13-4-19-c-5-7	442.263	632.41
5-18-7-17-4-19-13	c-5-18-7-c-17-4-19-c-13	361.711	813.376
18-17-7-5-19-13-4	c-18-17-c-7-5-19-c-13-4	565.935	629.522
18-5-13-17-4-19-7	c-18-5-13-c-17-4-19-c-7	369.796	741.803
13-4-18-17-19-5-7	c-13-4-c-18-17-c-19-5-7	538.139	632.006
18-17-13-19-4-7-5	c-18-17-c-13-19-4-c-7-5	415.204	661.122
4-19-13-17-18-5-7	c-4-19-13-c-17-18-5-c-7	345.113	937.323
18-17-13-19-4-5-7	c-18-17-c-13-19-4-c-5-7	421.737	652.677
4-19-13-17-5-18-7	c-4-19-13-c-17-5-18-c-7	333.451	975.738
4-19-13-17-18-7-5	c-4-19-13-c-17-18-c-7-5	351.435	928.3
18-5-7-19-4-17-13	c-18-5-7-c-19-4-17-c-13	388.454	682.357

TABLE 13: Pareto solution set of robust control parameters  $\Gamma=20$  for distribution center c.

Encoding	Decoding	Total risk	Total time
18-5-13-17-4-19-7	c-18-5-13-c-17-4-19-c-7	418.386	797.49
13-4-19-17-18-7-5	c-13-4-19-c-17-18-c-7-5	565.673	677.462
13-4-19-17-18-5-7	c-13-4-19-c-17-18-5-c-7	558.373	686.485
13-4-19-18-17-7-5	c-13-4-19-c-18-17-c-7-5	588.331	674.181
18-17-13-19-4-7-5	c-18-17-c-13-19-4-c-7-5	463.794	716.809
4-19-13-17-5-18-7	c-4-19-13-c-17-5-18-c-7	339.087	1301.559
18-17-19-4-13-7-5	c-18-17-c-19-4-13-c-7-5	481.205	705.332
4-19-13-17-18-5-7	c-4-19-13-c-17-18-5-c-7	350.749	1263.143
18-17-13-4-19-7-5	c-18-17-c-13-4-19-c-7-5	482.554	696.543
18-17-13-19-4-5-7	c-18-17-c-13-19-4-c-5-7	470.326	708.364
4-19-13-18-17-7-5	c-4-19-13-c-18-17-c-7-5	379.729	1250.84
4-19-13-17-18-7-5	c-4-19-13-c-17-18-c-7-5	357.071	1254.121
4-19-13-7-5-18-17	c-4-19-13-c-7-5-18-c-17	334.41	1394.464
4-17-19-18-5-7-13	c-4-17-19-c-18-5-7-c-13	405.101	1245.591
18-5-7-17-19-4-13	c-18-5-7-c-17-19-4-c-13	449.149	738.192
18-17-13-4-19-5-7	c-18-17-c-13-4-19-c-5-7	490.853	688.097
5-18-7-17-4-19-13	c-5-18-7-c-17-4-19-c-13	409.798	895.155
18-5-7-17-4-19-13	c-18-5-7-c-17-4-19-c-13	422.852	770.182

TABLE 14: Pareto optimal solutions for distribution center a when  $\Gamma = 0$ ,  $\Gamma = 10$  and  $\Gamma = 20$ .

$\Gamma$	Distribution center a	
	Risk optimal route	Time optimal route
0	a-9-20-a	a-15-14-16-a
	a-14-3-a	a-9-3-a
	a-15-16-a	a-20-a
10	a-9-20-a	a-15-14-16-a
	a-14-3-a	a-9-3-a
	a-15-16-a	a-20-a
20	a-9-20-16-a	a-15-14-16-a
	a-14-3-a	a-9-3-a
	a-15-a	a-20-a

TABLE 15: Pareto optimal solutions for distribution center b when  $\Gamma = 0$ ,  $\Gamma = 10$  and  $\Gamma = 20$ .

$\Gamma$	Distribution center b	
	Risk optimal route	Time optimal route
0	b-2-1-6-b	b-8-6-b
	b-12-10-11-b	b-1-12-2-b
	b-8-b	b-10-11-b
10	b-1-6-2-b	b-8-6-b
	b-12-10-b	b-2-12-1-b
	b-8-11-b	b-10-11-b
20	b-1-6-2-b	b-8-6-b
	b-12-10-b	b-2-12-1-b
	b-8-11-b	b-10-11-b

TABLE 16: Pareto optimal solutions for distribution center c when  $\Gamma = 0$ ,  $\Gamma = 10$  and  $\Gamma = 20$ .

$\Gamma$	Distribution center c	
	Risk optimal route	Time optimal route
0	c-17-4-19-c	c-7-4-13-c
	c-7-13-c	c-17-19-5-c
	c-18-5-c	c-18-c
10	c-4-19-13-c	c-13-4-c
	c-7-5-18-c	c-7-5-19-c
	c-17-c	c-18-17-c
20	c-4-19-13-c	c-13-4-19-c
	c-7-5-18-c	c-7-5-c
	c-17-c	c-18-17-c

optimal Pareto solutions sets for distribution center c when the robustness control parameters are taken 0, 10, and 20. The above tables are analyzed similar to Table 5. Figures 4, 5, and 6, respectively, show the Pareto optimal solutions for distribution centers a, b, and c when the robustness control parameters  $\Gamma = 0$ ,  $\Gamma = 10$ , and  $\Gamma = 20$ .

When the robustness control parameters  $\Gamma = 0$ ,  $\Gamma = 10$ , and  $\Gamma = 20$ , each objective optimal distribution route for all vehicles from distribution centers a, b, c is shown in Tables 14, 15, and 16. Aiming at distribution center a, when the robustness control parameters  $\Gamma = 0$  and  $\Gamma = 10$ , the total risk

TABLE 17: Performance comparison between FCMC-PSO algorithm and SPEA.

Index	FCMC-PSO			SPEA		
	0	20	40	0	20	40
$\Gamma$ value	0	20	40	0	20	40
Convergence Iterations	33	39	51	58	72	97
Run time/(s)	21	33	49	40	52	88

optimal routes have not changed; when  $\Gamma = 20$ , the total risk optimal route has changed relatively small; however, the total time optimal routes have not changed when the robustness control parameters  $\Gamma = 0$ ,  $\Gamma = 10$ , and  $\Gamma = 20$ , which means that uncertainty data are taken nominal value, and the total risk and time optimal routes always have some robustness. Aiming at distribution center b, when the robustness control parameters  $\Gamma = 10$  and  $\Gamma = 20$ , the total risk optimal routes have not changed; when the robustness control parameters  $\Gamma = 0$ ,  $\Gamma = 10$ , and  $\Gamma = 20$ , the total time optimal vehicle routes do not have any change, which shows the total risk optimal distribution route is relatively stable. Aiming at distribution center c, when the robustness control parameters  $\Gamma = 10$  and  $\Gamma = 20$ , the total risk optimal vehicle routes do not have any change, but the total time optimal vehicle routes are changed when the robustness control parameters  $\Gamma = 0$ ,  $\Gamma = 10$ , and  $\Gamma = 20$ , which shows when  $\Gamma = 10$  the total risk optimal distribution route is relatively stable, to a certain extent, which can be selected as distribution route, and the total time optimal vehicle route has relatively weak robustness, and if more stable distribution route is needed, it will be expected to increase the robustness control parameter to find the strong robustness distribution route.

For solution robustness, as the robustness control parameters get bigger, the corresponding Pareto solution robustness should be enhanced in theory. After a large number of analyses for the actual situation and basic data, we can determine the robustness control parameter value and obtain the corresponding candidate route set.

The strength Pareto genetic algorithm (SPEA) is used to test the efficiency of the FCMC-PSO algorithm. The parameters of the SPEA are set as follows: population size is 50, maximum evolution generation is 300, crossover rate is 0.8, and mutation rate is 0.05. The solution set can be obtained by calculation. The results are shown in Table 17. Compared with SPEA, the convergence iterations and operation time of FCMC-PSO algorithm are reduced. The results show that the FCMC-PSO algorithm designed in this paper not only can obtain a more satisfactory solution but also has faster convergence speed compared with the SPEA.

## 5. Conclusion

Hazmat transportation route optimization is an important link to ensure transportation safety of hazmat. In this paper, we take the hazmat transportation route problem with multidistribution center as the research object, considering the transportation risk and transportation time. In addition, an adjustable robustness transportation route multiobjective robust optimization model is established in

the end. Speaking of the solution, the FCMC-PSO algorithm is designed in this research. The demand points were assigned through FCMC algorithm in which transportation time and transportation risk are considered. The multiobjective route robust optimization model is solved by multiobject PSO algorithm based on adaptive archives grid. In the end, the example shows that the robust optimization model and FCMC-PSO algorithm can obtain different robustness Pareto solution sets. The robust optimization transportation routes of hazmat will provide basic theory support for safeguarding the transportation safety of hazmat.

In this paper, we only consider the two uncertainties including transportation risk and transportation time. However, there may be some other uncertainties such as customer demand and service time window in the real world. According to this situation, we need to establish the corresponding robust model for future study. Although most of the hazmats are transported by road transportation, the hazmat transportation risks induced by other modes cannot be ignored. The optimization research of multimodal transport modes for the hazmat needs to be further studied.

## Data Availability

The 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 research was funded by National Natural Science Foundation of China (no. 71861023 and no. 51808057), the Program of Humanities and Social Science of Education Ministry of China (no. 18YJC630118), Lanzhou Jiaotong University (no. 201804), and Hunan Key Laboratory of Smart Roadway and Cooperative Vehicle-Infrastructure Systems (no. 2017TP1016).

## References

- [1] W. R. Rhyne, *Hazmat Transportation Risk Analysis: Quantitative Approaches for Truck and Train*, Wiley, New York, NY, USA, 1994.
- [2] G. Walker, J. Mooney, and D. Pratts, "The people and the hazard: the spatial context of major accident hazard management in Britain," *Applied Geography*, vol. 20, no. 2, pp. 119–135, 2000.
- [3] M. Power and L. S. McCarty, "Risk-cost trade-offs in environmental risk management decision-making," *Environmental Science & Policy*, vol. 3, no. 1, pp. 31–38, 2000.
- [4] T. Wu, C. Low, and J. Bai, "Heuristic solutions to multi-depot location-routing problems," *Computers & Operations Research*, vol. 29, no. 10, pp. 1393–1415, 2002.
- [5] B. Y. Kara, E. Erkut, and V. Verter, "Accurate calculation of hazardous materials transport risks," *Operations Research Letters*, vol. 31, no. 4, pp. 285–292, 2003.
- [6] B. Fabiano, F. Currò, A. P. Reverberi, and R. Pastorino, "Dangerous good transportation by road: From risk analysis to emergency planning," *Journal of Loss Prevention in the Process Industries*, vol. 18, no. 4–6, pp. 403–413, 2005.
- [7] E. Erkut and A. Ingolfsson, "Transport risk models for hazardous materials: revisited," *Operations Research Letters*, vol. 33, no. 1, pp. 81–89, 2005.
- [8] R. Bubbico, G. Maschio, B. Mazzarotta, M. F. Milazzo, and E. Parisi, "Risk management of road and rail transport of hazardous materials in Sicily," *Journal of Loss Prevention in the Process Industries*, vol. 19, no. 1, pp. 32–38, 2006.
- [9] S. R. Lin, L. P. Cai, and D. Y. Lin, "Effects of electroacupuncture of "Zusanli" (ST 36) on gastric mucosal blood flow, NO and ET contents in gastric mucosal injury rats," *Acupuncture Research*, vol. 31, no. 2, pp. 110–112, 2006.
- [10] M. Verma, "A cost and expected consequence approach to planning and managing railroad transportation of hazardous materials," *Transportation Research Part D: Transport and Environment*, vol. 14, no. 5, pp. 300–308, 2009.
- [11] Y. P. Wang, W. C. Sun, and S. W. Li, "Route optimization model for urban hazardous material transportation based on Arc GIS," *Journal of Jilin University*, vol. 39, no. 1, pp. 45–49, 2009.
- [12] J. Jassbi and P. Makvandi, "Route selection based on soft MODM Framework in transportation of hazmat," *Applied Mathematical Sciences*, vol. 63, no. 4, pp. 3121–3132, 2010.
- [13] 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.
- [14] S.-W. Chiou, "A bi-objective bi-level signal control policy for transport of hazardous materials in urban road networks," *Transportation Research Part D: Transport and Environment*, vol. 42, pp. 16–44, 2016.
- [15] G. Assadipour, G. Y. Ke, and M. Verma, "A toll-based bi-level programming approach to managing hazardous materials shipments over an intermodal transportation network," *Transportation Research Part D: Transport and Environment*, vol. 47, pp. 208–221, 2016.
- [16] D. Pamučar, S. Ljubojević, D. Kostadinović, and B. Đorović, "Cost and risk aggregation in multi-objective route planning for hazardous materials transportation—A neuro-fuzzy and artificial bee colony approach," *Expert Systems with Applications*, vol. 65, pp. 1–15, 2016.
- [17] M. Mohammadi, P. Jula, and R. Tavakkoli-Moghaddam, "Design of a reliable multi-modal multi-commodity model for hazardous materials transportation under uncertainty," *European Journal of Operational Research*, vol. 257, no. 3, pp. 792–809, 2017.
- [18] A. Kheirkhah, H. Navidi, and M. M. Bidgoli, "A bi-level network interdiction model for solving the hazmat routing problem," *International Journal of Production Research*, vol. 54, no. 2, pp. 459–471, 2016.
- [19] G. A. Bula, C. Prodron, 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.
- [20] C. Ma, Y. Li, R. He, F. Wu, B. Qi, and Q. Ye, "Route optimisation models and algorithms for hazardous materials transportation under different environments," *International Journal of Bio-Inspired Computation*, vol. 5, no. 4, pp. 252–265, 2013.
- [21] C. Ma, W. Hao, F. Pan, W. Xiang, and X. Hu, "Road screening and distribution route multi-objective robust optimization for

- hazardous materials based on neural network and genetic algorithm,” *PLoS ONE*, vol. 13, no. 6, pp. 1–22, 2018.
- [22] A. Ben-Tal and A. Nemirovski, “Robust convex optimization,” *Mathematics of Operations Research*, vol. 23, no. 4, pp. 769–805, 1998.
- [23] D. Bertsimas and M. Sim, “Robust discrete optimization and network flows,” *Mathematical Programming*, vol. 98, no. 1–3, pp. 49–71, 2003.
- [24] C. Ma, W. Hao, R. He et al., “Distribution path robust optimization of electric vehicle with multiple distribution centers,” *PLoS ONE*, vol. 13, no. 3, pp. 1–16, 2018.
- [25] R. Yu, X. Wang, and M. Abdel-Aty, “A hybrid latent class analysis modeling approach to analyze urban expressway crash risk,” *Accident Analysis and Prevention*, vol. 101, pp. 37–43, 2017.
- [26] C. Ma, C. Ma, Q. Ye et al., “An improved genetic algorithm for the large-scale rural highway network layout,” *Mathematical Problems in Engineering*, vol. 2014, Article ID 267851, 6 pages, 2014.
- [27] Y. Wang, X. Ma, Z. Li, Y. Liu, M. Xu, and Y. Wang, “Profit distribution in collaborative multiple centers vehicle routing problem,” *Journal of Cleaner Production*, vol. 144, pp. 203–219, 2017.
- [28] C. Ma and R. He, “Green wave traffic control system optimization based on adaptive genetic-artificial fish swarm algorithm,” *Neural Computing and Applications*, vol. 26, no. 5, pp. 1–11, 2015.
- [29] Y. Wang, K. Assogba, Y. Liu, X. Ma, M. Xu, and Y. Wang, “Two-echelon location-routing optimization with time windows based on customer clustering,” *Expert Systems with Applications*, vol. 104, pp. 244–260, 2018.
- [30] Z. He, W. Guan, and S. Ma, “A traffic-condition-based route guidance strategy for a single destination road network,” *Transportation Research Part C: Emerging Technologies*, vol. 32, pp. 89–102, 2013.
- [31] C. Ma, R. He, W. Zhang, and X. Ma, “Path optimization of taxi carpooling,” *PLoS ONE*, vol. 13, no. 8, pp. 1–15, 2018.
- [32] X. Yang, A. Chen, B. Ning, and T. Tang, “A stochastic model for the integrated optimization on metro timetable and speed profile with uncertain train mass,” *Transportation Research Part B: Methodological*, vol. 91, pp. 424–445, 2016.
- [33] C. X. Ma, Y. Z. Li, R. C. He et al., “Bus-priority intersection signal control system based on wireless sensor network and improved particle swarm optimization algorithm,” *Sensor Letters*, vol. 10, no. 8, pp. 1823–1829, 2012.
- [34] C. A. Coello and M. S. Lechuga, “MPSO: a proposal for multiple objective particle swarm optimization,” in *Proceedings of IEEE Congress on Evolutionary Computation (CEC)*, pp. 1051–1056, IEEE Computer Society, New Jersey, NJ, USA, May 2002.

## Research Article

# Hydrological Layered Dialysis Research on Supply Chain Financial Risk Prediction under Big Data Scenario

Jia Liu, Shiyong Li , and Xiaoxia Zhu 

*School of Economics and Management, Yanshan University, Qinhuangdao 066004, China*

Correspondence should be addressed to Xiaoxia Zhu; zhuxiaoxia@ysu.edu.cn

Received 21 August 2018; Accepted 17 September 2018; Published 8 October 2018

Academic Editor: Lu Zhen

Copyright © 2018 Jia 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.

In recent years, internet development provides new channels and opportunities for small- and middle-sized enterprises' (SMEs) financing. Supply chain finance is a hot topic in theoretical and practical circles. Financial institutions transform materialized capital flows into online data under big data scenario, which provides networked, precise, and computerized financial services for SMEs in the supply chain. By drawing on the risk management theory in economics and the distributed hydrological model in hydrology, this paper presents a supply chain financial risk prediction method under big data. First, we build a "hydrological database" used for the risk analysis of supply chain financing under big data. Second, we construct the risk identification models of "water circle model," "surface runoff model," and "underground runoff model" and carry on the risk prediction from the overall level (water circle). Finally, we launch the supply chain financial risk analysis from breadth level (surface runoff) and depth level (underground runoff); moreover, we integrate the analysis results and make financial decisions. The results can enrich the research on risk management of supply chain finance and provide feasible and effective risk prediction methods and suggestions for financial institutions.

## 1. Introduction

Internet development provides new channels and opportunities for SME financing. Supply chain finance (SCF) is a recent stream of research aimed at optimizing financial flows through solutions implemented by financial institutions [1]. As a new theory, supply chain finance is still in the exploratory stage in practice, especially in risk management and control. The research on the theory of supply chain financial risk prediction is far behind the actual application requirements [2]. The relevant references are mainly concentrated on the following two aspects. The first category focuses on the conceptual and theoretical aspect of supply chain finance. Berger [3] first put forward some new ideas and new frameworks for the financing of SMEs. Kaplan [4] proposed the concept of "electronic center" and established a supply chain aggregation model. Buyers and sellers can provide financial services in the "electronic center" field. Corning [5] proposed that B2B e-commerce transactions created conditions for

real-time financing, and B2B transaction participants should form alliances with financial service organizations to develop credit, financing, and dispute processing services. Basu [6, 7] conducted a research on the prepayment financing model and pointed out that the prepayment of financing orders could effectively solve the lag problems of logistics, and established an online data analysis platform. He believed that online supply chain finance played an important role in shortening the repayment period. Swift [8] believed that online supply chain finance integrated all parties in the supply chain, improved information exchange data and business operation efficiency, and reduced financing costs. Templar [9] pointed out that SCF is at the evolutionary frontier of financial services that are closely related to the supply chain cycle. Kouvelis [10] studied the impact of credit ratings on operational and financial decisions of a supply chain with a supplier and a retailer interacting via an early payment discount contract. Hofmann [11, 12] proposed an approach that allows all parties involved in supply chain finance to

improve the working capital and creates a win-win situation; and he believed that blockchain technology played an important role in accelerating capital flow in online supply chain finance projects. The second category focuses on the risk prediction of supply chain finance. Tsai [13] proposed a cash flow risk model using the standard deviation between cash inflows, outflows, and net cash flows at different time periods to perform risk control to predict the company's risk. Ralf [14] believed that the core companies in the supply chain should make choices in the relationship between supplier capital occupation and supplier relationship maintenance in order to prevent risk transfer. David [15] comparatively analyzed the risks and risk management methods of supply chain financial management before and after shipment and established the infrastructure of supply chain financial management. Bandyal [16] pointed out that the risk occurrence in any node of supply chain finance was very likely to lead to the risk of the entire supply chain, resulting in the collapse of the entire supply chain. Trott [17] specifically proposed the role of supply chain financing in promoting the development of SMEs, pointing out that the risk sources mainly came from finance and operations. Peter Finch [18] pointed out that the nonimprovement of information systems would increase the risk exposure level of banks, and he specifically emphasized that it was necessary to strengthen the ability of small and micro enterprises to continue their operations and pay close attention to the capacity of supply chain financial information systems. Yu [19] used LDSPPC model to reveal the linear relationship of financial risk level of online supply chain. Yan [20] constructed a Stackelberg game in supply chain finance system and found that incorporating bank credit with a credit guarantee can effectively balance the retailer's financing risk between the bank and the manufacturer through interest rate charging and wholesale pricing. Chen [21] builds a two-stage supply chain financial decision model on the order quantity and wholesale price. Babich [22] pointed out that risk management is a promising direction for supply chain finance.

Different from the above research work, this paper addresses the key issues in risk management in the process of supply chain financing under big data. We adopt distributed hydrological model to build the "hydrological database" and, respectively, construct surface runoff model and underground runoff model to conduct the risk prediction from the breadth and depth dimensions. Finally, we integrate and aggregate the analysis results to determine the risk level of the financing enterprise, which help the financial institutions to make financing decisions.

## 2. Model Description

### 2.1. General Description of the Hydrologic Model

**2.1.1. Water Cycle Process.** The water cycle includes several major links, such as precipitation, runoff, canopy retention, infiltration, and evaporation. In many of these links, there will be conversions of matter and energy and the transformation of water volume. At the same time, the water cycle will also be

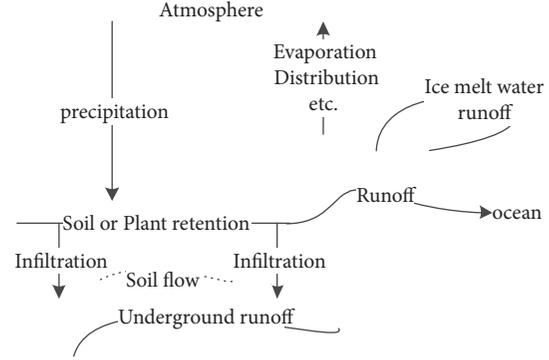


FIGURE 1: Schematic diagram of water circle cycle model.

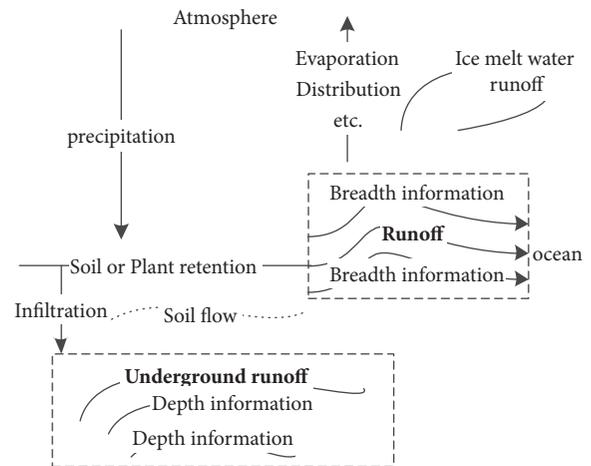


FIGURE 2: Schematic diagram of surface runoff model and underground runoff model.

influenced by atmospheric precipitation, weather and climate changes, and other factors such as landforms, topography, and human activities. Therefore, the water cycle is a very complicated process, which is shown in Figure 1.

All the solid, liquid, and gaseous trisate waters on Earth constitute the water circle of the earth. While the conditions are stable, the vapor in the atmosphere condenses and falls to the surface in the form of precipitation and under the action of the earth's gravity. Part of the precipitations return to the atmosphere through evaporation and distribution, and some precipitations form the surface runoff and finally flow into the ocean; and some precipitations flow into the soil to form a midstream stream; and some precipitations penetrate deep into the ground to form underground runoff [23]; some precipitation is formed at high latitudes and forms glacial snow, which melts and forms ice-melted-water runoff. The water bodies in the hydrosphere follow the cycle of gravity and solar radiation, and this process is called water circle cycle. The submodels included in the water cycle are as follows: "evapotranspiration" model; "surface runoff" model; "soil flow" model; "subterranean runoff" model; "ice-melting water runoff" model. Each model has its own characteristics, which is shown in Figure 2.

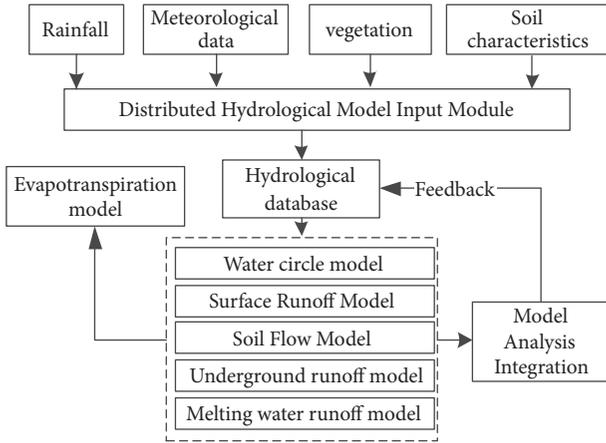


FIGURE 3: Risk prediction flow chart based on the distributed hydrology model.

2.1.2. *Distributed Hydrological Model.* Hydrology experts have invented and developed various methods and theories of flood forecasting from the unexplained rules of flood outbreaks that had caused mankind many great disasters. Hydrological model is one of them [23]. By using physical models and mathematical parameters to describe the actual hydrological system, it simulates and predicts the dynamic hydrological variables in the hydrological system under certain conditions. Among hydrological models, the distributed hydrological model is most widely used, and the mathematical model of the hydrological process of a certain watershed is simulated by the method of water cycle dynamics. The distributed hydrological model is a simplified representation of a complex hydrological system that is generally summarized and abstracted by academic community.

Distributed hydrological models have good early warning effects on floods. The working process is as follows: the river hydrological data (including rainfall, meteorological data, vegetation, and soil characteristics) are input into the distributed hydrological model, and the hydrological database summarizes and stratifies the data. The layered data are input into the water circle model, surface runoff model, soil flow model, underground runoff model, and melting water runoff model for analysis and integration to predict the flood. The final result is fed back to the hydrological database, as shown in Figure 3.

2.2. *Supply Chain Financing Risk Prediction Based on Hydrologic Model.* In this section, we apply the distributed hydrology model to the supply chain finance field and propose a supply chain financial risk prediction model based on hydrological layered dialysis method under big data scenario. We focus mainly on two models of surface runoff and underground runoff for risk prediction. The flow chart is shown in Figure 4.

The process of hydrological layered dialysis model for forecasting supply chain financial risk based on big data scenario is as follows: we firstly establish a “hydrological database” through collecting data on the trading volume of

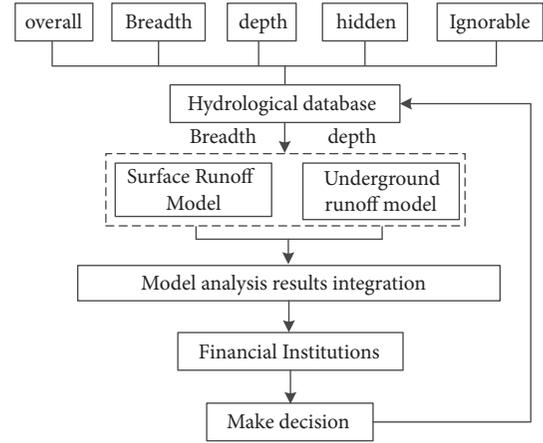


FIGURE 4: Financial risk prediction process based on hydrological layered dialysis method.

financing companies, bank cash, inflow and outflow data, and the value of changes in the credit index of small- and medium-sized enterprises. The “hydrological database” covers data for multiple dimensions of financing companies, including overall, breadth, depth, concealment, negligible, and external stimulus data. Among these, the overall data refers to the basic data of the financing enterprise itself for the current year; the breadth data refers to the data of other similar companies in comparison with financing companies; the in-depth data refers to the historical data that the financing companies themselves need to dig in depth; the covert data refers to the financing company’s own opaque, undisclosed data, and the negligible data refers to data that has minimal impact on the company’s own risk; the external stimulus data refers to data that cannot be controlled by the enterprise itself and is generated through external stimulation. In this paper, we mainly focus on the breadth and depth data for risk prediction. The collected breadth and depth information are used to determine the risk status of financing companies through “surface runoff” and “underground runoff” model analysis. Finally, through the integration of model analysis results, the bank makes final financing decisions.

2.3. *Hydrological Layered Dialysis Model Description for Supply Chain Financing Risk Prediction*

2.3.1. *“Water Circle” Model.* The risk evaluation indicator system in “water circle” model is established. In hydrology, people often set accommodation values for rivers, and there is a higher probability of flood peaks while this accommodation value is approached.

In the field of supply chain finance, banks also have such accommodation value while predicting financing risks. While the predicting risk value is close to the accommodation value, banks will probably face default risk. In this paper, we take risk theory and literature review as reference and construct a supply chain financing risk indicator system, which is shown in Table 1.

TABLE 1: Supply chain financing risk indicator system based on “water circle” model.

First-level indicators	Second-level indicators	Third-level indicators
Macro level	Business scenario risk indicators	Macro System Risk
		Industry System Risk
Meso level	Supply Chain Risk Indicators	Supply chain stability
		The degree of supervision of logistics enterprises
	Core corporate risk indicators	Enterprise strength
		Profitability
		Credit ability
Micro level	Financing Enterprise Risk Indicators	Short-term solvency
		Corporate quality
		Management capacity
		Currency funds (profitability)
	Collateral risk indicators	Solvency
		Development ability
		Inventory
		accounts receivable
		Prepayments

### 2.3.2. “Surface Runoff” Model

(1) *The Risk Evaluation Indicator System in “Surface Runoff” Model Is Established.* When extracting breadth information from hydrological database, considering the criterion of “easier to obtain and more representative compared with similar related financing enterprises,” we select six types of indicators with greater weight for “surface runoff” model establishment from the “water circle” indicator system. The indicator system contains the scale of financing companies, the cooperation closeness with upstream and downstream companies, the return rate of tangible assets, the asset-liability ratio, the quality of revenue from operating activities, and the ratio of cash flow liabilities. At the same time, the repayment risk levels of similar enterprises are extracted from the hydrological database. These seven type indicators are used as the basis for establishing the “surface runoff” model.

Based on this, we select a sample of financing data and adopt the decision tree analysis method where seven types of indicators are used as nodes. Firstly we divide the samples to form a decision tree and then transform it into a rule set and use the pessimistic estimation to determine whether a pruning is needed and a classification rule set is constructed; lastly we obtain a set of “surface runoff” credit risk classification rules, as shown in Table 2.

(2) *“Surface Runoff” Risk Prediction Model Based on Naive Bayesian.* Naive Bayes is a prediction method for risk classification. It is characterized by pruning and classifying various reference indicators and objectively predicts the risk of a certain object through model calculation. Therefore, it is widely used in the supply chain risk management field. This section adopts Naive Bayes method to construct a “surface runoff” risk prediction model.

The model establishment process is as follows: we assume that  $X$  is a data sample and  $C$  is the class label. The task of naive Bayes classification is to assign the unknown sample  $X$  to  $C_i$  with the highest posterior probability, where  $P(C_i | X) > P(C_j | X)$ ,  $1 \leq j \leq m$ ,  $i \neq j$ . At this point, the maximum  $P(C_i | X)$  corresponding to the highest posterior probability  $C_i$  is called the maximum posteriori assumption. According to Bayes’ theorem, we obtain that

$$P(C_i | X) = \frac{P(X | C_i) P(C_i)}{P(X)}. \quad (1)$$

Among these,  $P(C_i)$  represents the prior probability and  $P(X | C_j)$  represents the conditional probability. If the prior probability  $P(C_i)$  is unknown, then it is assumed that the classes are of equal probability; that is,  $P(C_1) = P(C_2) = P(C_3) = \dots = P(C_m)$ . If  $P(C_j)$  is known, then maximize  $P(X | C_j)P(C_j)$ , the prior probability is obtained with  $P(C_j) = s_j/s$ , where  $s_j$  is a training sample of  $C_j$ , and  $s$  is the total number of training samples. If the complex dataset has multiple attributes, then  $P(X | C_i)$  is difficult to be solved. At this time, we can make simple assumptions of independent conditions.

Based on the above steps, the risk status of the financing company can be calculated by adopting “surface runoff” model under “breadth information” condition.

### 2.3.3. “Underground Runoff” Model

(1) *The Risk Evaluation Index System in the “Underground Runoff” Model Is Established.* According to a 2008 Aberdeen Group Research Report on “Supply Chain Financial Market State,” four historical indicators that best reflect the financial performance of the supply chain are cited in the

TABLE 2: Credit risk classification rule set for corporate loan.

Attribute name	Ranges
Financing enterprise size	Large, medium and small
Cooperation with upstream and downstream companies	High, Medium and Low
Return on tangible assets	$>40, \leq 40$
Assets and liabilities	$>15, 8-15, <8$
Business activity revenue quality	$>90, \leq 90$
Cash flow to debt ratio	$>0.8, \leq 0.8$
Repayment risk (risk rate)	High, Medium and Low

establishment of the “surface runoff” model indicator system. We adopt monetary fund, accounts receivable, prepayment amount, and inventory amount as key indicators for analyzing the depth information of “underground runoff” model.

(2) “Underground Runoff” Risk Prediction Model Based on Cusp Mutation. The cusp mutation is a method of predicting the risk mutation. The characteristic is to discuss the risk through mutation. We use cusp mutation method to construct “underground runoff” risk prediction model.

The model establishment process is as follows: the cusp abrupt balance surface has a bifurcation set B. If the control variables in the surface cause the changing of the state variable while passing through the bifurcation set B and generate a sudden change, then the system will face unstable situation; if the control variables do not pass through the bifurcated set B, then the state variable is stable and the system is stable.

The cusp mutation function equation is  $V(x) = x^4 + mx^2 + nx + c$ . In the formula,  $m, n$  are control variables and  $x$  is a state variable. Assume that  $V'(x) = 0$ , and then we derive the equilibrium surface equation:  $4x^3 + 2mx + n = 0$ . Furthermore, we assume that  $V''(x) = 0$  and find the set of function singularities  $12x^2 + 2m = 0$ . After pairing the balanced surface with the singularity set, we obtain the bifurcation set,  $8m^3 + 27n^2 = 0$ , and the forked set is a set of points in a balanced surface that allows all  $x$  to jump. The real roots number of the equilibrium surface equation needs to be determined by the sign of the discrimination  $\Delta = 8m^3 + 27n^2$ . The results can be divided into four situations:  $\Delta > 0$  stands for the fact that the system is stable;  $\Delta = 0$  represents that the critical value of system is stable; while  $\Delta < 0$  and the error between the mutation and the normal value is within  $1/3$ , it means the system is relatively stable, and while  $\Delta < 0$  and the error between the mutation and the normal value is more than  $1/3$ , it means the system is instable. Based on the above analysis, it is possible to calculate the risk status of the financing company by “underground runoff” model under “depth information.”

### 3. Empirical Analysis

Shenli company (“S” for short) is a professional company specializing in the production of lifting equipment and supporting products. At present, in order to meet the customers’ demand, S puts forward the plans for expanding

the production scale and introducing high-quality products. However, it is difficult to quickly realize large-scale and specialized production based on its own funds. For this large funding gap, S applies 9 million RMB from M financial institution to ease financial pressure.

By investigation, M financial institution discovers that S relies on a core company in the supply chain. The cooperations between the two parties have been lasting for a long time and the business exchanges are stable. However, there are some certain risks in this loan. Therefore, M financial institution should analyze the financing risk of S. We extract the indicator data of the S’s breadth (surface runoff) and depth risk level (underground runoff) from the constructed “Hydrology Database” based on big data. Suppose that M financial institution has finished the “water circle” model selection, and then the risks of “surface runoff” and “underground runoff” models are needed to be constructed.

3.1. Application for “Surface Runoff” Risk Evaluation Model. The breadth level (surface runoff) data of S was analyzed by Naive Bayesian method. The M financial institution extracts the “surface runoff” indicators data of 15 companies of the same type from the “hydrological database,” which is shown in Table 3.

The surface runoff data of S is  $X = (\text{Financing enterprise size} = “6401”, \text{Business activity revenue quality} = “87”, \text{Cooperation with upstream and downstream companies} = “Medium”, \text{Return on tangible assets} = “41”, \text{Asset investment rate} = “18”, \text{Cash flow to debt ratio} = “1.1”)$ .

(1) Prior probability:  $P(\text{Repayment risk “Low”}) = 0.6$ ,  $P(\text{Repayment risk “High”}) = 0.4$ ;

(2) Conditional probability:

$P(\text{Financing enterprise size} = “6401”/\text{Repayment risk “Low”}) = 0.33$ ;

$P(\text{Business activity revenue quality} = “87”/\text{Repayment risk “Low”}) = 0.22$ ;

$P(\text{Cooperation with upstream and downstream companies} = “Medium”/\text{Repayment risk “Low”}) = 0.44$ ;

$P(\text{Return on tangible assets} = “41”/\text{Repayment risk “Low”}) = 0.67$ ;

$P(\text{Asset investment rate} = “18”/\text{Repayment risk “Low”}) = 0.44$ ;

$P(\text{Cash flow to debt ratio} = “1.1”/\text{Repayment risk “Low”}) = 0.78$ ;

Then:  $P(X/\text{Repayment risk “Low”}) = 0.00734$ ;

TABLE 3: Surface runoff indicators data for financing enterprises extracted from hydrological model database.

number	Financing enterprise size	Business activity revenue quality	Cooperation with upstream and downstream companies	Return on tangible assets	Asset investment rate	Cash flow to debt ratio	Repayment risk
1	≤4000	>85	Medium	>45	>15	≤1.3	Low
2	>40000	>85	High	≤45	>15	≤1.3	Low
3	>40000	>85	High	≤45	8-15	≤1.3	Low
4	4000-40000	≤85	Low	>45	8-15	≤1.3	High
5	4000-40000	>85	High	≤45	>15	≤1.3	Low
6	4000-40000	≤85	Medium	>45	8-15	>1.3	High
7	4000-40000	>85	Medium	>45	<15	<1.3	Low
8	4000-40000	>85	Low	≤45	<8	>1.3	High
9	4000-40000	≤85	Medium	≤45	<8	>1.3	High
10	4000-40000	≤85	Medium	≤45	<8	≤1.3	Low
11	≤4000	>85	Medium	≤45	8-15	>1.3	Low
12	≤4000	>85	Low	>45	<8	≤1.3	High
13	>40000	>85	High	>45	>15	>1.3	Low
14	≤4000	≤85	Low	>45	>15	>1.3	High
15	≤4000	>85	Low	>45	8-15	≤1.3	Low

TABLE 4: Underground runoff indicators data for financing enterprises extracted from hydrological model database.

Year	2010	2011	2012	2013	2014	2015	2016
Accounts receivable	170	1228	1841	3331	1784	1379	1499
Prepayments	661	1768	1361	1729	2123	2089	2361
currency funds	128	9	320	339	253	418	459
Inventory	1701	1231	1901	2339	2641	2491	2129

P (Financing enterprise size = “6401”/Repayment risk “High”) = 0.67;

P (Business activity revenue quality = “87”/Repayment risk “High”) = 0.67;

P (Cooperation with upstream and downstream companies = “Medium”/Repayment risk “High”) = 0.33;

P (Return on tangible assets = “41”/Repayment risk “High”) = 0.33;

P (Asset investment rate = “18”/Repayment risk “High”) = 0.33;

P (Cash flow to debt ratio = “1.1”/Repayment risk “High”) = 0.33;

Then: P (X/Repayment risk “High”) = 0.00532;

(3) The result is calculated:

P (X/Repayment risk “Low”)P(Repayment risk “Low”) = 0.004404;

P (X/Repayment risk “High”)P(Repayment risk “High”) = 0.002936;

From the above analysis, we find that the probability of low repayment risk of the financing enterprise is greater than the probability of high repayment risk. Therefore, through the analysis of (breadth level) surface runoff data through the hydrological layered dialysis model, the risk probability of the S is predicted to be “low.”

**3.2. Application for “Underground Runoff” Risk Evaluation Model.** The depth level (underground runoff) index data extracted by the “Hydrology Database” is shown in Table 4. We take the “year” as the independent variable X and take the four vertical index datasets as Y, using the least squares method in Matlab software to fit the quartic function of the cusp mutation.

Firstly, we calculate the fourth-order polynomial by fitting the accounts receivable, which is shown in formula (2):

$$Y_{\text{Accounts receivable}} = 39.28x^4 - 594.12x^3 + 2794.21x^2 - 3979.75x + 1944 \quad (2)$$

This is the fitting curve and error curve for 2010-2016, as shown in Figure 5. The blue line is the actual value, the red line is the fitting value, and the yellow line is the error curve value.

Secondly, we analyze the accounts receivable mutation. The functional equations of the cusp catastrophe model are derived from the polynomials through elementary changes. Set the polynomial as formula (3):

$$y(m) = a_1m^4 + a_2m^3 + a_3m^2 + a_4m + a_0 \quad (3)$$

Since there are no cubic terms in the cusp catastrophe model, we can make  $m = x - n$ ,  $n = a_2/4a_1$ , and the above formula can be changed to formula (4):

$$F(x) = b_1x^4 + b_2x^2 + b_4x + b_0 \quad (4)$$

among which, we get formula (5):

$$\begin{bmatrix} b_0 \\ b_4 \\ b_2 \\ b_1 \end{bmatrix} = \begin{bmatrix} n^4 & -n^3 & n^2 & -n & 1 \\ -4n^3 & 3n^2 & -2n & 1 & 0 \\ 6n^2 & -3n & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_0 \end{bmatrix} \quad (5)$$

Then we can make  $V(x) = F(x)/b_1$  to perform variable substitution, and then we get the standard function of the cusp catastrophe model as shown in formula (6):

$$V(x) = x^4 + ux^2 + vx + c \quad (6)$$

We can perform an elementary transformation on formula (2) and then get  $n = a_2/4a_1 = -3.78$ ,  $b_0 = 2756.20$ ,  $b_4 = 163.50$ ,  $b_2 = -575.63$ ,  $b_1 = 39.28$ . Then through variable substitution, we obtain  $u = -14.65$ ,  $v = 4.16$ ,  $c = 70.17$ .

The cusp mutation function is shown in formula (7):

$$V(x) = x^4 - 14.65x^2 + 4.16x + 70.17 \quad (7)$$

The bifurcation set is calculated as  $8u^3 + 27v^2 < 0$ .

Using the same cusp catastrophe analysis method, we obtained that the bifurcation sets for the prepayments, currency funds, and inventory are also less than zero by calculating the prepayments.

For S, despite the abrupt changes in accounts receivable, prepayments, currency funds, and inventory, we suppose that the reason is the sharp increase in business volume during the company’s own development process and the temporary shortage of working capital. Observing the data carefully, we can see that the absolute values of the error for the four indicators during the recent seven years are within one-third of the average value of the actual value, which indicates that the system mutation is benign despite the deterioration of the stability of the mutant system. Therefore, M financial institution can properly provide loan service for S. For the actual amount of financing provided by M financial institution, refer to Table 5.

TABLE 5: The actual financing amount provided by M financial institution.

error	0	>0 and <1/3
Financing amount	F	2F/3

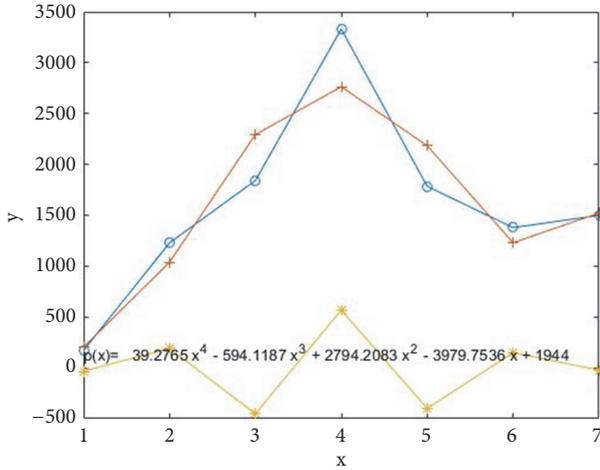


FIGURE 5: Fitting curve and error curve of accounts receivable.

3.3. *Integration Analysis for the Model Results.* The integration analysis process is as follows: firstly, the trader S will establish a “water circle” risk assessment indicator system and enter the “hydrological database.” Then, relevant horizontal (breadth level) data will be extracted from the “hydrological database” and be accepted by the “final runoff” classification prediction model of the M financial institution. If the M financial institution assesses that its risk is “low,” then it proceeds to the next step. If the risk is “high,” the M financial institution rejects the loan; finally, it extracts relevant vertical (depth) data from the “hydrological database” and accepts the “underground runoff” pointer mutation model screening. M financial institutions list three options for traders S: if condition 1 is satisfied, which means the fork set is greater than 0, then M financial institutions will provide traders with 9 million RMB of financing; if condition 2 is satisfied, which means the fork set is less than 0, but the error is within 1/3 of the average value of the actual value, then the financial institution will provide a financing amount of 6 million RMB. If condition 3 is satisfied, which means the fork set is less than 0 and the error is outside 1/3 of the actual value, then the financial institutions will refuse to provide loans.

Through the surface runoff model and the underground runoff model to analyze and forecast the financing risk of enterprise from the breadth and depth levels, the bank ultimately decides whether to finance the enterprise, so as to maximize the return of the bank under the premise of the lowest risk.

#### 4. Conclusions

The purpose of establishing a hydrological model is to predict the possible risk status of a certain river basin in advance. This paper applies the forecasting principle of the hydrological

model to the supply chain finance field and predicts the risk status of the financing enterprise in advance based on the hydrological layered dialysis model of the supply chain financial risk forecast under the big data scenario, which reduces the loan risk of financial institutions. Based on the hydrological database made by real time data, a quantitative analysis method is used to construct hydrological layered dialysis model and the financial institutions decide whether or not to lend financing companies through multidimensional and dynamic forecasting and evaluation of corporate financing risk based on supply chain under big data scenario.

We adopt the two dimensions of “surface runoff” and “underground runoff” to determine the financial risk; however, the indicators are not comprehensive. In the future research, the indicator system can be further improved by considering the risks in combination with the other models which can more effectively reduce the financing risk of financial institutions.

#### 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 would like to acknowledge the support from the Project Funded by Hebei Education Department (no. QN2018207), the National Natural Science Foundation of China (no. 71671159), and the Natural Science Foundation of Hebei Province (no. G2018203302).

#### References

- [1] S. Ma, “Differential dynamic evolutionary model of emergency financial service supply chain in natural disaster risk management,” *Discrete Dynamics in Nature and Society*, vol. 2016, Article ID 5103716, 6 pages, 2016.
- [2] J. Sun, C. Wang, X. Ji, and J. Wu, “Performance evaluation of heterogeneous bank supply chain systems from the perspective of measurement and decomposition,” *Computers & Industrial Engineering*, vol. 11, pp. 891–903, 2017.
- [3] A. N. Berger and G. F. Udell, “A more complete conceptual framework for SME finance,” in *Proceedings of the World Bank Conference on Small and medium enterprises: Overcoming growth constraints*, vol. 30, pp. 92–98, Washington: World Bank, 2004.
- [4] M. K. Sawhney, “E-hubs: the New B2B Marketplaces,” *Harvard Business Review*, vol. 78, no. 31, pp. 97–106, 2009.
- [5] O. Corning, “Decision rules for participating in B2B exchanges,” *Operations Management., Washington, Corporate Executive Board*, pp. 14–42, 2011.
- [6] P. Basu and S. K. Nair, “Supply Chain Finance enabled early pay: Unlocking trapped value in B2B logistics,” *International Journal*

- of Logistics Systems and Management*, vol. 12, no. 3, pp. 334–353, 2012.
- [7] P. Basu and S. K. Nair, “A decision support system for mean-variance analysis in multi-period inventory control,” *Decision Support Systems*, vol. 57, no. 1, pp. 285–295, 2014.
- [8] Swift, “A New Start for Supply Chain Finance,” [https://www.swift.com/si/sites/default/files/resources/swift\\_oursolutions\\_corporates\\_swift\\_new\\_start\\_supply\\_chain\\_finance\\_white\\_paper\\_er\\_201509.pdf](https://www.swift.com/si/sites/default/files/resources/swift_oursolutions_corporates_swift_new_start_supply_chain_finance_white_paper_er_201509.pdf), 2015.
- [9] S. Templar, E. Hofmann, and C. Findlay, *Financing the End-to-End Supply Chain: A Reference Guide for Supply Chain Finance*, Kogan Page, London, 2016.
- [10] P. Kouvelis and W. H. Zhao, “Who should finance the supply chain? Impact of credit ratings on supply chain decisions,” *Manufacturing & Service Operations Management*, pp. 19–35, 2017.
- [11] E. Hofmann and S. Zumsteg, “Win-win and no-win situations in supply chain finance: the case of accounts receivable programs,” *Supply Chain Forum*, vol. 16, no. 3, pp. 30–50, 2015.
- [12] E. Hofmann, U. M. Strewe, and N. Bosia, *Supply Chain Finance and Block chain Technology*, Springer-Verlag, Berlin, 2018.
- [13] C.-Y. Tsai, “On supply chain cash flow risks,” *Decision Support Systems*, vol. 44, no. 4, pp. 1031–1042, 2008.
- [14] R. W. Seifert and D. Seifert, “Financing the chain,” *International Commerce Review*, vol. 10, no. 1, pp. 32–44, 2011.
- [15] D. A. Wuttke, C. Blome, and M. Henke, “Focusing the financial flow of supply chains: An empirical investigation of financial supply chain management,” *International Journal of Production Economics*, vol. 145, no. 2, pp. 773–789, 2013.
- [16] D. Bandaly, A. Satir, and L. Shanker, “Integrated supply chain risk management via operational methods and financial instruments,” *International Journal of Production Research*, vol. 52, no. 7, pp. 2007–2025, 2014.
- [17] J. Rezaei, R. Ortt, and P. Trott, “How SMEs can benefit from supply chain partnerships,” *International Journal of Production Research*, vol. 53, no. 5, pp. 1527–1543, 2015.
- [18] P. Finch, “Supply chain risk management revisited,” *Supply Chain Management*, vol. 19, no. 3, pp. 142–156, 2016.
- [19] X. Yu, “Research on low dimensional successive projection pursuit modeling of online supply chain financial risk evaluation,” *Finance Theory and Practice*, vol. 12, pp. 49–52, 2017.
- [20] N. Yan, C. Liu, Y. Liu, and B. Sun, “Effects of risk aversion and decision preference on equilibriums in supply chain finance incorporating bank credit with credit guarantee,” *Applied Stochastic Models in Business and Industry*, vol. 33, no. 6, pp. 602–625, 2017.
- [21] J. Chen, “The optimal strategies of risk-averse newsvendor model for a dyadic supply chain with financing service,” *Discrete Dynamics in Nature and Society*, vol. 2017, Article ID 4861515, pp. 1–15, 2017.
- [22] V. Babich and P. Kouvelis, “Interface of operations and management: recent contributions and future directions,” *Manufacturing & Service Operations Management*, pp. 1–18, 2018.
- [23] Z. Xu, *Hydrological Models*, Science Press, Beijing, China, 2017.

## Research Article

# A Metaheuristic Method for the Task Assignment Problem in Continuous-Casting Production

Hongtao Hu <sup>1</sup>, Yiwei Wu,<sup>2</sup> and Tingsong Wang <sup>3</sup>

<sup>1</sup>College of Logistics Engineering, Shanghai Maritime University, Shanghai 201306, China

<sup>2</sup>School of Management, Shanghai University, Shanghai 200444, China

<sup>3</sup>School of Economics and Management, Wuhan University, Wuhan 430072, China

Correspondence should be addressed to Tingsong Wang; [emswangts@whu.edu.cn](mailto:emswangts@whu.edu.cn)

Received 2 July 2018; Accepted 14 September 2018; Published 2 October 2018

Guest Editor: Xinchang Wang

Copyright © 2018 Hongtao 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 steelmaking and continuous-casting (SCC) process in integrated iron and steel enterprises can be described as two stages: the upstream stage and downstream stage. Raw materials are transformed into molten steel in the upstream stage, while the downstream stage is responsible for transforming molten steel which is released at regular intervals and has a limited time for being turned into slabs. This article focuses on the task assignment problem in the downstream stage within the given information resulting from the upstream stage. This problem is formulated as a nonlinear mixed-integer programming model aimed at minimizing total tardiness within the resource constraints and time windows constraints for the tasks. An improved solution algorithm based on particle swarm optimization is developed to efficiently solve the proposed model. Finally, computational experiments are implemented to evaluate the performance of the solution algorithm in terms of solution quality and computational time.

## 1. Introduction

The iron and steel industry, as one of the most important economy sectors, is a large component of energy consumption and pollution emission. The energy consumption of iron and steel industry accounts for 10%–15% of the total energy consumption in the world, and the CO<sub>2</sub> emission accounts for about 7% of the total anthropogenic CO<sub>2</sub> emission [1]. Therefore, the iron and steel industry is an important study object because of the goal of emission reduction and energy saving. There are many scheduling tasks to be completed in the iron and steel industry. Because steelmaking and continuous-casting (SCC) process contributes the largest portion of the material loss in production and the SCC has important effects on the shape and quality of integrated iron and steel enterprises' final products (slabs), scheduling of SCC process is one of the most important tasks among different scheduling tasks [2]. Hence, studying the effective scheduling methods solving SCC problems has timely significance for improving the iron and steel industry productivity [3]. Figure 1 illustrates the material flow of an integrated iron and steel enterprise, which is a typical process industry, where final products are

produced by a sequence of processing operations (steelmaking, refining, and continuous-casting in sequence) to transform raw materials into solid slabs. The process can be described as two stages: an upstream stage and downstream stage. In the upstream stage, raw material is first turned into molten steel by a LDVC in order to initially reduce impurity content, which is called the steelmaking stage. And then, impurities are further eliminated by some refining techniques, such as RH and LF, which is called the refining stage. After the process of the upstream stage, the molten steel comes to the downstream stage, also called the continuous-casting stage, in which casting machines cast the molten steel into solid slabs uninterruptedly; the production process ends.

In the upstream stage, molten steel is produced in boilers. In order to save energy, upstream machines produce molten steel at regular intervals, which means that when boilers complete the production of a batch of molten steel, boilers continue to produce again rather than be cooled because it would require a lot of energy to restart boilers. Therefore, when upstream machines finish the production of molten steel, molten steel (upstream resource) is released immediately to the casting stage at regular intervals to service the casting machines processing of tasks.

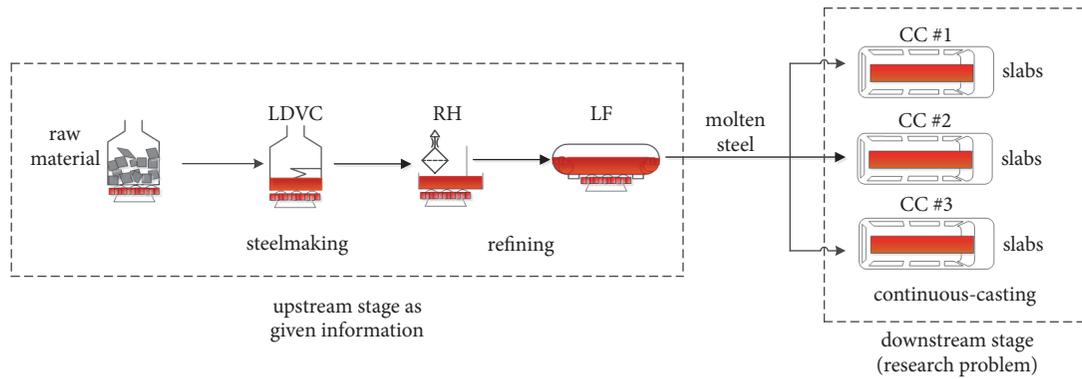


FIGURE 1: Steelmaking and continuous-casting process.

The temperature of molten steel supplied has very important effects on the quality of continuous-casting slabs. In order to prevent nozzle clogging caused by bad fluidity of low temperature molten steel, the continuous-casting system requires strict control of the temperature of molten steel. Even worse, low temperature molten steel can lead to freezing of nozzles, which would subsequently interrupt the casting. Therefore, once molten steel is released, it must be used within the resource available time, or the remaining resource will be cold and idle.

In this article, we focus on the casting plans of the continuous-casting process within given information resulting from the upstream stages of steelmaking and refining. The casting plan is to make a decision on how to distribute all tasks to casting machines while satisfying all constraints.

Our work is motivated by a bottleneck problem arising in the steelmaking and the continuous-casting stage of steel production. Due to lower production capacity, the SCC production is always regarded as a bottleneck and restricts enterprises' development. In steel production, molten steel is supplied to parallel casting machines in the continuous-casting stage, where every task needs its own amount of resources, and the total resource demanded does not exceed the resource supplied at any point in time. Therefore, during the whole processing time, resources are released at regular intervals, where the resource quantity at each resource release time is fixed.

The contributions of this article are as follows:

- (1) This study takes account of some realistic factors such as the resource available time for use, nonlinear supply resource, nonrenewable resource, and tardiness minimization. There are related studies that considered some of these factors, but few previous studies have covered all of these factors.
- (2) We formulate the task assignment problem as a nonlinear mixed-integer programming model and propose an efficient solution algorithm to solve the proposed model which combines the particle swarm optimization algorithm and artificial bee colony algorithm.

The remainder of this article is organized as follows. Section 2 provides an overview of related work. Section 3

describes the problem in detail and presents an integer programming model for the problem. In Section 4 we present an iterated local search approach for solving the problem. Section 5 presents computational experiments. Final remarks are given in Section 6.

## 2. Literature Review

Because the scheduling of SCC process plays an important role in the iron and steel industry, there exist numerous related studies on the SCC process. We review two streams of related literature on SCC scheduling problems. Readers interested in overviews can refer to Tang et al. [4], Tang et al. [5], Tang et al. [6], Pan [3], and Mori and Mahalec [2] to obtain a comprehensive overview on the scheduling of SCC process.

The first research stream is related to the task assignment problem (TAP). The TAP problem is a challenging combinatorial optimization problem which aims to assign different tasks to some distributed processors in a system in order to increase the effectiveness of the system and minimize a certain cost. The beginning of the development of practical solution methods solving the classic TAP is proposed by Kuhn [7]. Numerous solving methods have been presented by scholars for the generalized TAP which aims to find the minimum assignment cost. Studying the problem of assigning drivers to cover tasks with service time windows and uncertain task durations, Cheung et al. [8] propose an adaptive labelling solution procedure which incorporates various practical constraints and work rules. Yan and Tang [9] develop a heuristic approach which is embedded in a framework aimed for stochastic airport gate assignment under uncertain flight delays. Zhen et al. [10] provide a stochastic programming model and a robust optimization model to deal with task workload requirements. Zhen et al. [11] study a strategic-level decision problem on assigning production stages to minimize the fixed cost and the expected value costs under stochastic demands of customers. Task assignment in this article also belongs to the parallel machine scheduling problem which is one of the classical problems in production systems.

The second topic of related works is concerned with the parallel machine scheduling (PMS). The PMS problem is a relatively simple multimachine scheduling problem in

actual production and is applied in a very wide range of production, so it is studied by many researchers. According to the statistical data through the web of science (SCIT + CPCI), there are about 160,000 articles related on the study of the parallel machines scheduling problem, and the number increases every year. Although there have been many researchers developing the parallel machine theory and practical applications, PMS still attracts more and more researchers' interests. There are three main research directions related to the parallel machine problem now: the optimization of the parallel machine scheduling problem whose goal is to minimize total tardiness or makespan, the parallel machine scheduling problem considering a product family, machine processing equipment, etc., and the parallel machine scheduling problem with resource constraints. This article studies the parallel machine scheduling problem with resource constraints which aims to minimize total tardiness.

Many researchers have studied PMS problems and applied them to a variety of manufacturing cases due to their theoretical and practical significance, for example, in the semiconductor manufacturing industry [12] and the electronic industry [13]. However, machines are the only resource considered in most PMS studies. In most real-life manufacturing systems, additional resources, such as automated guided vehicles, machine operators, tools, pallets, and industrial robots, are required for some jobs [14].

The first systematic approach to PMS problems was proposed in the mid-1950s, which was further extended from single machine scheduling problems. Since then, thousands of articles have investigated this problem. Cheng and Sin [15] provide a comprehensive analysis of PMS. In the 1990s, many articles extended PMS problems. Allahverdi et al. [16] provide a comprehensive review of such problems, including dynamic and stochastic problem settings for single and parallel machines, flow shops, and job shops. Mokotoff [17] develops a new local search heuristic algorithm for parallel machine problems which minimizes the total weighted completion time. Pfund et al. [18] review and summarize the literature on nontraditional parallel machine scheduling problems.

Comprehensive surveys for the tardiness minimization scheduling problem are provided by Dileepan [19] and Koulamsa [20]. Compared to Emmons et al. [21] who study tardiness in a single machine scheduling problem, Du and Leung [22] propose a better solution using different algorithms and describe the optimal solution of the problem. Lawler [23] proposes a decomposition algorithm that decomposes the problem into multiple noninterfering subproblems. Later, Shim and Kim [24] propose to use the branch-and-bound method to solve a single machine problem aimed at tardiness minimization. More recently, Lee and Kim [25] apply the branch-and-bound method to parallel machine scheduling and tardiness minimization. In addition, Mensendiek et al. (2015) develop a mathematical model for optimal sequences on parallel machines with fixed delivery dates and present heuristics to solve the problem.

Apart from this theoretical analysis, there are many researchers studying real problems related to industrial activities and considering the problems of resource constraints.

Blazewicz et al. [26] point out that job processing not only needs a machine, but also needs automatic guided vehicles, machine operators, tools, pallets, abrasives, industrial robots, and other additional resources. Blazewicz et al. [27] discuss the single machine scheduling problem with a continual resource and also discuss the nonrenewable resources proposed by Shabtay and Kaspi [14]. Grigoriv et al. (2005) consider the scheduling problem with raw material resource constraints, mainly with the following three constraints: each job has its own raw material, all jobs require a generic resource, and each job has many kinds of raw materials and has to choose one or some materials. The objective of different models in the above articles is to minimize the processing time, or minimize the maximum number of late jobs. Briskorn et al. [28] study the single machine scheduling problem aimed at minimizing the processing time or the maximum number of late jobs when the production process should stop if the resource demand cannot be met. Briskorn et al. [29] study the single machine scheduling problem with inventory constraints, and the goal is to minimize the total processing time. In order to solve the above problems, the branch-and-bound method and dynamic programming method are used to find the optimal solution. It was found that the heuristic algorithm outperforms a state-of-the-art commercial mixed integer programming solver even with 20 jobs. Györgyi et al. [30] study resource constraints in single machine scheduling, such as a nonrenewable raw material, where the goal is to minimize the total processing time. Wang et al. [31] study the PMS problem with two kinds of resource constraints (machines and molds), and the goal is to minimize the total processing time. It discusses preventive maintenance activities for the flexibility of resources.

In summary, many previous studies on SCC optimization problems disregarded the relationship between the PMS problem and the TSP problem. Although few studies have considered this relationship, they did not consider many realistic factors, such as the nonlinear supplied and required resource, and resource available time. Therefore, this article focuses on the scheduling problem with nonlinear required and consumed resource constraints aimed at minimization of total tardiness under parallel machines, and the production resource is a nonrenewable resource and has limited time for use.

This study conducts a comprehensive investigation of the SCC problem by considering several realistic factors, such as nonlinear supplied and required resource, and resource available time; these factors are related to the practical operation and production management in the iron and steel industry. In addition, an improved algorithm is also presented. In comparison to the existing literature, the model proposed in this study can provide more reasonable SCC scheduling plans in real-world application.

### 3. Problem Statement and Mathematical Model

*3.1. Problem Description.* The production scheduling problem studied in this article can be described as follows. In iron and steel enterprises, the main production and scheduling

problems are the furnace plans and casting plans. The furnace plan determines how many batches of molten steel should be produced in the upstream stage. The casting plan means the pouring order of all tasks which are distributed to the same casting machine. This article focuses on how to distribute casting machines in the casting stage, which implies that, in order to achieve multifurnace casting and reduce energy consumption, the continuous-casting process in an iron and steel enterprise should have each casting machine produce as many casting sets as possible. Therefore, the main problem to be solved in this article is to determine each task's casting machine and then determine the pouring order of each casting machine. Specifically, there are many tasks waiting to be processed by parallel casting machines. Tasks must be completed one by one because each machine can only process one task at a time. Additionally, each task requires a certain amount of common resource that is provided by the upstream stage.

The key of continuous-casting production schedules is the upstream resource constraints. An example of a production schedules is shown in Figure 2 which illustrates the implications of resource constraints over time. There are 6 tasks which should be processed by any one of two available machines in the continuous-casting process. At each resource's release time, a fixed amount of resource from the upstream stage will be released to service casting machines processing tasks within the resource available time. (1, *a*) in Figure 2 represents that task 1 is processed on machine *a*, and (2, *b*) represents that task 2 is processed on machine *b*. At time 0, upstream resources are released first. At time  $C_1$ , task 1 is finished on machine *a*, so task 3 is processed on machine *a*, and the total resource requirement increases. At time  $C_3$  where task 3 is finished on machine *a*, because the remaining resource is not enough to process another task, machine *a* will stop processing till the next resource release time. At time  $C_2$ , task 2 is finished on machine *b*, but the remaining resource is not enough to process any more tasks, so machine *b* has to wait till the next resource release time, too.  $f$  is the fixed quantity of released resource at each resource release time. The cumulative amount of resources required cannot exceed the cumulative amount of resource supplied at any time, and both of them are nonlinear over time. We assume that each task consumes resources at the beginning of the process.

As can be seen from Figure 3, since the cumulative resource demanded cannot exceed the cumulative resource supplied at any time, tasks may have a waiting time when they arrive and cannot be processed, and the remaining resource at the last resource release time is idle. Besides, since the upstream resource has a limited time for use, it will also be idle because of low temperature. Let  $N$  denote the set of tasks to be processed by the parallel machines. Each task is characterized by its due date  $d_i$ , processing time  $p_i$ , and the quantity  $q_i$  of resource required by the machine to process the task. The resource is supplied in the fixed quantity  $f$  at each resource release time  $\delta$ . The resource supplied can only be used within the resource available time  $\emptyset$ , because, after  $\emptyset$  time, the resource will be cold and will not meet the temperature requirements. The goal of this article is to find the optimal solution for task assignments which considers

the resource constraints (resource release time and resource available time for use) and aims at minimization of total tardiness.

Before formulating the mathematical model for this problem, the assumptions of this study are clarified as follows:

- (1) Every machine can only process one task at any time and can be used if it is available during the whole production.
- (2) Resource is released and consumed at the beginning of each resource released cycle.
- (3) The preparation time is not necessary for all tasks.
- (4) The completion times of tasks are unknown. The given information is only each task's processing time, amount of resource required, and due date.
- (5) All tasks are independent and must be processed once without interruption.

### 3.2. Mathematical Model

#### Notation

0: virtual starting artifacts

$d_i$ : the due date of task  $i$

$f$ : the fixed amount of the resource supplied each time

$i$ : the index of a task,  $i = 1, 2, \dots, |N|$ , where  $|N|$  is the total number of tasks

$k$ : the index of a resource released time; there are  $|K|$  times of resource release

$m$ : the index of a machine,  $m = 1, 2, \dots, |Q|$ , where  $|Q|$  is the total number of machines

$M$ : a large positive number

$n + 1$ : virtual ending artifacts, where  $n$  is the number of tasks

$N^- N \cup \{0\}$

$N^+ N \cup \{n + 1\}$

$\overline{N} N^- \cup N^+$

$p_i$ : the processing time of task  $i$

$q_i$ : the resource required by task  $i$

$\delta$ : the resource release cycle

$\emptyset$ : the resource availability time for use

#### Decision Variables

$C_i$ : the completion time of task  $i$

$T_i$ : the tardiness of task  $i$

$X_{ijm}$ : a binary variable, which equals 1 if task  $j$  is scheduled to be processed immediately after task  $i$  on machine  $m$ ; it is 0, otherwise

$Y_{ikm}$ : a binary variable, which equals 1 if task  $i$  is scheduled to be processed at the  $k$ th resource release time on machine  $m$ ; it is 0, otherwise

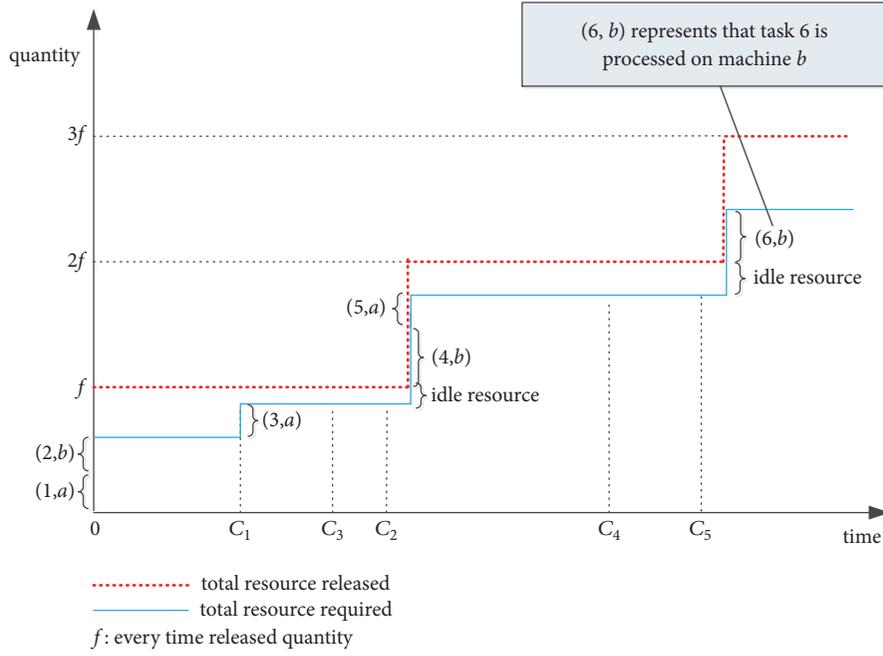


FIGURE 2: Resource released time constraints in continuous-casting production.

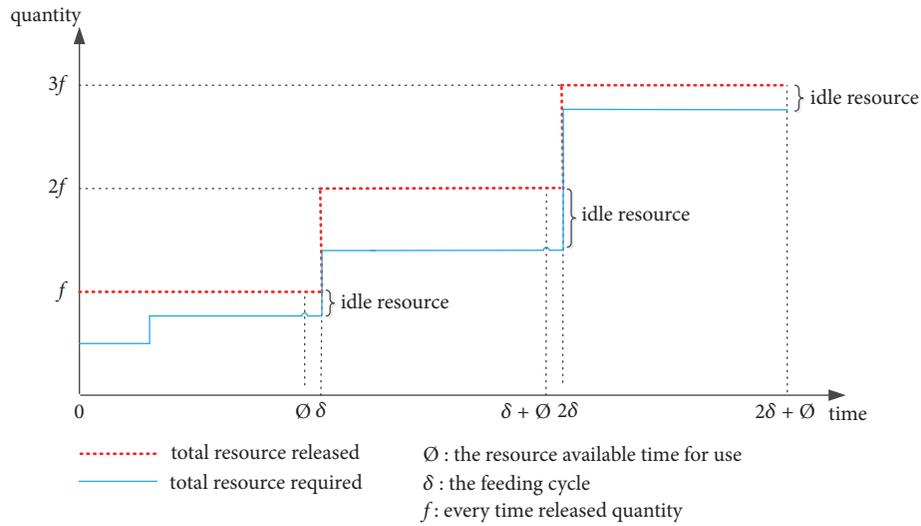


FIGURE 3: Resource available time constraints in continuous-casting production.

The problem can be formulated as the following mathematical model:

$$[M1] \min \sum_{i \in N} T_i \quad (1)$$

subject to

$$\sum_{m \in Q} \sum_{i \in N^-(i \neq j)} X_{ijm} = 1 \quad j \in N \quad (2)$$

$$\sum_{m \in Q} \sum_{j \in N^+(i \neq j)} X_{ijm} = 1 \quad i \in N \quad (3)$$

$$C_j > C_i + p_j - (1 - X_{ijm})M \quad (4)$$

$$i \in N^-, j \in N^+ (i \neq j), m \in Q$$

$$T_i \geq C_i - d_i \quad i \in \bar{N} \quad (5)$$

$$f \geq \sum_{m \in Q} \sum_{i \in N} q_i Y_{ikm} \quad k \in K \quad (6)$$

$$\sum_{m \in Q} \sum_{k \in K} Y_{ikm} = 1 \quad i \in N \quad (7)$$

$$\sum_{k' \in K(k' \geq k)} Y_{jk'm} \geq X_{ijm} Y_{ikm} \quad i \in N^-, j \in N^+ (i \neq j), k \in K, m \in Q \quad (8)$$

$$\sum_{k \in K} Y_{ikm} \geq X_{ijm} \quad i \in N^-, j \in N^+ (i \neq j), m \in Q \quad (9)$$

$$\sum_{k \in K} Y_{jkm} \geq X_{ijm} \quad i \in N^-, j \in N^+ (i \neq j), m \in Q \quad (10)$$

$$\sum_{j \in N^-(i \neq j)} X_{jim} - \sum_{j \in N^+(i \neq j)} X_{ijm} = 0 \quad i \in N, m \in Q \quad (9)$$

$$\sum_{j \in N^+} X_{0jm} = 1 \quad m \in Q \quad (10)$$

$$\sum_{i \in N^-} X_{i(n+1)m} = 1 \quad m \in Q \quad (11)$$

$$C_i - p_i \geq \delta \sum_{m \in Q} \sum_{k \in K} k Y_{ikm} \quad i \in N \quad (12)$$

$$C_i \leq \delta \sum_{m \in Q} \sum_{k \in K} k Y_{ikm} + \theta \quad i \in N \quad (13)$$

$$X_{ijm} \in \{0, 1\} \quad i \in N^-, j \in N^+, m \in Q \quad (14)$$

$$Y_{ikm} \in \{0, 1\} \quad i \in \bar{N}, k \in K, m \in Q \quad (15)$$

$$T_i, C_i \geq 0 \quad i \in \bar{N} \quad (16)$$

Objective (1) is to minimize the total tardiness of all tasks. Constraints (2) and (3) ensure that tasks are finished one by one. Constraints (4) state that if task  $j$  is processed after task  $i$  immediately on machine  $m$ , the beginning processing time of task  $j$  is later than the completion time of task  $i$ . Constraints (5) restrict the tardiness of each task  $i$ . Constraints (6) are nonlinear supply resource constraints and ensure that the total resource supplied at any time is not less than the resource required in each resource release time. Constraints (7) ensure that each task is fulfilled exactly once. Constraints (8) ensure that if task  $j$  is processed after task  $i$  immediately on machine  $m$ , the index of task  $j$ 's resource release time is not less than that of task  $i$ . It is noted that Constraints (8) are nonlinear constraints and will be transformed to the linear form. Constraints (9) are the flow conservation constraint. Constraints (10) and (11) ensure that each machine begins to work with virtual task 0 and ends with virtual task  $n + 1$ . Constraints (12) and (13) ensure that if task  $i$  is processed at the  $k$ th resource released time, the starting processing time of task  $i$  is later than this resource release time, and it should be processed within the resource available time for use  $\theta$ . Constraints (14) to (16) restrict the domain of the decision variables.

**3.3. Linearization of Constraints (8).** Constraints (8) contain a nonlinear part  $X_{ijm} Y_{ikm}$ , which is the product of two binary variables. We define a new binary variable  $\varphi_{ijkm}$  to replace the nonlinear part.

*Newly Defined Variables.*  $\varphi_{ijkm}$  binary variable is equal to one if and only if task  $j$  is processed after task  $i$  immediately on machine  $m$ , and the index of task  $i$ 's resource release time is  $k$ .

Constraints (8) become

$$\sum_{k' \in K(k' \geq k)} Y_{jk'm} \geq \varphi_{ijkm} \quad (17)$$

$$i \in N^-, j \in N^+ (i \neq j), k \in K, m \in Q$$

In addition, some more constraints need to be defined so that the newly defined variable  $\varphi_{ijkm}$  can replace the function of  $X_{ijm} Y_{ikm}$ .

$$\varphi_{ijkm} \geq X_{ijm} + Y_{ikm} - 1 \quad (18)$$

$$i \in N^-, j \in N^+ (i \neq j), k \in K, m \in Q$$

$$\varphi_{ijkm} \leq X_{ijm} \quad (19)$$

$$i \in N^-, j \in N^+ (i \neq j), k \in K, m \in Q$$

$$\varphi_{ijkm} \leq Y_{ikm} \quad (20)$$

$$i \in N^-, j \in N^+ (i \neq j), k \in K, m \in Q$$

$$\varphi_{ijkm} \in \{0, 1\} \quad (21)$$

$$i \in N^-, j \in N^+ (i \neq j), k \in K, m \in Q$$

Constraints (17) are linear and now the following [M2] is a MIP model, which can be solved by off-the-shelf solvers such as CPLEX.

[M2] An MIP model: Objective (1), subject to Constraints (2)–(7) and (9)–(21).

## 4. Solution Approach

**4.1. PSO Algorithm.** For some small-scale instances, the proposed model can be solved by using CPLEX solver directly. However, large-scale instances cannot be solved by CPLEX solver within a reasonable time, or can lead to an OutOfMemory error.

Particle swarm optimization (PSO) is a population-based optimization algorithm, where each particle has a position and velocity representing a solution. The position reflects the quality of the solution, and the velocity determines where the particle will move in the next iteration. The ordinary updating formulae for velocities and positions are as follows:

$$v_{ik}^{t+1} = v_{ik}^t + c_1 r_1 (x_{pbest_{ik}}^t - x_{ik}^t) + c_2 r_2 (x_{gbest_{ik}}^t - x_{ik}^t) \quad (22)$$

$$x_{ik}^{t+1} = x_{ik}^t + v_{ik}^t \quad (23)$$

where  $x_{ik}^{t+1}$  and  $x_{ik}^t$  represent the current and previous positions of particle  $i$  in dimension  $k$ , respectively;  $v_{ik}^{t+1}$  and  $v_{ik}^t$  represent the current and previous velocity of particle  $i$  in dimension  $k$ , respectively; the weights  $c_1$  and  $c_2$  are acceleration constants which are generally viewed as the cognitive

```

i ← 1
k ← maxEpoch
while i ≤ k do
    decode the particle to generate genes, obtain the reduced MIP and solve it by using CPLEX
    compare and update each swarm's best position and particle's best position
    update the velocity and position accordingly
    i ← i + 1
end while

```

ALGORITHM 1: PSO algorithm main steps.

acceleration coefficient and social acceleration coefficient, respectively;  $r_1$  and  $r_2$  are random numbers in  $[0, 1]$ ;  $x_{pbest_{ik}}^t$  is the best position of particle  $i$  in dimension  $k$  up to iteration  $t$ ; and  $x_{gbest_{ik}}^t$  is the best position of the whole swarm in dimension  $k$  up to iteration  $t$ .

The pseudocode of our PSO algorithm is shown in Algorithm 1. The approach starts by generating an initial solution, which can be any sequence of tasks. It then initializes an iteration counter and repeats the same setup until the iteration counter reaches a given limit. In each iteration, the approach selects the next neighborhood operator from a randomly generated sequence of operators. For each operator, all possible moves are examined. If an improved solution is found, this solution is accepted as the new incumbent solution. If no improvement can be obtained by any of the operators, a locally optimal solution is found. The search process is repeated in order to be able to escape from a local optimum. Before continuing with the next iteration, the incumbent solution is perturbed in order to obtain a new solution that potentially can be improved using the operators.

Apart from the main steps, this article also designs a new encoding and decoding method to improve solution quality. This encoding and decoding method is specific to  $X_{ijm}$  and is illustrated as follows. As Figure 4 shows, we have 5 tasks, and the number of available machines is 3. We first generate 5 random numbers in  $[0, 1]$  as genes, then use the value of genes times the total number of machines, round down fractions, and obtain each task's processing machine. Then, we calculate the due date for each task and determine its urgency degree based on its due date. The task with an earlier due date is prior. We compare the urgency degrees of all tasks distributed to the same machine, and tasks are discharged from the processing sequence according to their degree of urgency. For example, as in Figure 4, tasks 2, 3, and 4 are arranged to be processed on machine 2, task 1 is arranged to machine 0, and task 5 is arranged to machine 1. Then, as Table 1 shows, each machine's processing order is arranged according to the degree of urgency. So, just as Table 2 shows, the processing order on machine 0 is task 1, the order on machine 1 is task 5, and the order on machine 2 is tasks 3, 2, and 4. Therefore,  $X_{322} = X_{242} = 1$ .

**4.2. Algorithm Improvement.** The neighborhood structure is a key component of the algorithmic framework, and neighborhood search has an effect on the performance of heuristic algorithm. Therefore, we design an improved neighborhood

search method for the PSO algorithm in order to improve solving ability. According to the experimental analysis of the example, we find that when CPLEX solves this problem directly, the time that the Gap value (percent) drops from 100% to 10% is very fast. In most cases, almost 80% of the whole computation time is spent on dropping the Gap value from 10% to 0. Therefore, it may need to take a long time to find the neighborhood of the optimal solution. And if the local search of the PSO algorithm does not find an improved solution, we employ another local search method to find it again. We use the behavior of bees in nature to explain the core idea of our improved algorithm more easily. During the searching process, a food source represents a candidate solution and the best food source found by bees is the model's optimal solution. The initial solution  $x_i (i = 1, 2, \dots, SN, SN$  is the number of nectar source) of swarm is a vector of  $D$  dimension, where  $D$  is the number of the parameters to be optimized. Each  $x_i$  has  $n$  variables ( $x_{id}, d = 1, 2, \dots, n$ ) and these variables will all be optimized to make the problem's objective minimal. So all the bees' first solution is generated by formula (24). In later iterations, scouts move to other solutions according to formula (25). And greedy selection is applied between the old solution and the new solution to find a better solution. After that, onlooker bees go out to exploit better solutions in their surroundings. Roulette wheel selection is applied to create further generation of new solutions. New solutions are updated by formula (26). When the number of iterations reaches the limit, the current global optimal solution is taken as the solution, and solution fitness is defined as the ABC algorithm's solution tardiness.

$$X_{id}^0 = L_d + r(U_d - L_d) \quad (24)$$

where  $X_{id}$  is the  $i$ th potential solution in the swarm,  $U_d$  and  $L_d$  are bounds of  $X_i$  in the  $d$ th dimension, and  $r$  is a random number in the interval  $[0, 1]$ .

$$new\_X_i^j = X_i^j + \varphi(X_i^j - X_k^j) \quad (25)$$

where  $i$  is the current solution,  $k (k \neq i)$  is a randomly selected solution among the  $SN$  solutions, and  $\varphi$  is a random number in the interval  $[-1, 1]$ .

$$X_{id}^t = L_d + r(U_d - L_d) \quad (26)$$

where  $X_{id}^t$  is the  $t$ th iteration solutions,  $U_d$  and  $L_d$  are bounds of  $X_i$  in the  $d$ th dimension, and  $r$  is a random number in the interval  $[0, 1]$ .

		index of tasks				
		1	2	3	4	5
genes		0.2	0.9	0.8	0.7	0.6
	genes $\times$ number of machines	0.6	2.7	2.4	2.1	1.8
	machine(round down)	0	2	2	2	1

FIGURE 4: The encoding and decoding method.

TABLE 1: An example of urgency degree.

Index of tasks	Due date	Processed machine	Urgency degree
1	15	0	1
2	14	2	2
3	8	2	1
4	20	2	3
5	16	1	1

TABLE 2: An example of decoding results.

Index of machines	Task processing order
0	1
1	5
2	3-2-4

This local search method searches iteratively in a feasible neighborhood till an upper limit on time, and then it starts to search in another feasible neighborhood. So, too much time is spent doing local search, which leads to poor results. Therefore, this article conjectures that the improved local search method is more suitable for a local search to solve the proposed model. Considering the global search ability of the particle swarm algorithm, and the local search ability of the abovementioned method, this article proposes an improved algorithm, called PSOB, to improve the existing algorithm, accelerate the resolving speed, improve the resolving ability, and generate better solutions.

The pseudocode of the PSOB algorithm is shown in Algorithm 2. In order to find the optimal neighborhood near the optimal solution as soon as possible, the PSOB algorithm searches the solution globally quickly like the PSO algorithm. Then, it uses the abovementioned local search algorithm to do a local search. The PSOB algorithm's solving process is as follows. When the first part of the solving process is finished, the second part's initial global optimal position is set as the current GlobalMin and GlobalParas[ $i$ ]. Then, the abovementioned local search algorithm is iterated. After maxCycle times, the final result is the PSOB algorithm's optimal solution.

TABLE 3: Parameter values of the model.

Parameter	Value	Parameter	Value	Parameter	Value
$d_i$	$U(10,60)$	$f$	15	$\delta$	25
$p_i$	$U(2,15)$	$q_i$	$U(2,7)$	$\theta$	23

TABLE 4: Parameter values of PSO algorithm and the proposed algorithm.

Parameter	Value
<i>max iteration</i>	15
<i>tabu list size</i>	12
<i>N size</i>	10

## 5. Computational Experiments

This section presents the computational experiments conducted to evaluate the proposed algorithm. Several experiments are performed on a PC (Intel Core i7, 2.6 G Hz; Memory, 8 G). Each algorithm based heuristic is coded by using C#. In addition, Tables 3 and 4 list the parameter values of the proposed model, and PSO algorithm and the proposed algorithm, respectively.

The quality of solutions is affected by the number of machines  $m$  and the number of tasks  $i$ . Tables 5 and 6 summarize the small-scale instances and large-scale instances' computational experiments solved by CPLEX, the PSO algorithm method, and the PSOB algorithm method in terms of the average computation time (CPU) and the solution quality. Each set has ten randomly generated cases and we record their corresponding average values.

```

s ← initialsolution()
i ← 1, N ← 20, k ← 30,
while i ≤ k do
  repeat
    select neighborhood operator φ
    for all solutions s' that can be generated by applying φ to s do
      if tardiness(s') < tardiness(s) then
        s ← s'
      end if
    end for
  until s is locally optimal for all operators
  s ← update()
  i ← i+1
end while
BestSolution ← GlobalMin, Solution[s] ← GlobalParas[i]
initialsolution(): iter ← 1, N ← 20, limit ← 30, i ← 1, maxCycle ← 20
k ← maxEpoch
while iter ≤ limit do
  scouts: decode the particle to generate genes, obtain the reduced MIP and solve it by
  CPLEX
  calculate fitness
  generate employed bees and onlooker bees: trail(i) ← 0
  select neighborhood operator φ
  for all solutions s' that can be generated by applying φ to s do
    if tardiness(s') < tardiness(s) then
      s ← s'
      trial(i)=0
    else
      trial(i)=trail(i)+1
    end if
    calculate Probability(i)
    if Probability(i) > rand()
      employed bees: select neighborhood operator
      for all solutions s' that can be generated by applying φ to s do
        if tardiness(s') < tardiness(s) then
          s ← s'
          trial(i) ← 0
        else
          trial(i) ← trail(i)+1
        end if
      end if
    if trail(i) > limit
      scouts: generate the new food source
      update BestSolution
      iter ← iter+1
    end if
    if iter > maxCycle
      find BestSolution
    else
      generate employed bees and repeat
    end if
  end while
end while

```

ALGORITHM 2: PSOB algorithm main steps.

As Tables 5 and 6 show, the average computation time required by CPLEX increases sharply with an increase in the number of tasks. For the instance with 2 machines, the 20-task case is already very difficult to solve; the instances with 3 machines and 4 machines are very difficult to solve with

16 tasks and 18 tasks, respectively. Analyzing Tables 5 and 6, the total number of instances is 27, and the solutions of PSOB algorithm are closer to the exact solutions obtained by CPLEX than that of PSO algorithm in all 27 instances, including 7 instances with equal solutions. Besides, the computation

TABLE 5: Computational experiments of solving small-scale instances by CPLEX, PSO, and PSOB algorithms.

Number Of Machines	Number Of Tasks	CPLEX		PSO		PSOB	
		Best Solution	CPU(s)	Best Solution	CPU(s)	Best Solution	CPU(s)
2	6	1.2	0.1	1.2	17.3	1.2	17.2
	8	17.3	9.9	18.5	25.4	18.5	21.8
	10	43.5	9.0	51.5	38.2	43.5	27.4
	12	83.5	124.3	95.9	44.9	90.8	39.3
	14	118.0	702.1	151.5	49.8	143.2	52.3
Average		53.0	169.0	64.0	35.0	59.0	32.0
3	8	10.7	2.5	10.7	68.3	10.7	43.6
	10	39.1	6.6	39.1	171.9	39.1	84.8
	12	76.2	42.6	76.5	359.8	76.3	344.9
	14	101.0	149.7	149.9	2191.7	101.0	1067.7
	16	318.3	424.3	329.7	2658.7	321.0	1499.1
Average		109.1	125.1	121.2	1090.1	109.6	608.0
4	10	39.0	9.9	39.0	222.5	39.0	199.0
	12	75.9	56.5	75.9	765.6	75.9	759.1
	14	116.0	206.8	116.8	1476.6	116.0	1376.6
	16	318.3	792.7	323.3	2462.7	318.3	2162.7
	18	324.8	2036.0	331.3	2316.0	330.0	2176.8
Average		174.8	620.4	177.3	1448.7	175.8	1334.8

TABLE 6: Computational experiments of solving large-scale instances by CPLEX, PSO, and PSOB algorithms.

Number Of Machines	Number Of Tasks	CPLEX		PSO		PSOB	
		Best Solution	CPU(s)	Best Solution	CPU(s)	Best Solution	CPU(s)
2	20	575	3600	598.0	105.0	584.0	104.0
	22	643727	3600	867.7	116.3	820.3	469.0
	25	607637	3600	1345.0	202.3	894.3	263.3
	30	16612880	3600	2218.0	347.3	1990.0	337.3
Average		-	3600	1257.2	192.7	1072.2	293.4
3	20	382432	3600	465.0	3042.0	410.7	2719.0
	22	629671	3600	708.7	2965.0	697.8	2778.9
	25	10491201	3600	1128.0	3234.3	1101.3	3008.9
	30	-	3600	2010.7	3517.0	1050.3	3379.1
Average		-	3600	1078.1	3189.6	815.0	2971.5
4	20	390400	3600	463.0	3354.0	424.0	3155.2
	22	635660	3600	689.7	3248.7	689.7	3145.4
	25	10341439	3600	1175.0	3440.0	1095.0	3456.0
	30	-	3600	3972.0	3549.2	2477.4	3600.0
Average		-	3600	1574.9	3398.0	1171.5	3339.2

Note. The en dash means we did not find any solution within the time limits. Two values of best solutions solved by CPLEX are the lower bound and upper bound of the corresponding instance.

time by the PSO algorithm is longer than that of the PSOB algorithm. Therefore, the computational experiments show that the PSOB algorithm outperforms the PSO algorithm in terms of both solution quality and computation time.

In total, the heuristic algorithms are able to find equally good solutions for small-scale instances and can find better

solutions than the CPLEX solver within a reasonable time. Tables 5 and 6 show that the PSOB algorithm's solutions are closer to CPLEX's solution than those of the PSO algorithm. Therefore, this article argues that, in the continuous-casting stage of iron and steel enterprises, the distribution of casting machines and casting molten steel tundish processing order

problems based on upstream resource constraints are more suitable to be solved by the PSOB algorithm for large-scale problems.

## 6. Conclusions

The steelmaking and continuous-casting problems have been extensively studied in literature and widely adopted in industry. However, most studies only consider a single machine which has a linear resource supplied. In practice, the resource supply rate and consumption rate are nonlinear, and the available time for use of supplied resource is limited. Accordingly, we assume that the upstream resource is released at regular intervals and has limited time for use. The release and consumption of the resource happen at the beginning of each resource release cycle.

This article studies task assignment for continuous-casting production in integrated iron and steel enterprises. This problem differs from other similar problems without resource constraints because the optimal solution can be influenced by the resource constraints. The distribution of casting machine and casting molten steel tundish processing order problems are important planning problems for daily operational management in iron and steel enterprises. In order to deal with these problems efficiently, we present a mathematical model minimizing total tardiness and an improved heuristic algorithm to generate solutions automatically. We design a new encoding and decoding method and use the CPLEX, PSO algorithm, and proposed PSOB algorithm to solve the problem. From the computational experiments, we can find that the PSOB algorithm is able to find equally good solutions for small-scale instances and can find better solutions than the CPLEX solver and the PSO algorithm within a reasonable time.

Throughout this article, we assume that the resource is supplied at regular intervals and has a limited time for use. While these assumptions are essential and practical, the heuristic approach can easily be adapted to be used for different scenarios with arbitrary resource constraints. Therefore, the proposed approach can be used for a variety of industry production processes in which schedulers must take into account resource constraints.

## Data Availability

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

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China [grant numbers 71771143 and 71771180], Innovation Program of Shanghai Municipal Education Commission [No. 14YZ111], Shanghai Young Eastern Scholar

Programme [QD2015041], Shanghai Pu Jiang Program (No. 13PJ066), and Human Arts & Social Science Foundation of Ministry of Education of China (No. 16YJC630112).

## References

- [1] L. Chen, B. Yang, X. Shen, Z. Xie, and F. Sun, "Thermodynamic optimization opportunities for the recovery and utilization of residual energy and heat in China's iron and steel industry: A case study," *Applied Thermal Engineering*, vol. 86, pp. 151–160, 2015.
- [2] J. Mori and V. Mahalec, "Planning and scheduling of steel plates production. Part II: scheduling of continuous casting," *Computers & Chemical Engineering*, vol. 6, pp. 438–446, 2016.
- [3] Q.-K. Pan, "An effective co-evolutionary artificial bee colony algorithm for steelmaking-continuous casting scheduling," *European Journal of Operational Research*, vol. 250, no. 3, pp. 702–714, 2016.
- [4] L. Tang, J. Liu, A. Rong, and Z. Yang, "Mathematical programming model for scheduling steelmaking-continuous casting production," *European Journal of Operational Research*, vol. 120, no. 2, pp. 423–435, 2000.
- [5] L. Tang, P. B. Luh, J. Liu, and L. Fang, "Steel-making process scheduling using Lagrangian relaxation," *International Journal of Production Research*, vol. 40, no. 1, pp. 55–70, 2002.
- [6] L. X. Tang, Y. Zhao, and J. Y. Liu, "An improved differential evolution algorithm for practical dynamic scheduling in steel-making-continuous casting production," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 2, pp. 209–225, 2014.
- [7] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Research Logistics Quarterly*, vol. 2, pp. 83–97, 1955.
- [8] R. K. Cheung, D. D. Hang, and N. Shi, "A labeling method for dynamic driver-task assignment with uncertain task durations," *Operations Research Letters*, vol. 33, no. 4, pp. 411–420, 2005.
- [9] S. Yan and C.-H. Tang, "A heuristic approach for airport gate assignments for stochastic flight delays," *European Journal of Operational Research*, vol. 180, no. 2, pp. 547–567, 2007.
- [10] L. Zhen, "Task assignment under uncertainty: Stochastic programming and robust optimisation approaches," *International Journal of Production Research*, vol. 53, no. 5, pp. 1487–1502, 2015.
- [11] L. Zhen, D. Zhuge, and S.-L. Zhu, "Production stage allocation problem in large corporations," *OMEGA - The International Journal of Management Science*, vol. 73, pp. 60–78, 2017.
- [12] C.-F. Chien and C.-H. Chen, "Using genetic algorithms (GA) and a coloured timed Petri net (CTPN) for modelling the optimization-based schedule generator of a generic production scheduling system," *International Journal of Production Research*, vol. 45, no. 8, pp. 1763–1789, 2007.
- [13] W. Kaczmarczyk, "Proportional lot-sizing and scheduling problem with identical parallel machines," *International Journal of Production Research*, vol. 49, no. 9, pp. 2605–2623, 2011.
- [14] D. Shabtay and M. Kaspi, "Parallel machine scheduling with a convex resource consumption function," *European Journal of Operational Research*, vol. 173, no. 1, pp. 92–107, 2006.
- [15] T. C. E. Cheng and C. C. S. Sin, "A state-of-the-art review of parallel-machine scheduling research," *European Journal of Operational Research*, vol. 47, no. 3, pp. 271–292, 1990.
- [16] A. Allahverdi, J. N. D. Gupta, and T. Aldowaisan, "A review of scheduling research involving setup considerations," *Omega*, vol. 27, no. 2, pp. 219–239, 1999.

- [17] E. Mokotoff, "Parallel machine scheduling problems: a survey," *Asia-Pacific Journal of Operational Research*, vol. 18, no. 2, pp. 193–242, 2001.
- [18] M. Pfund, J. W. Fowler, and J. N. D. Gupta, "A survey of algorithms for single and multi-objective unrelated parallel-machine deterministic scheduling problems," *Journal of the Chinese Institute of Engineers*, vol. 21, no. 3, pp. 230–241, 2004.
- [19] T. Sen, J. M. Sulek, and P. Dileepan, "Static scheduling research to minimize weighted and unweighted tardiness: a state-of-the-art survey," *International Journal of Production Economics*, vol. 83, no. 1, pp. 1–12, 2003.
- [20] C. Koulamas, "The single-machine total tardiness scheduling problem: review and extensions," *European Journal of Operational Research*, vol. 202, no. 1, pp. 1–7, 2010.
- [21] H. Emmons, "One-machine sequencing to minimize certain functions of job tardiness," *Operations Research*, vol. 17, no. 4, pp. 701–715, 1969.
- [22] J. Du and J. Y. Leung, "Minimizing total tardiness on one machine is NP-hard," *Mathematics of Operations Research*, vol. 15, no. 3, pp. 483–495, 1990.
- [23] E. L. Lawler, "A pseudopolynomial algorithm for sequencing jobs to minimize total tardiness," *Annals of Discrete Mathematics*, vol. 1, no. 8, pp. 331–342, 1977.
- [24] S.-O. Shim and Y.-D. Kim, "Scheduling on parallel identical machines to minimize total tardiness," *European Journal of Operational Research*, vol. 177, no. 1, pp. 135–146, 2007.
- [25] J.-Y. Lee and Y.-D. Kim, "A branch and bound algorithm to minimize total tardiness of jobs in a two identical-parallel-machine scheduling problem with a machine availability constraint," *Journal of the Operational Research Society*, vol. 66, no. 9, pp. 1542–1554, 2015.
- [26] J. Blazewicz, J. K. Lenstra, and A. H. Rinnooy Kan, "Scheduling subject to resource constraints: classification and complexity," *Discrete Applied Mathematics*, vol. 5, no. 1, pp. 11–24, 1983.
- [27] J. Blazewicz, K. H. Ecker, E. Pesch, G. Schmidt, and J. Weglarz, *Handbook on Scheduling*, Springer, 2007.
- [28] D. Briskorn, B.-C. Choi, K. Lee, J. Leung, and M. Pinedo, "Complexity of single machine scheduling subject to nonnegative inventory constraints," *European Journal of Operational Research*, vol. 207, no. 2, pp. 605–619, 2010.
- [29] D. Briskorn, F. Jaehn, and E. Pesch, "Exact algorithms for inventory constrained scheduling on a single machine," *Journal of Scheduling*, vol. 16, no. 1, pp. 105–115, 2013.
- [30] P. Györgyi and T. Kis, "Approximability of scheduling problems with resource consuming jobs," *Annals of Operations Research*, vol. 235, pp. 319–336, 2015.
- [31] S. Wang and M. Liu, "Multi-objective optimization of parallel machine scheduling integrated with multi-resources preventive maintenance planning," *Journal of Manufacturing Systems*, vol. 37, pp. 182–192, 2015.

## Research Article

# Inventory and Production Dynamics in a Discrete-Time Vendor-Managed Inventory Supply Chain System

Yongchang Wei <sup>1</sup>, Fangyu Chen <sup>1</sup>, and Hongwei Wang <sup>2,3</sup>

<sup>1</sup>School of Business Administration, Zhongnan University of Economics and Law, Wuhan, China

<sup>2</sup>School of Management, Huazhong University of Science and Technology, Wuhan, China

<sup>3</sup>School of Automation, Huazhong University of Science and Technology, Wuhan, China

Correspondence should be addressed to Hongwei Wang; [hwwang@hust.edu.cn](mailto:hwwang@hust.edu.cn)

Received 18 June 2018; Accepted 1 September 2018; Published 19 September 2018

Academic Editor: Lu Zhen

Copyright © 2018 Yongchang Wei 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 presents some analytical results on production and order dynamics in the context of a discrete-time VMI supply chain system composed of one retailer and one manufacturer. We firstly derive the lower bound and upper bound on the range of inventory fluctuations for the retailer under unknown demand. We prove that the production fluctuations can be interestingly smoothed and stabilized independent of the delivery frequency of the manufacturer used to satisfy the retailer's demand, even if the retailer subsystem is unstable. The sufficient and necessary stability condition for the whole supply chain system is obtained. To further explore the production fluctuation problem, the bullwhip effect under unknown demand is explored based on a transfer function model with the purpose of disclosing the influences of parameters on production fluctuations. Finally, simulation experiments are used to validate the theoretical results with respect to inventory and production fluctuations.

## 1. Introduction

Vendor-managed inventory (VMI) is a well-known collaborative program [1–3], in which a retailer propagates inventory information and point of sales data to its supplier, while the upstream supplier makes replenishment decision for both the retailer and itself. Since it was pioneered by Wal-Mart in the early 1980s [4], VMI has been successfully implemented by many leading firms from different industries, such as Glaxosmithkline [5], Nestle and Tesco [6], Boeing and Alcoa [7], and Shell Chemical [8].

The benefits of VMI for both suppliers and retailers have been extensively studied in the literature [9–11]. For vendors, the VMI program improves decision-making flexibility since they gain the right to determine the amount and timing of the order for retailers, which further leads to reduced bullwhip effect [12] and better utilization of manufacturing capacity, as well as better synchronization of replenishment planning [13]. For retailers, VMI saves ordering cost and inventory cost and ensures the customer service level through signing VMI contracts. For example, a supplier will be penalized

once the inventory of retailer exceeds certain ranges [14]. To successfully implement the VMI program, how to make use of the flexibility of decision-making in inventory control to benefit both upstream and downstream members is a highly challenging problem.

Most of the existing literature has focused on game theory or optimization models [15–19], while this paper explores the impact of replenishment parameters on the inventory and production dynamics [20], which are closely related to the total costs. Significant costs can be incurred by production fluctuations due to frequently switching production quantities up and down, which further complicate the activities in labor force arrangement, capacity adjustment, and equipment maintenance and management. However, in a dynamic system with uncertain demand, deriving optimal solutions from an optimization model incorporating all these cost components is impractical. Furthermore, in practice, through information sharing even with order batching [21–24], VMI has become one of the frequently employed programs for eliminating bullwhip effect [23, 25, 26]. This implies that production can be smoothed and transportation cost can be

saved via the implementation of VMI program. However, we notice that the majority of existing bullwhip literature has ignored the interaction between inventory and production fluctuations.

In a VMI system, the upstream supplier has the flexibility to determine the amount and timing of the replenishment order for the retailer. However, as mentioned, the supplier might be penalized once the retailer's inventory exceeds the predetermined range. Therefore, in order to reduce penalty cost, the supplier should know how the replenishment parameters for the retailer affect the range of inventory fluctuations of the retailer. In addition, the supplier should also consider the production or ordering cost and inventory cost for the whole system when designing its own replenishment policies. In this paper, we aim to answer three questions for the implementation of VMI: (1) How the replenishment rules designed by the supplier for itself and retailer affect the stability, inventory fluctuations, and production smoothing for the entire system under uncertain demand? (2) What is the trade-off between inventory fluctuations and production fluctuations? Specifically, what actions can be taken to limit inventory and order fluctuations. This question will be addressed in terms of stability. As investigated in the literature, stability is a fundamental problem for any dynamical system including supply chain systems [21, 27, 28]. After being disrupted, the state of a stable system will return to its steady state gradually. Actually, a stable supply chain system tends to be resilient in the presence of demand disruptions [29]. Furthermore, stability analysis simplifies the parameters selection and performance optimization. For a stable supply chain system, stability limits the fluctuation of inventory and order in a certain range after the sudden change of demand. By contrast, unstable designs will cause undesirable fluctuations for order and inventory, which leads to high cost due to inventory accumulation or stockouts. (3) How the replenishment parameters affect the magnitude of production smoothing for the manufacturer.

Specifically, we attempt to explore the inventory and production fluctuations in the context of a VMI supply chain system composed of a manufacturer and a retailer. The manufacturer uses a reorder point policy to manage the retailer's inventory and the APIOBPCS (automatic pipeline, inventory, and order based production control system) for production control [12, 30]. The contributions of this paper lie in threefold. Firstly, we analytically studied the impact of reorder point, delivery frequency, lead time, and demand characteristics on the range of the retailer's inventory under uncertain demand by extending a more generalized replenishment policy. Secondly, we derived the necessary and sufficient stability conditions of the VMI system for arbitrary lead times by considering the interaction between the retailer subsystem and the manufacturer subsystem, which are significant due to the development of global supply chains and complicated manufacturing process [31]. We prove that production fluctuations are independent of the replenishment parameters in the reorder point policy. Besides, we demonstrate that although the production of the upstream manufacturer can be perfectly smoothed without paying extra attention to the quantity and frequency of

replenishment for the retailer, the parameters in the reorder point policy have significant influences on the inventory fluctuations for both retailer and manufacturer. Finally, to further quantify the production smoothing magnitude, the impact of all the decision parameters on the robust bullwhip metric is further studied based on a transfer function model. The results of this research provide insightful guidelines on the implementation of VMI program.

The remainder of this paper is organized as follows. Section 2 introduces related literature. Section 3 describes the VMI model in terms of the retailer subsystem and manufacturer subsystem. Section 4 studies the stability and inventory oscillations for the retailer subsystem. The production smoothing effect and inventory fluctuation for the manufacturer are studied in Section 5. Section 6 focuses on the robustness of the bullwhip effect problem under uncertain demand. This paper is concluded in Section 7.

## 2. Literature Review

Over the past decades, the advantages of VMI program have been extensively discussed. The shift of the authority of inventory control from downstream enables the supplier to choose the time and quantity for each replenishment order and also the route to transport the required goods to multiple retailers. It was well recognized that VMI brings more advantages over information sharing alone [4, 32–35]. In Fry et al. [14], a  $(z, Z)$  type contract was included in a VMI supply chain and the authors found that such a supply chain performs better than traditional supply chains. Lee et al. [3] studied the benefits and coordination problem for four business scenarios: traditional system, integrated system, VMI with stockout-cost sharing, and VMI with fixed transfer payments. Ben-Daya et al. [15] compared the benefits for three distinct supply chains, two of which were in the context of vendor-managed inventory consignment partnership. However, the supplier and retailer might belong to different companies or a vertically integrated company. The result showed that flexibility of the supplier has significant influences on the benefits of the vendor. Dong et al. [36] examined the benefits of the VMI program via empirical analysis, from which the results showed that inventory and stockout reductions might be realized at different times following the implementation of VMI. That is, inventory reduction initially may be the major benefit to distributor from VMI, while the benefits of stockout reduction may be realized after the first year. Lu et al. [9] studied the impact of two kinds of overconfidence on the inventory decision-making from a behavioral perspective. The study demonstrated behavioral factors might, to some extent, enhance the supplier's profit through continuous effort.

Although the opportunities of VMI are evident in many industries, the benefit allocation and coordination problem has been argued continuously [3, 37, 38]. A number of authors argued that the majority of the benefits from VMI tend to flow to the supplier, rather than to the retailer [32]. Lee and Cho [39] examined the problem of designing a VMI contract with consignment stock and stockout-cost sharing in a  $(Q, r)$  inventory system between a supplier

and a retailer. The result showed that the retailer may not always benefit from VMI. However, Mishra and Raghunathan [40] found that VMI intensifies the competition among manufacturers of competing brands, thus providing benefits to retailers. Pan and So [41] analyzed the interaction among the assembler and two component suppliers under a VMI contract. One supplier has uncertainty in the supply process, in which the actual number of components available for assembly is equal to a random fraction of the production quantity. Under deterministic and stochastic demand, the optimal component prices offered by the assembler are derived.

In addition to the previous literature, optimal replenishment policies of VMI supply chain systems with multiple retailers have been studied frequently over the recent years [17, 42–45]. Nevertheless, we see that the existing literature has ignored the dynamic fluctuations of order and inventory, which can be hardly incorporated into the previous optimization or game-theory based models. Although our VMI model is similar to an early work by Disney and Towill [12], we generalized their model by allowing for free adjustment of the amount and timing of replenishment frequency. Another important difference is that we pay attention to the trade-off between the production smoothing and inventory fluctuations by providing some analytical results, while their work is focused on simulation experiments.

### 3. Model Description and Analysis

**3.1. Description for the VMI System.** Consider a periodic-review, single item VMI supply chain system composed of a retailer and a manufacturer. In a VMI program, the retailer should provide the manufacturer with demand and inventory information. On the basis of these information, the manufacturer makes replenishment decision to control the inventory for both the manufacturer and the retailer.

The timing of the events in each period is assumed as follows. At the beginning of each period  $t$ , the manufacturer puts its newly finished work-in-progress (WIP) inventory into a warehouse. Assume that there is a fixed production lead time,  $L_f$ , for the manufacturer to finish an order. At the same time, the retailer also receives its replenishment order from the manufacturer to recover its inventory. It takes the manufacturer  $L_r$  periods to dispatch each replenishment order for the retailer. Note that one extra period is included in  $L_f$  and  $L_r$  to cope with the events order because the replenishment decisions are made at the beginning of each period. After putting final products into a warehouse, the manufacturer forecasts the demand for the coming period according to realized demand data. Then the manufacturer makes the production decision for itself and the replenishment decision for the retailer by reviewing the inventory information regarding the retailer's inventory level, in-transient inventory, and reorder point, as well as the manufacturer's own inventory level and WIP inventory. This is the essential difference in comparison with traditional supply chain models. To keep the model linear, we also assume that all the orders of the retailer

which can not be satisfied immediately are backlogged. Finally, during the remaining time of each period, the retailer fulfils customer demand and any unfilled orders are also backlogged.

To be more general, we assume that the demand of the retailer is unknown but bounded by  $D_t \in [d_{min}, d_{max}]$ , in which  $d_{min}$  and  $d_{max}$  represent the lower bound and upper bound, respectively. The demand uncertainty is caused by a lot of factors, such as promotions, technological innovations, seasonal changes, and economical or political events. To cope with these uncertain factors, the manufacturer uses an adaptive reorder point policy to manage the retailer's inventory. The reorder point  $R_t$  is updated by

$$R_t = R_{t-1} + \theta_r (G \times D_{t-1} - R_{t-1}), \quad (1)$$

where  $\theta_r$ ,  $0 \leq \theta_r \leq 1$ , is a coefficient to smooth the reorder point while the parameter  $G$  is used to adjust service level. We note that the reorder point satisfies  $Gd_{min} \leq R_t \leq Gd_{max}$  as long as  $Gd_{min} \leq R_1 \leq Gd_{max}$ . Once the inventory position of the retailer,  $IP_{r,t}$ , is below the reorder point, an replenishment order will be triggered automatically or manually. The amount of the retailer's replenishment order  $Q_t$  is determined by the following reorder point policy:

$$Q_t = \begin{cases} \rho D_{t-1} & \text{if } IP_{r,t} < R_t, \\ 0 & \text{if } IP_{r,t} \geq R_t, \end{cases} \quad (2)$$

in which the parameter  $\rho$  is used to control the replenishment frequency for the retailer. Note that the reorder policy in Disney and Towill [12] is a special case when  $\rho = 1$ , which imposes a constraint on the amount of each order and limits the decision-making flexibility for the manufacturer to adjust the replenishment frequency to achieve economical scale. In our model, we add this parameter  $\rho$  allowing for freely adjusting the frequency of replenishment for the retailer, which essentially affects the performance of the whole system. Actually, the retailer might determine the value for  $\rho$  according to its inventory capacity and demand characteristics. We argue that the parameter  $\rho$  plays a different role in contrast to the parameter  $G$  in (1). The parameter  $\rho$  is used to determine the replenishment frequency, while  $G$  is used to set the reorder point and balance the trade-off between inventory cost and service level. In practice, if the distance between the manufacturer and the retailer is large, the manufacturer might choose a large  $\rho$  to save transportation cost. Finally, we advocate that the parameter setting for  $G$  and  $\rho$  should consider all the cost components.

In a VMI system, a vendor makes replenishment decisions based on information sharing with respect to downstream member's demand and inventory. In our periodic inventory model, the balanced equations for the retailer's inventory level,  $IL_{r,t}$ , and in-transient inventory,  $GIT_t$ , are, respectively, represented as

$$IL_{r,t} = IL_{r,t-1} + Q_{t-L_r} - D_{t-1} \quad (3)$$

and

$$GIT_t = GIT_{t-1} - Q_{t-L_r} + Q_{t-1}. \quad (4)$$

The inventory position of the retailer is further obtained as  $IP_{r,t} = IL_{r,t} + GIT_t$ . Similarly, the difference equations for the manufacturer's inventory level and WIP are, respectively, represented as

$$IL_{m,t} = IL_{m,t-1} + P_{t-L_f} - Q_{t-1} \quad (5)$$

and

$$WIP_t = WIP_{t-1} - P_{t-L_r} + P_{t-1}. \quad (6)$$

The exponential smoothing algorithm is accurate for short-term forecasting and easy to implement [30]; thus the manufacturer uses it to forecast the demand for the coming period. It should be noted that the change of reorder point should be incorporated into the forecasting process. The forecasting algorithm is obtained as

$$F_t = F_{t-1} + \theta_f [D_{t-1} + \theta_r (G \times D_{t-1} - R_{t-1}) - F_{t-1}], \quad (7)$$

where  $\theta_f$  is the smoothing coefficient with  $0 \leq \theta_f \leq 1$ . A small  $\theta_f$  helps smoothing order fluctuations, while a large  $\theta_f$  is more appropriate to respond to volatile demand.

The replenishment rule used by the manufacturer for production decision is called APIOBPCS [30]. In traditional supply chains, the replenishment decisions for upstream members are made by themselves, whereas the replenishment decision in a VMI model should exploit systematic information. For this purpose, we define system inventory as

$$SI_t \triangleq IL_{r,t} + IL_{m,t} + GIT_t - R_t. \quad (8)$$

The replenishment rule for the manufacturer is represented as

$$P_t = F_t + \alpha^i (TINT - SI_t) + \alpha^{wip} (L_m \times F_t - WIP_t), \quad (9)$$

where  $\eta$  is the parameter to establish the target value of inventory level  $TINT = \eta \times F(t)$ ,  $L_m$  is the production lead time,  $L_m \times F_t$  is the desired value of WIP level, and  $\alpha^i$  and  $\alpha^{wip}$  are two replenishment parameters for recovering the inventory level and the WIP level. As the demand is uncertain, the target level for inventory and WIP should be updated adaptively. Essentially, the replenishment rule (9) is a proportional controller with two feedback loops. The two parameters significantly increases the flexibility to meet practical requirements. However, the feedback mechanism, the inventory interaction between the retailer and the manufacturer, coupled with the production lead time, makes the dynamics of VMI system highly complicated.

**3.2. A Transfer Function Model via Transformation of Difference Equations.** In the literature, the production smoothing behavior is usually characterized by bullwhip effect metrics

[46]. To facilitate the study of production smoothing in terms of bullwhip effect, we develop a transfer function model by treating the demand of the retailer as input and the production of the manufacturer as output. Unlike existing methods in the literature, in which transfer functions are derived by combining the components of block graphs [30], we directly obtain the transfer function by formulating a state space model and manipulating the difference equations using z-transform [47]. This change avoids the complicated computation process of the well-known Mason's gain formula [48].

From the replenishment rule (9), we obtain

$$\begin{aligned} P_t - P_{t-1} &= [1 + \alpha^i \eta + \alpha^{wip} (L_f - 1)] \times (F_t - F_{t-1}) \\ &\quad - \alpha^i (SI_t - SI_{t-1}) \\ &\quad - \alpha^{wip} (WIP_t - WIP_{t-1}). \end{aligned} \quad (10)$$

By the definition of system inventory, we can derive

$$SI_t - SI_{t-1} = P_{t-L_f} - D_{t-1} - \theta_r (GD_{t-1} - R_{t-1}). \quad (11)$$

Substituting (6) and (11) into (10), we further have

$$\begin{aligned} P_t - P_{t-1} &= [1 + \alpha^i \eta + \alpha^{wip} (L_f - 1)] \\ &\quad \cdot \theta_f [D_{t-1} + \theta_r (G \times D_{t-1} - R_{t-1}) - F_{t-1}] \\ &\quad - \alpha^i [P_{t-L_f} - D_{t-1} - \theta_r (GD_{t-1} - R_{t-1})] \\ &\quad - \alpha^{wip} (P_{t-1} - P_{t-L_f}). \end{aligned} \quad (12)$$

From (12), it is very interesting to observe that for arbitrary demand processes, the parameter  $\rho$  is irrelevant to the production dynamics. In traditional supply chains, order batching is frequently used to achieve transportation scale and reduce transportation cost. However, it is well recognized that order batching with low delivery frequency is one of the main causes of production fluctuations. In contrast to traditional supply chains, (12) demonstrates that, under the VMI program, the manufacturer can dispatch goods to the retailer with a low frequency to reduce transportation cost while not disrupting the execution of production schedules. In this sense, we conclude that transportation cost and production cost can be reduced simultaneously in a VMI system.

A state space model can be developed directly from (1), (7), and (12) by defining  $X_t \triangleq (R_t, F_t, P_t)^T$  as the state vector,  $U_t \triangleq D_t$  as input variable, and  $Y_t$  as the output variable. Define  $\psi(\alpha^i, \alpha^{wip}, L_f) \triangleq 1 + \alpha^i \eta + \alpha^{wip} (L_f - 1)$ , a state space model is obtained as

$$\begin{aligned} X_t &= AX_{t-1} + A_d X_{t-L_f} + BU_t, \\ Y_t &= CX_t, \end{aligned} \quad (13)$$

where

$$\begin{aligned}
 A &= \begin{bmatrix} 1 - \theta_r & 0 & 0 \\ -\theta_f \theta_r & 1 - \theta_f & 0 \\ -\alpha^i \theta_r - \theta_r \theta_f \times \psi(\alpha^i, \alpha^{wip}, L_f) & -\theta_f \times \psi(\alpha^i, \alpha^{wip}, L_f) & 1 - \alpha^{wip} \end{bmatrix}, \\
 A_d &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha^{wip} - \alpha^i \end{bmatrix}, \\
 B &= \begin{bmatrix} G\theta_r \\ \theta_f(1 + \theta_r G) \\ \theta_f(1 + \theta_r G) \times \psi(\alpha^i, \alpha^{wip}, L_f) + \alpha^i(1 + \theta_r G) \end{bmatrix},
 \end{aligned} \tag{14}$$

and

$$C = [0 \ 0 \ 1]. \tag{15}$$

Using the z-transform in Jury [47], the state space model (13) is converted as

$$X(z) = (I - Az^{-1} - A_d z^{-L_f}) Bz^{-1}U(z). \tag{16}$$

Finally, we can obtain the following transfer function:

$$\begin{aligned}
 G(z) &= \frac{P(z)}{D(z)} = \frac{Y(z)}{U(z)} \\
 &= C(I - Az^{-1} - A_d z^{-L_f})^{-1} Bz^{-1}.
 \end{aligned} \tag{17}$$

As mentioned in Lalwani [49], state space model (13) lays a solid foundation for studying the controllability and observability problem, and it might open a door for studying nonlinear dynamics and incorporating time-varying lead time into the APIOBPCS model, which makes the model more realistic and the results appealing. In contrast to the existing literature, the transfer function (17) in this study is formulated in a matrix form, which facilitates the implementation in computers. More importantly, the derivation process is simpler than obtaining the transfer function by combining block graphs with Mason's gain formula [48], which enables us study the production smoothing behavior in terms of worst-case bullwhip for unknown demand.

#### 4. Inventory Fluctuations of the Retailer

Before we discuss the dynamics of the VMI system, we shall firstly introduce the definition of stability. By stability, we mean that the output of a system is bounded if the input is bounded. The customer demand  $D_t$  is treated as the input. Because the fluctuations of inventory and order are of great importance, we define that the retailer subsystem is stabilized if  $IL_{r,t}$ , or  $IP_{r,t}$ , is bounded. Note that  $Q_t$  is easily bounded because customer demand is bounded by nature. The manufacturer subsystem is stable if  $IL_{m,t}$  and  $P_t$  are bounded.

Although the stability of the VMI system has been discussed in Disney and Towill [30], the obtained results are limited to small or specified lead times. Furthermore, the previous literature has ignored the interaction between the retailer and manufacturer. In particular, they failed to consider the dynamics of the retailer, which might be due to the fact that the model in Disney and Towill [30] only considers the special case  $\rho = 1$ , which simplifies the dynamics of the retailer subsystem.

In this section, we begin by studying the stability and inventory oscillation of the retailer. In the VMI program, the retailer may assume the risk of incurring high inventory cost or poor customer service level. To avoid such a problem, the retailer can increase its benefits by signing effective contractual agreements, which are characterized by the penalty cost once the retailer's inventory beyond an upper bound or a lower bound. To reduce the penalty cost and maintaining satisfactory service level for the retailer, the manufacturer must clarify how the replenishment rule affects the range of inventory fluctuations of the retailer. To this end, as a first step, we shall now discuss the inventory fluctuation range of the retailer under uncertain demand with Theorem 1.

**Theorem 1.** Assume that  $R_1 < I_r \leq (G + \rho)d_{max} - d_{min}$  and  $\rho \geq 1$ . Define  $IL'_{r,t} = IL_{r,t} - D_t$  as the inventory level of the retailer at the end of each period, then the retailer's inventory position satisfies

$$Gd_{min} - d_{max} \leq IP_{r,t} \leq (G + \rho)d_{max} - d_{min}; \tag{18}$$

and  $IL'_{r,t}$  satisfies

$$\begin{aligned}
 Gd_{min} - (L_r \rho + 2)d_{max} &\leq IL'_{r,t} \\
 &\leq (G + \rho)d_{max} - 2d_{min}.
 \end{aligned} \tag{19}$$

*Proof.* Let  $I_r$  be the initial inventory of the retailer and  $R_1$  be the reorder point in the first period. We set  $I_r > R_1$  and  $Gd_{min} \leq R_1 \leq Gd_{max}$ . Because the inventory position is less than the reorder point, the manufacturer will not dispatch goods to the retailer until the retailer's inventory position drops below the reorder point. The dynamics for the retailer's

inventory level and reorder point are depicted in Figure 1. Without loss of generality, we assume that  $IP_{r,t} \geq R_t$  for  $t < t_1^*$ ,  $IP_{r,t} < R_t$  for  $t_1^* \leq t < t_2^*$ , and  $IP_{r,t} \geq R_t$  when  $t = t_2^*$ . In the first stage  $t < t_1^*$ , we have  $Q_t = 0$  and  $Gd_{min} \leq R_t \leq IP_{r,t}$ . In the second stage  $t_1^* \leq t < t_2^*$ , because  $IP_{r,t} < R_t$  the manufacturer should send the replenishment order to the retailer, which is given by  $Q_t = \rho D_{t-1}$ . For  $t = t_1^*$ , we have

$$IP_{r,t_1^*} = IP_{r,t_1^*-1} - D_{t_1^*-1} \geq Gd_{min} - d_{max}. \quad (20)$$

When  $t_1^* < t < t_2^*$

$$IP_{r,t_1^*+1} = IP_{r,t_1^*-1} - D_{t_1^*-1} + \rho D_{t_1^*-1} - D_{t_1^*} \quad (21)$$

and

$$IP_{r,t} = IP_{r,t_1^*-1} + \sum_{i=0}^{t-1-t_1^*} (\rho - 1) D_{t_1^*-1+i} - D_{t-1}. \quad (22)$$

Because  $\rho \geq 1$ , the upper bound and lower bound for the inventory position  $IP_{r,t}$  are obtained by

$$Gd_{min} - d_{max} \leq IP_{r,t} \leq Gd_{max}. \quad (23)$$

Inequality (23) implies that there exists a lower bound for  $IP_{r,t}$  once it drops below the reorder point. In the third stage  $t \geq t_2^*$ , we still have  $Gd_{min} \leq R_t \leq IP_{r,t}$  and

$$\begin{aligned} IP_{r,t_2^*} &= IP_{r,t_2^*-1} + \rho D_{t_2^*-2} - D_{t_2^*-1} \\ &\leq (G + \rho) d_{max} - d_{min}. \end{aligned} \quad (24)$$

Further, we can obtain

$$IP_{r,t_2^*+1} = IP_{r,t_2^*-1} + \rho D_{t_2^*-2} - D_{t_2^*-2} - D_{t_2^*-1} \leq IP_{r,t_2^*}. \quad (25)$$

It means that once the inventory position is higher than the reorder point, it starts to decrease. The inventory level at the end of each period for the retailer is represented as

$$IL'_{r,t} = IP_{r,t} - \sum_{i=1}^{L_r} Q_{t-i} - D_t \quad (26)$$

Further, we directly obtain

$$\begin{aligned} Gd_{min} - (L_r \rho + 2) d_{max} &\leq IL'_{r,t} = IP_{r,t} - \sum_{i=1}^{L_r} Q_{t-i} - D_t \\ &\leq (G + \rho) d_{max} - 2d_{min}. \end{aligned} \quad (27)$$

It completes the proof of Theorem 1.  $\square$

From Theorem 1, we see that the retailer subsystem is BIBO stable if and only if  $\rho \geq 1$ . When  $\rho < 1$ , the retailer subsystem will be destabilized because its demand can not be sufficiently satisfied and its inventory will decrease continuously. In addition, Theorem 1 also provides us some managerial insights. The retailer's service level is mainly determined by the parameter  $G$ , which is employed to set the reorder point. A large  $G$  results in better customer service

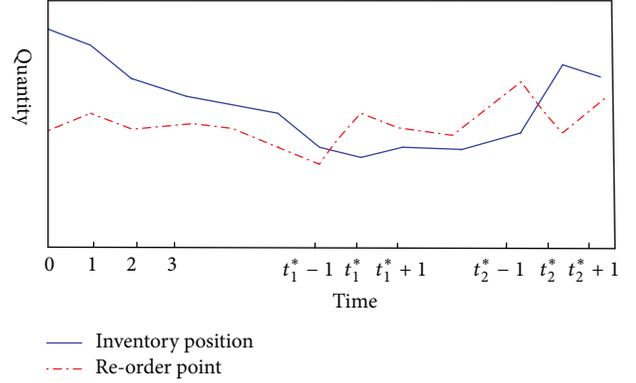


FIGURE 1: The schematic dynamic curves for the retailer's inventory position and reorder point.

level but incurs high inventory cost and leads to negative effect to cause huge inventory oscillations. As mentioned, the delivery frequency for the retailer is determined by  $\rho$ ; that is, a large value of  $\rho$  can be exploited to reduce transportation cost. However, high delivery frequency also increases inventory fluctuations. As the replenishment lead time increases, the retailer's inventory cost might increase. Actually, it is well recognized that lead time is one of the main causes of the bullwhip effect. Finally, demand uncertainties are also key drivers for high inventory cost.

## 5. Inventory and Production Fluctuations of the Manufacturer

Under the VMI program, the manufacturer manages both its own inventory and the retailer's inventory based on the systematic information. The shifting of the authority of inventory control and information sharing, two main properties of a VMI supply chain system, must affect the dynamics properties of the entire system. Disney and Towill [30] studied the stability of a similar VMI supply chain model using transfer function approach. However, they only provided the stability results for small lead times while ignoring inventory fluctuations of the VMI supply chain system. For example, the shipment policy used by the manufacturer to control the retailer's inventory will certainly affect the retailer's dynamics, as emphasized in the previous section. Meanwhile, the retailer's dynamics have great influences on the dynamics of the manufacturer. Wei et al. [21] studied the stability of the discrete APIOBPCS model for a single manufacturer, the methods in which can be extended to a two-echelon VMI supply chain system in Section 4.

As mentioned, if  $0 < \theta_r < 1$  and  $0 < \theta_f < 1$ , the two variables  $F_t$  and  $R_t$  will be bounded within certain ranges. Because final demand and demand forecast signal can be treated as external perturbations, from (12), we see that the stability of the production process of the manufacturer is determined by

$$P_t = (1 - \alpha^{wip}) P_{t-1} + (\alpha^{wip} - \alpha^i) P_{t-L_f}. \quad (28)$$

We recall that  $\rho < 1$  will result in more and more inventory shortages for the retailer. In this situation, the retailer subsystem can be hardly stable. Then, an interesting phenomenon arises. The manufacturer's production process can be stable even if the retailer's inventory is unstable. Here, we must distinguish the stability of the manufacturer's production process from the stability of the manufacturer subsystem. We say that the production process is stable as long as the variable  $P_t$  is bounded, while the stability for the manufacturer subsystems requires both its inventory related variables, such as  $SI_t$  and  $IL_{m,f}$ , and production process  $P_t$  are all bounded. Actually, this interesting result is due to the fact that the production decision of the manufacturer is made based on system inventory  $SI_t$ , which is defined as  $SI_t = IP_{r,t} - R_t + IL_{m,f}$ , rather than  $IL_{m,f}$ . Thus, we can get the conclusion that even though  $\rho < 1$ , the production process can be stable, but the manufacturer's inventory will increase continuously. Thus, we can draw an initial conclusion that there must be a trade-off between inventory fluctuations and production fluctuations. This result is also the key difference between VMI systems and traditional supply chain systems. This result will be further discussed with more details by simulation experiments later. As a result, if  $\rho \geq 1$ , the stability of (28) guarantees the stability for the entire VMI supply chain system.

In fact, the stability of (28) has been discussed in Wei et al. [21] by studying a single manufacturer. Here we can obtain a necessary and sufficient stability condition for the VMI supply chain by Theorem 2.

**Theorem 2.** *Let  $L_f \geq 2$ . Then the VMI supply chain is stable if and only if  $|1 - \alpha^{wip}| < L_f / (L_f - 1)$ ,  $\rho \geq 1$ , and*

- (a)  $|1 - \alpha^{wip}| - 1 < \alpha^i - \alpha^{wip} < [2 + (\alpha^{wip})^2 - 2\alpha^{wip} - 2|1 - \alpha^{wip}| \cos \phi]^{1/2}$  for even  $L_f$ , or
- (b)  $|1 - \alpha^i| < 1$  and  $|\alpha^i - \alpha^{wip}| < [2 + (\alpha^{wip})^2 - 2\alpha^{wip} - 2|1 - \alpha^{wip}| \cos \phi]^{1/2}$  for odd  $L_f$ ,

where  $\phi$  is the solution in  $(0, \pi/L_f)$  of  $|1 - \alpha^{wip}| = \sin(L_f \phi) / \sin(L_f \phi - \phi)$ .

It should be remarked that  $|1 - \alpha^{wip}| < L_f / (L_f - 1)$  guarantees only one solution for  $|1 - \alpha^{wip}| = \sin(L_f \phi) / \sin(L_f \phi - \phi)$  in  $(0, \pi/L_f)$ . From Theorem 2, we see that the stability boundaries for even or odd production lead times are quite different. Based on Theorem 2, we can draw precise stability boundaries or stable regions for arbitrary production lead times.

In order to simplify the selection of replenishment parameters ensuring the system is stable, we can further obtain the stability condition independent of the production lead time, which is introduced in Theorem 3. The proof for Theorem 3 is similar to the proof for the Proposition 4.3 in Wei et al. [21] and we will not include the details here.

**Theorem 3.** *The VMI supply chain is stable if  $\rho \geq 1$  and*

$$|1 - \alpha^{wip}| + |\alpha^{wip} - \alpha^i| < 1. \quad (29)$$

In fact, the precise stable regions gradually converge to the stable region independent of production lead times as the production lead time increases. Figure 2 shows the stable regions in the  $\alpha^{wip} - \alpha^i$  plane for different production lead times, in which  $P$ -Region represents precise stable regions tailored for specific lead times with Theorem 2, and  $S$ -Region represents the stable region independent of lead times. It is obvious that the area of  $S$ -Region will cover most of the areas of the precise stable region  $P$ -Region as the production lead time increases. It implies that the  $S$ -Region simplifies the selection of replenishment parameters.

From above, we see that the production process of the manufacturer can be smoothed even with low delivery frequency determined by the manufacturer to control retailer's inventory. In this sense, the manufacturer can reduce transportation cost and production cost simultaneously. However, the manufacturer must be cautious to determine the replenishment frequency or the size of order batching, because there still exists a trade-off between inventory cost and production cost caused by production fluctuations. This trade-off will be further explored via simulation experiments in the next section.

## 6. Simulation Experiments

In the simulation experiments, we will firstly validate the theoretical results on the range of inventory fluctuations of the retailer subsystem. Then, we will study the dynamics interaction between the retailer and the manufacturer. Finally, the impact of different parameters on the production smoothing will be numerically studied with the worst-case bullwhip measure.

**6.1. Inventory Fluctuations of the Retailer.** We will validate the theoretical results with uniformly distributed demand  $D_t \sim [d_{min}, d_{max}]$ . The reason we choose uniformly distributed demand is that the upper bound and lower bound can be readily known. It is worth noting that the above results can be applied to other demand functions as long as we know the maximal value and the minimal value of the demand data. Based on Theorem 1, we define  $\wp_1 \triangleq (G + \rho)d_{max} - 2d_{min}$  and  $\wp_2 \triangleq Gd_{min} - (L_r \rho + 2)d_{max}$  as the upper bound and lower bound of the retailer's inventory, respectively. Further, we measure the range of inventory fluctuation as

$$\begin{aligned} \Delta_{\wp} &\triangleq \wp_1 - \wp_2 \\ &= (G + 2)(d_{max} - d_{min}) + \rho(1 + L_r)d_{max}. \end{aligned} \quad (30)$$

In the simulations, we compute the retailer's system performance by four metrics: replenishment frequency,  $FR_r$ , the customer service level metric,  $SL_r$ , the maximal value of inventory level,  $IL_{max}$ , and the minimal value of inventory level,  $IL_{min}$ . Assume that we simulate the retailer's dynamics for  $N$  periods. Then the metric  $FR_r$  is represented by

$$FR_r \triangleq \frac{C(Q_t)}{N}, \quad (31)$$

where  $C(Q_t)$  records the number of periods that the manufacturer should deliver a replenishment order to the retailer

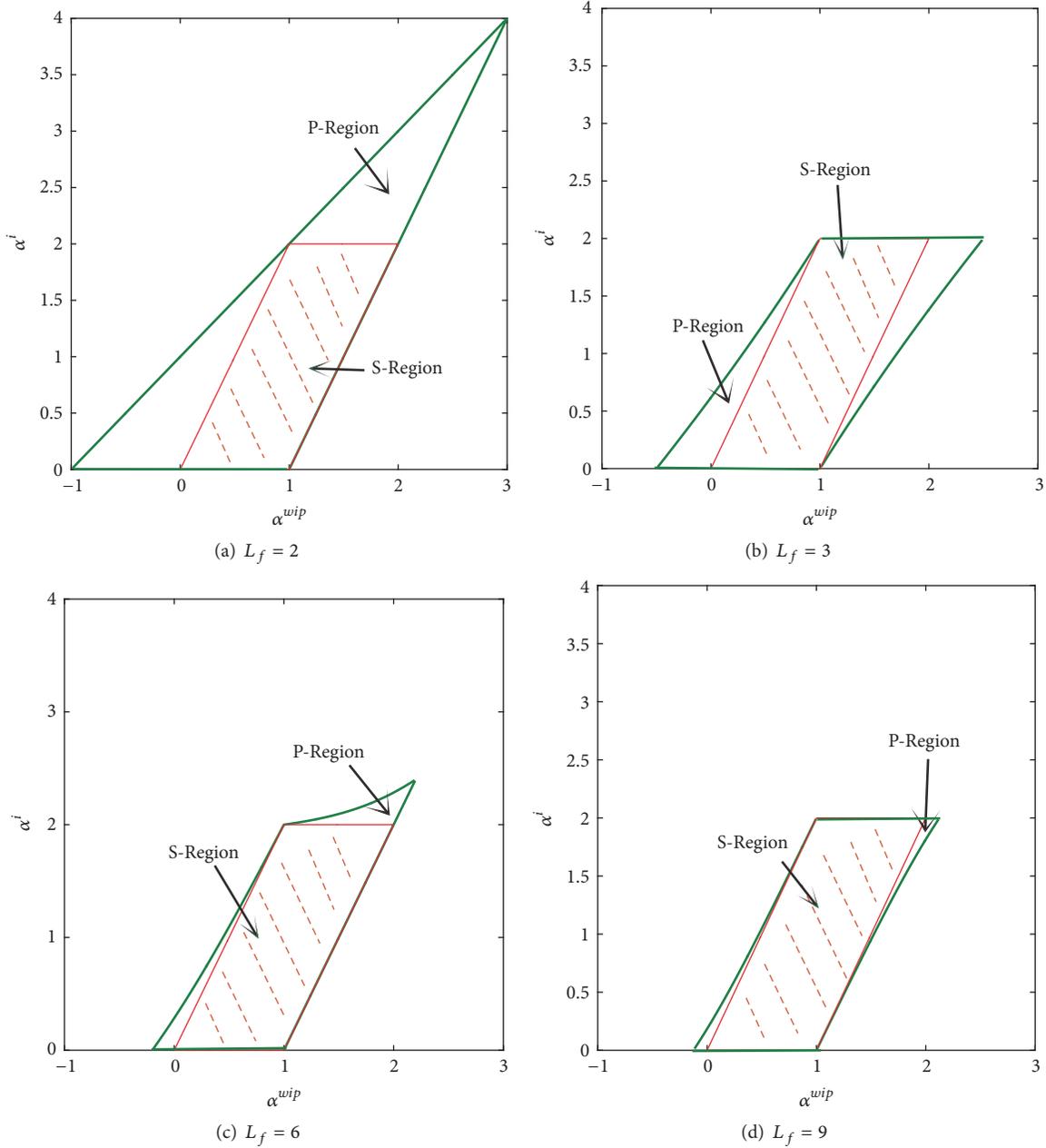


FIGURE 2: The stability boundaries for the VMI supply chain system.

( $Q_t > 0$ ). It means that a higher  $FR_r$  will result in a higher frequency of replenishment order. As a result, the metric  $FR_r$  is closely relevant to the transportation cost. The transportation cost can be saved for a small  $FR_r$ . The customer service level metric  $SL_r$  is represented as

$$SL_r = \frac{S(IL'_{r,t})}{N}, \quad (32)$$

where  $S(IL'_{r,t})$  is the number of periods satisfying  $IL'_{r,t} > 0$ . The service level is high if the value of  $SL_r$  is large.

The other two metrics  $IL_{max}$  and  $IL_{min}$  are defined as

$$IL_{max} = \max \{IL'_{N_0}, IL', \dots, IL'_N\} \quad (33)$$

and

$$IL_{min} = \max \{IL'_{N_0}, IL', \dots, IL'_N\}, \quad (34)$$

in which  $N_0$  is the starting period used to eliminate the influence of initial data on system performance. Based on these four metrics, four groups of simulations are designed to observe the impact of demand fluctuations,  $G$ ,  $\rho$ , and  $L_r$  on system performance. In the simulation, we set  $N = 5000$ ,  $N_0 = 200$ , and  $I_r = 100$ . Theoretical results and simulation results are compared in Table 1.

Table 1 well validates the results obtained previously. The first group of simulation experiment  $G_1$  demonstrates that demand fluctuations are closely related to inventory

TABLE 1: The comparison between simulation results and theoretical results for the retailer subsystem.

Groups	Parameter values				Theoretical results				Simulation results			
	$d_{min}$	$d_{max}$	$G$	$\rho$	$L_r$	$\varrho_1$	$\varrho_2$	$\Delta_\varrho$	$FR_r$	$SL_r$	$IL_{max}$	$IL_{min}$
$G_1$	<b>10</b>	<b>40</b>	4	1.5	2	-160	200	360	0.57	0.99	128.30	-27.11
	<b>10</b>	<b>80</b>	4	1.5	2	-360	420	780	0.52	0.97	265.67	-63.48
	<b>10</b>	<b>120</b>	4	1.5	2	-560	640	1200	0.50	0.95	372.84	-125.67
$G_2$	10	40	<b>2</b>	1.5	2	-180	120	300	0.58	0.43	-60.96	64.32
	10	40	<b>6</b>	1.5	2	-100	260	360	0.52	1.00	185.89	7.28
	10	40	<b>10</b>	1.5	2	-180	300	480	0.54	1.00	298.74	98.37
$G_3$	10	40	6	<b>1</b>	2	-100	260	360	1.00	0.48	44.30	-40.93
	10	40	6	<b>2</b>	2	-180	300	480	0.41	1.00	199.87	19.29
	10	40	6	<b>5</b>	2	-420	420	840	0.16	1.00	307.03	17.53
$G_4$	10	40	4	1.5	<b>3</b>	-220	200	420	0.43	0.81	101.57	-53.32
	10	40	4	1.5	<b>5</b>	-340	200	540	0.41	0.16	72.70	-108.38
	10	40	4	1.5	<b>7</b>	-460	200	660	0.565	0.01	22.90	-173.32

fluctuations. Meanwhile, demand uncertainty could result into poor customer service level. The second group  $G_2$  shows that increasing the parameter  $G$  will improve customer service level. However, high inventory fluctuations will occur if the parameter  $G$  is sufficiently large. In the third group of simulation, it is obvious that increasing  $\rho$  increases the delivery frequency for the retailer. This means that the generalized replenishment policy (2) is more flexible in determining the replenishment frequency than  $\rho = 1$ . However, Theorem 1 illustrates that a large  $\rho$  contributes to inventory fluctuations. Actually, in the third group of simulation  $G_3$ , it is also interesting to observe that  $\rho = 1$  can make the manufacturer lose control of the retailer's inventory because  $FR_r = 1$ . This situation leads to low customer service level because the reorder point loses its ability to limit the retailer's inventory. The final simulation group  $G_4$  validates the result that long lead time brings difficulties for inventory control.

**6.2. Production and Inventory Fluctuations of the Manufacturer.** We shall now simulate the system with a unit step signal as the demand input, which is represented by

$$D_t = \begin{cases} 1 & \text{if } t \geq 1, \\ 0 & \text{if } t < 1. \end{cases} \quad (35)$$

The step signal is widely used in testing dynamical systems. Actually, the step signal can be used to model sudden change of customer demand due to accidental events. The well-known Supply Chain Operations Reference (SCOR) model also measures supply chain performance in response to a step change in demand. In this subsection, we select a common group of parameter setting:  $\theta_r = \theta_f = 1/9$ ,  $G = 4$ ,  $L_r = 2$ ,  $I_r = 5$ ,  $R_1 = 0$ , and  $I_f = 0$ . In the following, in order to figure out the reason why the production process can be stabilized even if the retailer subsystem is destabilized with  $\rho < 1$ , we will firstly explore the dynamic properties under two different scenarios, namely,  $\rho < 1$  and  $\rho > 1$ . Then, we will validate the stability results derived previously.

**6.2.1. Production and Inventory Dynamics.** In the case when  $\rho < 1$ , the retailer subsystem is destabilized. In the simulation experiments, we firstly set  $\rho = 0.8$ ,  $\alpha^i = 0.25$ , and  $\alpha^{wip} = 0.5$ . The production and inventory dynamics for the VMI system in response to the step signal input are shown in Figures 3(a) and 3(b). Figure 3(a) shows that the production process,  $P_t$ , and system inventory,  $SI_t$ , are stable. However,  $\rho < 1$  causes serious outcome that the inventory levels for both the retailer and the manufacturer are divergent or unstable. In particular, the customer demand of the retailer can never be fully satisfied because the replenishment amount is always less than realized demand. Even if the production process is smoothed, the manufacturer accumulates high inventory, which will further lead to extra inventory cost. Thus, we should avoid selecting the parameter setting  $\rho < 1$ .

The advantages of  $\rho > 1$  are shown in Figures 3(c) and 3(d) by setting  $\rho = 3$ ,  $\alpha^i = 0.25$ , and  $\alpha^{wip} = 0.5$ . This parameter setting falls into the stable region determined by Theorems 2 and 3. In this situation, the delivery frequency is decreased to reduce transportation cost. Although the replenishment order for the retailer fluctuates with a large magnitude, the production dynamics is still smoothed. In this sense, the production cost due to machine startup or labor force scheduling can be saved. However, the manufacturer should be cautious on selecting the delivery frequency or determining the value for  $\rho$ , as the inventory fluctuations magnitude might be large for both the retailer and the manufacturer, which increases inventory cost for the retailer and penalty cost for the manufacturer. As introduced in the previous section,  $\rho = 1$  may lead to a situation that the reorder point makes no sense for large value of  $G$ , while incorporating the parameter  $\rho$  improves the flexibility for the manufacturer to manage the retailer's inventory greatly. In practice, managers should make a balanced decision towards transportation cost, inventory cost, and production cost.

**6.2.2. Validation for Stability Results.** Here we simply test the stability results for  $\rho = 1$ ,  $L_f = 9$ , and  $\alpha^{wip} = 1.2$ . In such

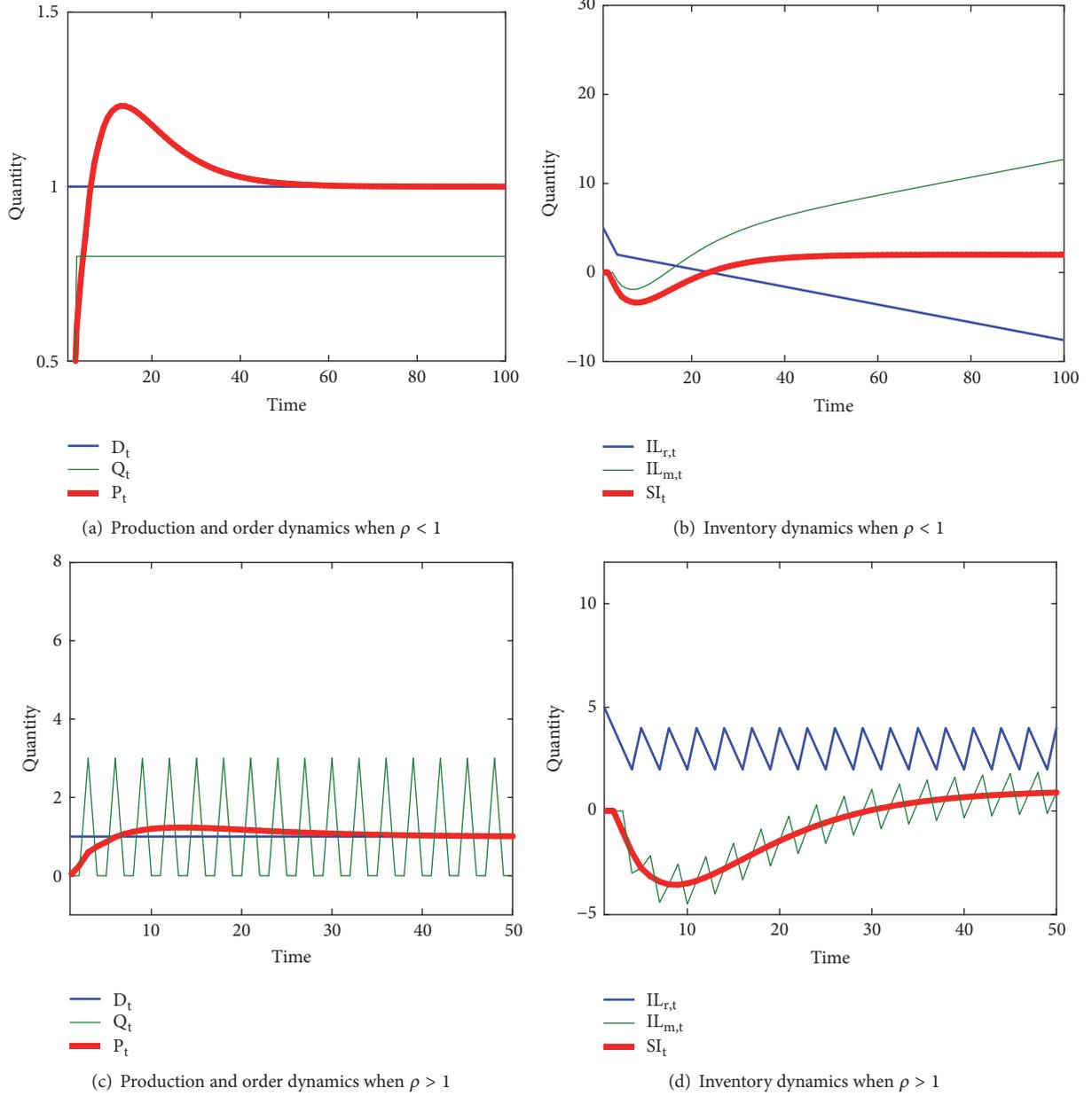


FIGURE 3: The inventory and production dynamics for the VMI supply chain system.

a situation, the necessary and sufficient stable condition for the VMI supply chain system becomes  $0.3858 < \alpha^i < 2$ . In the simulation, we consider two scenarios: unstable scenario,  $\alpha^i = 0.32$ , and stable scenario,  $\alpha^i = 0.42$ . The inventory dynamics for the VMI supply chain system are shown in Figure 4. Figure 4(a) shows the dynamics for  $\alpha^i = 0.32$ , in which all the state variables diverge as time passes. Unstable design always bring high inventory and production cost as a result of inventory fluctuations. Thus, unstable behavior should be avoided by selecting appropriate replenishment parameters. In contrast, Figure 4(b) shows the stable behavior for  $\alpha^i = 0.42$ . In a stable system, inventory and production will return to steady state even under sudden change of customer demand.

**6.3. A Worst-Case Bullwhip Effect Study.** For a manufacturer, reducing production fluctuations is significant because it is closely related to production cost. Smoothing production process helps arranging labor force, deciding the manufacturing capacity, and also enlarging the life of production machines. It is well recognized that production fluctuations can be smoothed by mitigating the bullwhip effect [46]. Actually, stability is a fundamental step for smoothing production fluctuations because it makes no sense to study the bullwhip for an unstable supply chain system. In the following, based on the stability conditions in Theorems 2 and 3, we will further study the production smoothing behavior in terms of the worst-case bullwhip effect based on the transfer function model (17).

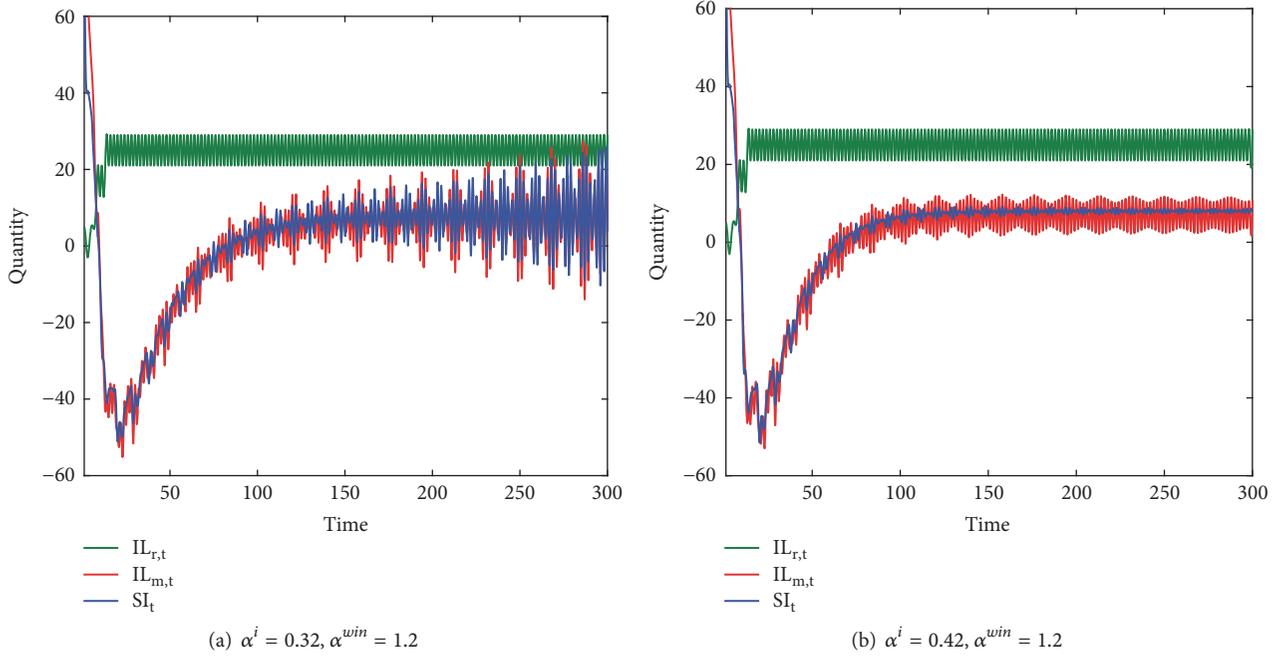


FIGURE 4: Validation of stability results.

The bullwhip effect problem has received considerable attention since it causes high cost for upstream firms in a supply chain firm due to the distortion of demand information. The majority of the literature focuses on the bullwhip effect of supply chain system for specific demand with statistical method and control theory method [50]. However, the customer demand in real life is highly uncertain over time and the probability distribution used to approximate real life demand can be inaccurate or unreliable [51]. To this end, we aim to understand the bullwhip effect of the VMI supply chain by combining the state space model and transfer approach for unknown demand. As mentioned, the state space model can be readily constructed from the difference model, which can be conveniently used to derive transfer functions. As mentioned, it is complicated to derive transfer functions by combining block graphs. Here we study the bullwhip effect for unknown demand using transfer function because it is easier to calculate the robust bullwhip metric.

Assume that  $G(z)$  is the transfer function by selecting  $P_t$  as output and  $D_t$  as input in the frequency domain. Then the worst-case bullwhip effect metric [21, 52] is represented as

$$W_I^2 := \left( \sup_{\omega \in [0, 2\pi)} |G(z)|_{z=e^{j\omega}} \right)^2, \quad (36)$$

where  $|G(z)|$  is the modulus of the transfer function  $G(z)$ . From the definition (36), we see that the metric  $W_I^2$  is the squared worst-case amplification of  $G(z)$  across all  $\omega \in [0, 2\pi)$ . The most important property is that  $W_I^2$  is independent of customer demand. In particular, the VMI supply chain system experiences no bullwhip effect when  $W_I^2 < 1$ .

Using Matlab 7.0, we can calculate the value of  $W_I^2$  based on the transfer function model (17) accurately. The following numerical analysis will disclose the influences of replenishment parameters, production lead time, and forecasting parameters on the worst-case bullwhip effect for unknown demand. We see that the parameter  $\rho$  is not included in the state space model (13) and thus it has no influence on the smoothness of production fluctuations. Meanwhile, the robust metric  $W_I^2$  is irrelevant to the retailer's lead time  $L_r$ . Here we focus on the influences of  $\alpha^i$ ,  $\alpha^{wip}$ ,  $G$ , and  $\theta_f$  on the worst-case bullwhip effect.

When  $L_f = 3$ ,  $\eta = 1$ ,  $G = 4$ , and  $\theta_r = \theta_f = 0.2$ , Figures 5(a) and 5(b) show the relationship between replenishment parameters and the worst-case bullwhip  $W_I^2$ . For a specified  $\alpha^{wip}$ , the bullwhip effect will be very serious if we increase the value of  $\alpha^i$ . In contrast, for a specified  $\alpha^i$ , the relationship between  $\alpha^{wip}$  and  $W_I^2$  roughly follows a convex function. It means that, for fixed  $\alpha^i$ , there exists a optimal  $\alpha^{wip}$  to dampen production fluctuations. Figure 5(c) shows the impact of parameter  $G$  on the bullwhip metric  $W_I^2$ . It has been emphasized that the parameter  $G$  can be used to adjust the customer service level of the retailer. A large  $G$  produces a high reorder point and thus brings extra inventory cost. However, Figure 5(c) indicates that increasing the value of  $G$  will cause serious production fluctuations. Figure 5(d) shows the impact of demand forecasting parameter  $\theta_f$  on the robust performance. It is quite clear to see that a large value of  $\theta_f$  causes higher bullwhip effect for unknown demand. This phenomenon can be intuitively explained by the forecasting algorithm (7); that is, a small  $\theta_f$  smooths the demand signal. A large  $\theta_f$  implies that demand forecast follows the

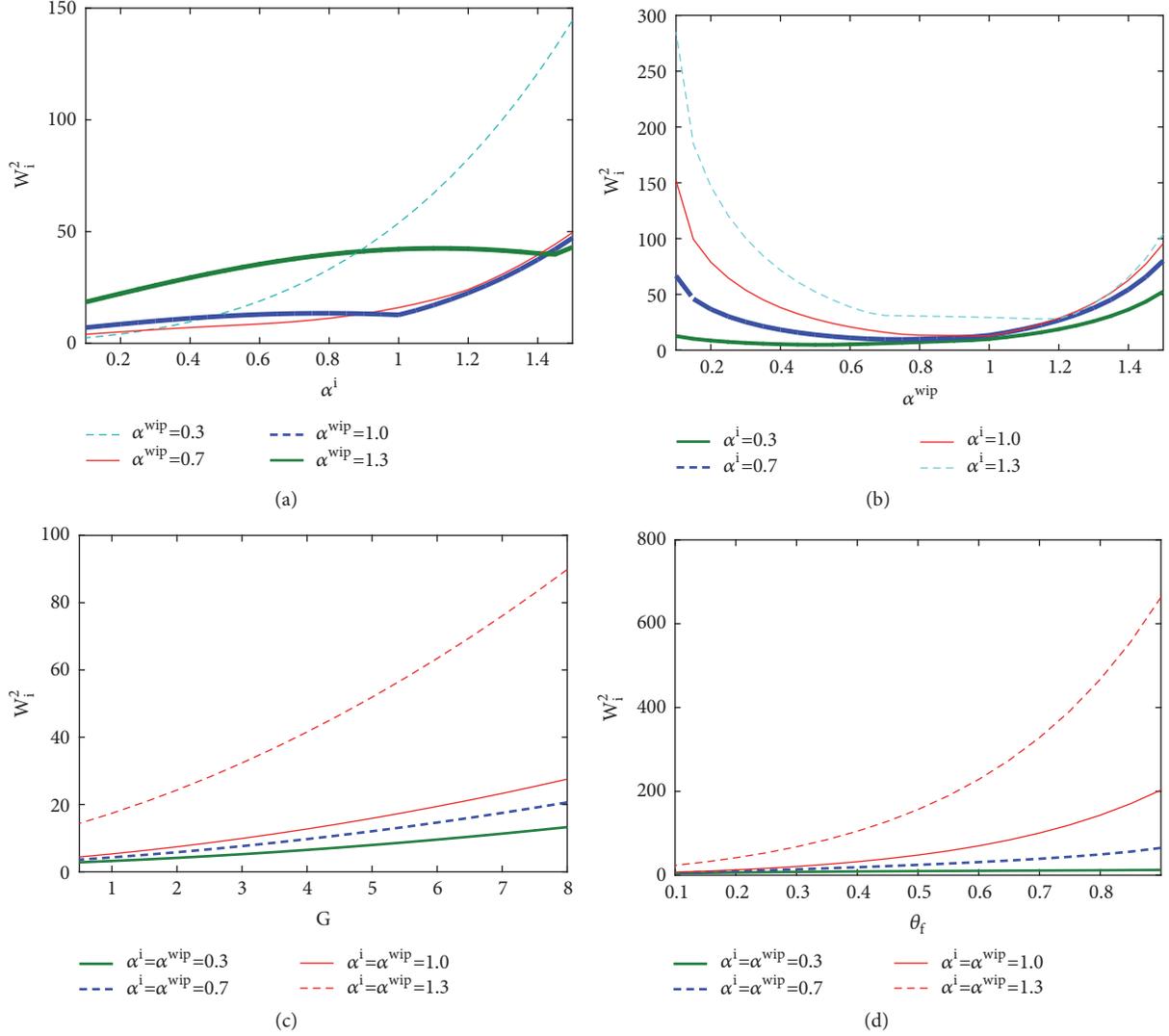


FIGURE 5: The impact of system parameters on the worst-case bullwhip.

historical demand immediately and causes high magnitude of production fluctuations.

## 7. Discussions and Conclusions

This paper has investigated the dynamics of a VMI supply chain system composed of one manufacturer and one retailer. To provide guidelines for the manufacturer to better manage the retailer's inventory, we firstly explored the fluctuation range of the retailer's inventory. We found that the retailer's inventory fluctuations are mainly influenced by demand characteristics, delivery frequency, lead time, and other parameters. Specifically, the magnitude of the retailer's inventory fluctuation is an increasing function of the parameter  $\rho$ , which reflects the delivery frequency. We notice that  $\rho = 1$  might cause poor customer service because in this situation, the reorder point might lose its ability to control the retailer's inventory. This result means that incorporating the parameter  $\rho$  into the VMI model not only enhances the flexibility

of decision-making, but also is useful in improving system performance. In practice, the upstream manufacturer should make replenishment decisions by optimizing inventory cost, transportation cost, and production cost simultaneously.

The shifting of the authority to manage the retailer's inventory and information sharing certainly affect the dynamics for the manufacturer. Based on the difference equations, we investigated the dynamical properties of the entire VMI supply chain system. It is interesting to demonstrate that, under the VMI program, the production process can be stabilized even if the retailer subsystem is unstable. This result seems to be favorable for the manufacturer; however, the retailer will face serious shortages because the manufacturer fails to satisfy the demand of the retailer. Until present, the majority of the literature on the dynamics of the VMI supply chain system is limited to simulation studies. For example, Disney et al. [53] showed that, under the VMI program, the manufacturer can choose a large delivery frequency to achieve economical scale without causing serious production

fluctuations. It means that production cost and transportation cost can be reduced simultaneously, which is one of the advantages of VMI. In this paper, we confirm this result through theoretical analysis. However, we emphasize that managers should be cautious on determining the delivery frequency used to satisfy the retailer's demand. Although a high delivery frequency might be helpful in reducing transportation cost, the inventory cost of the retailer and the manufacturer can increase and the customer service level of the retailer can be very poor.

Based on the difference equations, we also studied the stability of the VMI supply chain. Stability is a fundamental problem for production smoothing. Although the production decision of the manufacturer is based on system inventory, we found that the stability condition of the VMI supply chain is very similar to a one-echelon production and inventory control system, as has been studied in Wei et al. [21]. To simplify the selection of replenishment parameters, we also obtained the stable region independent of the production lead time. Furthermore, we found that the stable region independent of the production lead time approximates the precise stable region as the production lead time increases. We further also studied production smoothing in terms of the worst-case bullwhip effect of the VMI supply chain system for unknown demand based on a transfer function method. Based on the transfer function of the VMI supply chain system, we numerically validate the effectiveness of the robust metric.

It is worth noting that this paper focuses on a VMI supply chain with a relatively simple structure and specific assumptions. In practice, the VMI program can be implemented for multiple supplier systems or distribution networks. In addition, a realistic supply chain system must be nonlinear [54]; thus incorporating nonlinearities into VMI models is a significant problem deserving future research.

## Data Availability

The paper is focused on modeling and simulation. All the data is generated by simulation experiments with different parameter settings or available from the corresponding author upon request.

## Conflicts of Interest

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

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (nos. 71401181 and 71701213) and the MOE (Ministry of Education in China) Project of Humanities and Social Sciences (nos. 14YJC630136 and 15YJC630008).

## References

- [1] K. Govindan, "Vendor-managed inventory: A review based on dimensions," *International Journal of Production Research*, vol. 51, no. 13, pp. 3808–3835, 2013.
- [2] A. Angulo, H. Nachtmann, and M. A. Waller, "Supply chain information sharing in a vendor managed inventory partnership," *Journal of Business Logistics*, vol. 25, no. 1, pp. 101–120, 2004.
- [3] J.-Y. Lee, R. K. Cho, and S.-K. Paik, "Supply chain coordination in vendor-managed inventory systems with stockout-cost sharing under limited storage capacity," *European Journal of Operational Research*, vol. 248, no. 1, pp. 95–106, 2016.
- [4] S. Çetinkaya and C.-Y. Lee, "Stock replenishment and shipment scheduling for vendor-managed inventory systems," *Management Science*, vol. 46, no. 2, pp. 217–232, 2000.
- [5] P. Danese, "Beyond Vendor Managed Inventory: the Glaxo-smithkline Case," *Supply Chain Forum: An International Journal*, vol. 5, no. 2, pp. 32–40, 2004.
- [6] E. Watson, "Nestle switches to vendor managed inventory with tesco," *Food Manufacture*, vol. 80, no. 8, p. 20, 2008.
- [7] V. A. Micheau, "How boeing and alcoa implemented a successful vendor managed inventory program," *The Journal of Business Forecasting*, vol. 24, no. 1, p. 17, 2005.
- [8] J. H. Bookbinder, M. Gumus, and E. M. Jewkes, "Calculating the benefits of vendor managed inventory in a manufacturer-retailer system," *International Journal of Production Research*, vol. 48, no. 19, pp. 5549–5571, 2010.
- [9] X. Lu, J. Shang, S.-y. Wu, G. . Hegde, L. Vargas, and D. Zhao, "Impacts of supplier hubris on inventory decisions and green manufacturing endeavors," *European Journal of Operational Research*, vol. 245, no. 1, pp. 121–132, 2015.
- [10] S. Birim and C. Sofyalioglu, "Evaluating vendor managed inventory systems: how incentives can benefit supply chain partners," *Journal of Business Economics and Management*, vol. 18, no. 1, pp. 163–179, 2017.
- [11] J. Ru, R. Shi, and J. Zhang, "When Does A Supply Chain Member Benefit from Vendor-Managed Inventory?" *Production Engineering Research and Development*, vol. 27, no. 5, pp. 807–821, 2018.
- [12] S. M. Disney and D. R. Towill, "The effect of vendor managed inventory (VMI) dynamics on the Bullwhip Effect in supply chains," *International Journal of Production Economics*, vol. 85, no. 2, pp. 199–215, 2003.
- [13] M. Waller, M. E. Johnson, and T. Davis, "Vendor-managed inventory in the retail supply chain," *Journal of Business Logistics*, vol. 20, no. 1, p. 183, 1999.
- [14] M. J. Fry, R. Kapuscinski, and T. L. Olsen, "Coordinating Production and Delivery under a (z, Z)-Type Vendor-Managed Inventory Contract," *Manufacturing and Service Operations Management*, vol. 3, no. 2, pp. 151–173, 2001.
- [15] M. Ben-Daya, E. Hassini, M. Hariga, and M. M. Aldurgam, "Consignment and vendor managed inventory in single-vendor multiple buyers supply chains," *International Journal of Production Research*, vol. 51, no. 5, pp. 1347–1365, 2013.
- [16] L. C. Coelho and G. Laporte, "An optimised target-level inventory replenishment policy for vendor-managed inventory systems," *International Journal of Production Research*, vol. 53, no. 12, pp. 3651–3660, 2015.
- [17] A. A. Taleizadeh, M. Noori-Daryan, and L. E. Cárdenas-Barrón, "Joint optimization of price, replenishment frequency, replenishment cycle and production rate in vendor managed inventory system with deteriorating items," *International Journal of Production Economics*, vol. 159, pp. 285–295, 2015.
- [18] J. Cai, P. R. Tadikamalla, J. Shang, and G. Huang, "Optimal inventory decisions under vendor managed inventory: substitution effects and replenishment tactics," *Applied Mathematical*

- Modelling: Simulation and Computation for Engineering and Environmental Systems*, vol. 43, pp. 611–629, 2017.
- [19] M. Khan, M. Y. Jaber, S. Zanoni, and L. Zavanella, “Vendor managed inventory with consignment stock agreement for a supply chain with defective items,” *Applied Mathematical Modelling*, vol. 40, no. 15–16, pp. 7102–7114, 2016.
- [20] D. E. Cantor and E. Katok, “Production smoothing in a serial supply chain: A laboratory investigation,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 4, pp. 781–794, 2012.
- [21] Y. C. Wei, H. W. Wang, and C. Qi, “On the stability and bullwhip effect of a production and inventory control system,” *International Journal of Production Research*, vol. 51, no. 1, pp. 154–171, 2013.
- [22] Y. C. Wei, H. W. Wang, and C. Qi, “The impact of stock-dependent demand on supply chain dynamics,” *Applied Mathematical Modelling*, vol. 37, no. 18–19, pp. 8348–8362, 2013.
- [23] H. L. Lee, V. Padmanabhan, and S. Whang, “Information distortion in a supply chain: the bullwhip effect,” *Management Science*, vol. 43, no. 4, pp. 546–558, 1997.
- [24] F. Chen, Y. Wei, and H. Wang, “A heuristic based batching and assigning method for online customer orders,” *Flexible Services and Manufacturing Journal*, pp. 1–46, 2017.
- [25] X. Wang and S. M. Disney, “The bullwhip effect: progress, trends and directions,” *European Journal of Operational Research*, vol. 250, no. 3, pp. 691–701, 2016.
- [26] V. Giard and M. Sali, “The bullwhip effect in supply chains: A study of contingent and incomplete literature,” *International Journal of Production Research*, vol. 51, no. 13, pp. 3880–3893, 2013.
- [27] M. Udenio, E. Vatamidou, J. C. Fransoo, and N. Dellaert, “Behavioral causes of the bullwhip effect: An analysis using linear control theory,” *IIE Transactions*, vol. 49, no. 10, pp. 980–1000, 2017.
- [28] Y. Wei, F. Chen, and F. Xiong, “Dynamic Complexities in a Supply Chain System with Lateral Transshipments,” *Complexity*, vol. 2018, Article ID 3959141, 15 pages, 2018.
- [29] G. Huang, J. Wang, C. Chen, C. Guo, and B. Zhu, “System resilience enhancement: Smart grid and beyond,” *Frontiers of Engineering Management*, vol. 4, no. 3, p. 271, 2017.
- [30] S. M. Disney and D. R. Towill, “A discrete transfer function model to determine the dynamic stability of a vendor managed inventory supply chain,” *International Journal of Production Research*, vol. 40, no. 1, pp. 179–204, 2002.
- [31] B. Brunaud and I. E. Grossmann, “Perspectives in multilevel decision-making in the process industry,” *Frontiers of Engineering Management*, vol. 4, no. 3, p. 256, 2017.
- [32] G. Cachon and M. Fisher, “Campbell soup’s continuous replenishment program: Evaluation and enhanced inventory decision rules,” *Production Engineering Research and Development*, vol. 6, no. 3, pp. 266–276, 1997.
- [33] Y. Yao and M. Dresner, “The inventory value of information sharing, continuous replenishment, and vendor-managed inventory,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 44, no. 3, pp. 361–378, 2008.
- [34] J.-H. Kang and Y.-D. Kim, “Coordination of inventory and transportation managements in a two-level supply chain,” *International Journal of Production Economics*, vol. 123, no. 1, pp. 137–145, 2010.
- [35] S. Çetinkaya, E. Tekin, and C. Lee, “A stochastic model for joint inventory and outbound shipment decisions,” *Institute of Industrial Engineers (IIE). IIE Transactions*, vol. 40, no. 3, pp. 324–340, 2008.
- [36] Y. Dong, M. Dresner, and Y. Yao, “Beyond information sharing: An empirical analysis of vendor-managed inventory,” *Production Engineering Research and Development*, vol. 23, no. 5, pp. 817–828, 2014.
- [37] Y. Yao, P. T. Evers, and M. E. Dresner, “Supply chain integration in vendor-managed inventory,” *Decision Support Systems*, vol. 43, no. 2, pp. 663–674, 2007.
- [38] M. Nagarajan and S. Rajagopalan, “Contracting under vendor managed inventory systems using holding cost subsidies,” *Production Engineering Research and Development*, vol. 17, no. 2, pp. 200–210, 2008.
- [39] J.-Y. Lee and R. K. Cho, “Contracting for vendor-managed inventory with consignment stock and stockout-cost sharing,” *International Journal of Production Economics*, vol. 151, pp. 158–173, 2014.
- [40] B. K. Mishra and S. Raghunathan, “Retailer- vs. Vendor-managed inventory and brand competition,” *Management Science*, vol. 50, no. 4, pp. 445–457, 2004.
- [41] W. Pan and K. C. So, “Component procurement strategies in decentralized assembly systems under supply uncertainty,” *Institute of Industrial Engineers (IIE). IIE Transactions*, vol. 48, no. 3, pp. 267–282, 2016.
- [42] G. Kannan, M. C. Grigore, K. Devika, and A. Senthilkumar, “An analysis of the general benefits of a centralised VMI system based on the EOQ model,” *International Journal of Production Research*, vol. 51, no. 1, pp. 172–188, 2013.
- [43] A. Diabat, “Hybrid algorithm for a vendor managed inventory system in a two-echelon supply chain,” *European Journal of Operational Research*, vol. 238, no. 1, pp. 114–121, 2014.
- [44] M. Hariga, M. Gumus, and A. Daghfous, “Storage constrained vendor managed inventory models with unequal shipment frequencies,” *Omega*, vol. 48, pp. 94–106, 2014.
- [45] A. Mateen, A. K. Chatterjee, and S. Mitra, “VMI for single-vendor multi-retailer supply chains under stochastic demand,” *Computers & Industrial Engineering*, vol. 79, pp. 95–102, 2015.
- [46] G. P. Cachon, T. Randall, and G. M. Schmidt, “In search of the bullwhip effect,” *Manufacturing & Service Operations Management*, vol. 9, no. 4, pp. 457–479, 2007.
- [47] E. I. Jury, *Theory and Application of the z-Transform Method*, Wiley, 1964.
- [48] C. R. Phillips and N. T. Nagle, *IEEE Transactions on Systems, Man, and Cybernetics*, Prentice Hall Press, 2007.
- [49] C. S. Lalwani, S. M. Disney, and D. R. Towill, “Controllable, observable and stable state space representations of a generalized order-up-to policy,” *International Journal of Production Economics*, vol. 101, no. 1, pp. 172–184, 2006.
- [50] J. Dejonckheere, S. M. Disney, M. R. Lambrecht, and D. R. Towill, “Measuring and avoiding the bullwhip effect: a control theoretic approach,” *European Journal of Operational Research*, vol. 147, no. 3, pp. 567–590, 2003.
- [51] Y. Wang, R. Shou, L. H. Lee, and E. P. Chew, “A case study on sample average approximation method for stochastic supply chain network design problem,” *Frontiers of Engineering Management*, vol. 4, no. 3, p. 338, 2017.
- [52] Y. Ouyang, “The effect of information sharing on supply chain stability and the bullwhip effect,” *European Journal of Operational Research*, vol. 182, no. 3, pp. 1107–1121, 2007.

- [53] S. M. Disney, M. M. Naim, and A. Potter, "Assessing the impact of e-business on supply chain dynamics," *International Journal of Production Economics*, vol. 89, no. 2, pp. 109–118, 2004.
- [54] J. Chang and L. Zhao, "Complexity analysis of dynamic cooperative game models for supply chain with the remanufactured products," *Discrete Dynamics in Nature and Society*, Art. ID 9246934, 10 pages, 2018.

## Research Article

# Dynamic Strategies on Firm Production and Platform Advertisement in Crowdfunding considering Investor's Perception

Ying Ji , Ju Wei , Zhong Wu, Shaojian Qu , and Baojun Zhang 

*Business School, University of Shanghai for Science and Technology, Shanghai, China*

Correspondence should be addressed to Ju Wei; [wj725496@163.com](mailto:wj725496@163.com)

Received 27 June 2018; Accepted 13 August 2018; Published 18 September 2018

Academic Editor: Lu Zhen

Copyright © 2018 Ying Ji 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.

Taking investor's perception into account, the optimal decisions about the product quality and platform advertisement are investigated in a dynamic model in the context of crowdfunding. Researches in the literature, however, usually set investor's perception as a fixed value and rarely consider the important phenomenon that the online information has some influences on investor's perception. Considering the effects of information about product quality and platform advertisement on the investor's perception, a dynamic decision model is proposed. Firstly, investment desire and reference price of the investor are introduced in two dynamic settings to describe investor's perception. Then, the optimal decisions about the product quality and platform advertisement are formulated under two circumstances: the sponsor and the platform make decisions independently and they cooperate as a system. Finally, the influences of reference price and cost-sharing ratio on the optimal results are compared and the data simulation experiment verifies the necessity of the study. Some new insights can be drawn for the operations management of the firm in crowdfunding as follows: (i) it is more profitable for the firm to cooperate with the platform when investors pay more attention to their reference price; (ii) it is optimal for the firm to share a larger proportion of platform cost when the profit-sharing ratio is low.

## 1. Introduction

As an effective way of financing, crowdfunding has rapid development in the investment market. Crowdfunding is a general term that refers to the activities of raising funds from common investors through the Internet. In one of the most popular crowdfunding websites, Kickstarter.com, more than 3.5 million people have participated in 108,437 creative projects, and over 2.47 billion dollars has been pledged from the date Kickstarter.com founded till now. More than 115 numbers of crowdfunding platforms have been built and over 1 billion yuan has been raised at the end of 2014 in China [1]. Therefore, how to produce and operate effectively is becoming an increasingly prominent problem for crowdfunding firms with the popularity of crowdfunding activities.

There are four types of crowdfunding projects depending on what the investors can receive from the project, including

donation-based, reward-based, lending, and equity crowdfunding [2–4]. In reward-based crowdfunding, the sponsor publishes the product's material and function information on the platform to introduce the products in detail. The information provides accurate product quality, and the cases that the sponsor provides faulty information to mislead readers are not considered in the present study. After accessing to the information, investors will form a certain valuation about the product quality and they will compare the price of the product with their reference price when they make purchase decisions. If the published online information can not attract investors effectively, the investment will be failing.

Investor's perception frequently appears in the field of financial investment. When investors consider whether to participate in a crowdfunding project or not, the process of making decisions is similar to making purchase decisions in financial market. Therefore, it is reasonable to apply investor's perception to analyze investor's behavior in the context

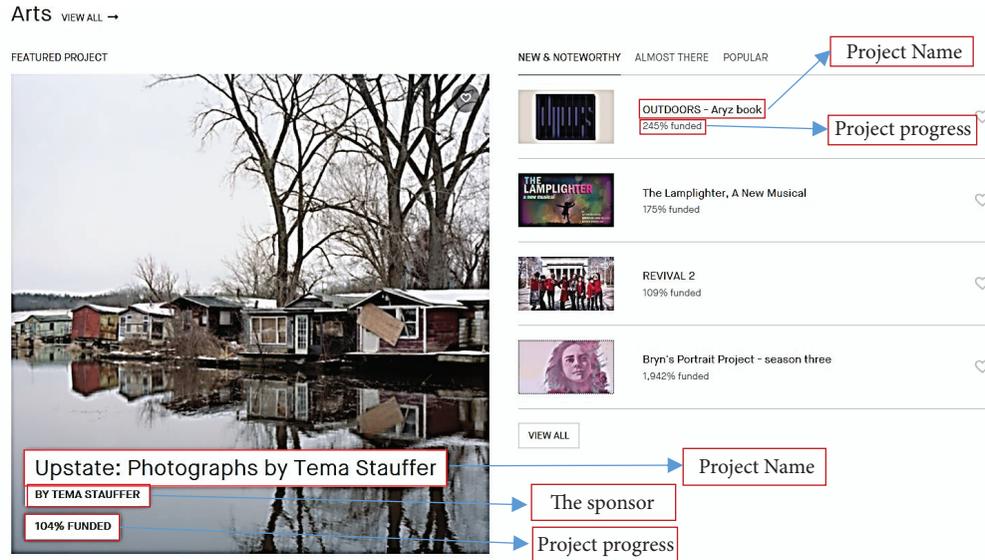


FIGURE 1: One webpage on Kickstarter.com.

of crowdfunding. Researches in the literature usually set investor's perception as a fixed value, and they rarely consider the important phenomenon that the online information will cause some influences on investor's perception.

As the most important way of information dissemination, personal networks and entrepreneur's social network have significant effects on crowdfunding performance [5]. Bi [1] illustrated that the signals of product quality and electronic word of mouth have caused great influences on funder's investment decisions by solving ELM model. Undoubtedly, the information about the product quality is the most important factor that affects investors to buy the product or not. Figure 1 illustrates that the advertisements of the project on platform can also cause some influences on the product valuations of the investors.

Figure 1 is one webpage of art projects on Kickstarter.com. There are five art projects in the picture. It is obvious that the project on the left side, Upstate: Photographs by Tema Stauffer, can attract us mostly when we click on this webpage since it takes more space than the other four projects.

As shown in Figure 1, the location and space of the project advertisement will cause some important influences on investor's perception. Platforms of crowdfunding such as Kickstarter.com decide the location and space of the project introductions in a whole webpage. Further, platforms provide the services of matching fund raisers with investors and they also provide some investment style to fund raisers. Therefore, it is reasonable to suppose the advertisement on the platform as a risk factor that can cause some influences on investor's perception.

Considering the effects of online information on investor's perception, investor's perception applied in the present study is that the investors will continuously change their own valuations about the product based on the information they got. In detail, the investment desire and

reference price of the investor are introduced in two dynamic settings to describe investor's perception.

Therefore, taking investor's perception into account the problems of product design are illustrated in the present study. Product design refers to firms which make suitable designs about product quality, price, marketing strategy, etc. Product design has been proved to be an efficient tool for the firms to produce their product, and the advisable product design shown was an effective way for the actual firms to maintain a long-term competitive advantages [6]. When dynamic investor's perception is taken into consideration, we want to analyze the following four questions:

- (1) How do the online information of the product quality and platform advertisement affect investor's investment decisions?
- (2) What are the optimal decisions of product quality for the sponsor in crowdfunding?
- (3) What are the optimal decisions of platform advertisement in crowdfunding?
- (4) How do the cost-sharing ratio, profit-sharing ratio, and reference price affect the optimal decisions?

To solve the above questions, we propose and analyze a dynamic model in which the sponsor designs the product quality to cater investor's preference, and the platform formulates the strategies about project advertisement to optimize their own profit based on the changeable investor's perception. Firstly, the information about product quality and platform advertisement is identified as the two risk factors that affect investor's perception. Then, investment desire and reference price are used to illustrate investor's perception in detail which are described in two dynamic equations. Finally, the optimal decisions about the product quality and platform advertisement are formulated under two circumstances: the sponsor and the platform make decisions independently and they cooperate as a system.

The rest of the paper is organized as follows. Literature reviews are presented in Section 2. In Section 3, investor's perception is introduced into the dynamic product quality design model and the optimal results are given in two circumstances, respectively. Based on the model, the influences of investor's perception on the optimal results are analyzed, and some profitable strategies for the sponsor are given. In Section 4, some sensitive analyses for parameters are given and the simulation is described in Section 5. Section 6 describes some concluding remarks of the present study.

## 2. Literature Reviews

It is difficult for many innovative companies to get financial support by traditional financing methods since small innovative companies have no power to make public offerings. Crowdfunding provides a solution for the small innovative companies to raise the initial startup capital on the Internet [7, 8]. It is found that Internet financing was a good way to solve the traditional financing difficulties of small companies effectively [9].

The preliminary literatures of crowdfunding focus on the following two aspects. First, some studies have discussed the definition and the business model of the crowdfunding. Crowdfunding was defined as a concept that invites individuals to invest in various section [10]. The concept of crowdfunding originates from the idea of the crowdsourcing which can be seen as the wisdom of a crowd, while the difference in the crowdfunding focuses on the financial perspective of crowdsourcing [11, 12]. For the business model, there are three participators involved in crowdfunding: Internet platforms, fund raisers, and investors. Internet platforms provide the services of matching fund raisers with investors and provide some investment style to fund raisers [13, 14]. As shown in the research of Lukkarinen [15], the driving force of crowdfunding can be summarized as two aspects. One is the characteristics of the project such as the early fund raised by media networks and the value of minimum investment. The other includes the total target, the project's duration, and the provision of financial information.

Second, some studies have investigated which factors affect the crowdfunding mostly. The period of fund raising and the amount of goal were identified as the directive factors related to a project [3, 16–19]. Many researches revealed that an appropriate fund raising goal and an appropriate geographical location of a crowdfunding project have caused some valid influences on the success of crowdfunding [1, 3, 5, 18, 19]. It was found that the amount of the target is negative to the success of the project while the projects that attract the interest of investors are more likely to have a higher success rate of a crowdfunding [16, 18]. Some researches claimed frequent updates, and a high level of communication through updates and the information about the investors type also caused great influence on the success of crowdfunding [20, 21].

There are few researches focusing on the production and operation problem of the firms in crowdfunding when the investor's product valuation is changing over time. The preliminary researches have focused on analyzing the effects

of the information on investor's perception in financial market, and accordingly some studies have explored the effects of the investor's perception on the investment decisions. It was found that the low risk of fund and the liquidity of fund scheme have a great impact on investor's perception for investing in the mutual fund [22]. Joshi and Patel [23] found that the adverse investor's perception has prevented the information being reflected into the prices of stock. Investors may form their own valuations about the product according to their previous purchase experience and the information they mastered in the market. How to produce the products with a reasonable product quality to attract investors to participate in a crowdfunding is our focus. Hu [24] studied the optimal product pricing strategy in a reward-based crowdfunding, and the product line was proved to be more optimal than providing a single product. Based on the study of Hu, Chen [25] proposed the optimal pricing strategy for the sponsor when facing strategic and myopic consumers in a crowdfunding of green product.

Based on the above studies, we give the dynamic model to describe how to design the product quality when investor's perception is taken into account in the next following sections.

## 3. The Basic Model

The decision system considered in the present study consists of a sponsor and a platform. The sponsor raises funds through the crowdfunding on the platform and provides some rewards to the platform when the raised funds reach the amount of target. Different from the real sales, investors understand the product based on the online information they perceived from the platform. In crowdfunding, sponsor designs the product quality to cater investor's preference, and platform makes decisions about the advertisement of the project. Therefore, investor's perception is considered as the function of the product quality and the platform advertisement in the present study. For understanding easily, the description of parameters is illustrated in Table 1.

Since the information updates continuously, investor's perception changes over time. In detail, investment desire and reference price are introduced to describe investor's perception. First, the changing of the investment desire follows the Nerlove-Arrow framework [26]:

$$\begin{aligned} \dot{g}(t) &= \theta_1 q + \theta_2 e - \delta g, \\ g(0) &= g_0, \end{aligned} \quad (1)$$

where  $g(t)$  is the accumulated investment desire over time  $t$  and  $g(0) > 0$  is the initial investment desire.  $\theta_1$  and  $\theta_2$  are positive constants that reflect both the product quality and platform advertisement which have an positive effects on the investment desire, and  $\delta$  is the diminishing rate of investment desire.

Then, product quality and platform advertisement can also affect investor's reference price. The reference price is formed after investors get the information about the product attributes. The changeable reference price in the present study is supposed to be affected by product quality,

TABLE 1: Parameters description.

Symbol	Explanation	Variable type
$q$	The quality of product	Decision variable
$e$	The advertisement of platform	Decision variable
$V_t$	Investor's product valuations ( $V_t=H, L$ )	-
$g$	Investment desire	-
$r$	Reference price	-
$d$	The demand of product	-
$\Pi_s$	The profit of the sponsor	-
$\Pi_p$	The profit of the platform	-
$J_s$	The present value of sponsor's profit	Decision goal
$J_p$	The present value of platform's profit	Decision goal
$\varepsilon$	The ratio of the profit-sharing	-
$\phi$	The ratio of the cost-sharing	Decision variable
$k$	The probability of success	-
$m, \theta_1, \theta_2, \delta, \mu_1, \mu_2, \lambda_1, \lambda_2, \alpha, \beta$	Parameters	-

platform advertisement, and product price. The following equation is introduced to illustrate the changeable procession of reference price:

$$\begin{aligned} \dot{r}(t) &= \beta(p - r) + \mu_1 q + \mu_2 e, \\ r(0) &= r_0, \end{aligned} \quad (2)$$

where  $r(0) = r_0$  is the initial reference price and  $\beta, \mu_1, \mu_2$  are all constants. The item  $\beta(p - r)$  means the price influence; a higher  $\beta$  implies that the investors have short memory of the product. The two items  $\mu_1 q$  and  $\mu_2 e$  represent the influences of the product quality and the platform advertisement on the reference price. Higher product quality and platform advertisement can enhance the consumer's valuations on the product, so  $\mu_1 > 0$  and  $\mu_2 > 0$  are supposed.

Further, investors are divided into short and long sightedness which means that investors have less or deep impression on the product purchase experience in the past. Based on the research of *Hu* [24], combining the consumer's type, i.e., short or long sightedness, we give the following equation to illustrate the different investor's product evaluations:

$$V_t = \begin{cases} H & \text{with probability } \alpha(1 - \beta), \\ L & \text{with probability } \alpha\beta, \end{cases} \quad (3)$$

where  $H > L > 0$  represents the product valuations of the investor.

The price of product is determined by the sponsor at the beginning of the crowdfunding, and then the platform advertises the product online. This paper focuses on choosing the optimal product quality and platform advertisement when investor's perception is taken into account. The price of product  $p$  is assumed as a fixed value when the model is solved, and the different results caused by the changes of price will be discussed based on the research of *Hu* [24]. *Hu* provided several price strategies for the product crowdfunding; accordingly the optimal product quality and platform advertisement are analyzed in the present study based on his research.

Generally, the reference price, the product quality, and the platform advertisement are assumed to have positive effects on the sale. Thus, the demand  $d(t)$  of the product in a crowdfunding is supposed to satisfy the following equation:

$$d(t) = m(r - p) + g + \lambda_1 q + \lambda_2 e, \quad (4)$$

where  $m, \lambda_1$ , and  $\lambda_2$  are all positive constants. The item  $m(r - p)$  represents the influence of reference price on demand. The influence on the demand is positive when  $r > p$  and the influence is negative otherwise. A high  $m$  implies that the investors are more sensitive to the gap between the reference price and the real price. The other two items  $\lambda_1 q$  and  $\lambda_2 e$  indicate that both the product quality and the platform advertisement have positive influences on the demand.

Similar to the previous research of Jørgensen [27], the cost is in the form of quadratic function. The product cost of the sponsor is  $Cs = (1/2)q^2$ , and the cost of the platform is  $Cp = (1/2)e^2$ . Platforms can get some rewards when the raised funds reach the goal amount while there are no rewards provided to the platform if the crowdfunding fails. Therefore, it is advisable for the sponsor to share a part of the platform's cost to encourage the platform to make their efforts to improve the crowdfunding.  $\phi \in (0, 1)$  is the cost-sharing ratio that means the sponsor is willing to undertake the platform's advertisement cost, then the profit of the sponsor is

$$\Pi_s = kp(1 - \varepsilon)d - \frac{1}{2}q^2 - \frac{1}{2}\phi e^2, \quad (5)$$

where  $k \in (0, 1)$  is the probability that the raised funds reach the goal amount which is often called the successful probability of a crowdfunding.  $\varepsilon \in (0, 1)$  is called the ratio of profit sharing. So the profit of the platform is

$$\Pi_p = kp\varepsilon d - \frac{1}{2}(1 - \phi)e^2, \quad (6)$$

where the item  $k\varepsilon p d$  ( $\varepsilon \in (0, 1)$ ) is the reward of the platform when the crowdfunding raised the goal amount of funds.

The goals of the sponsor and the platform are to maximize the present value of their profit, respectively,

$$\begin{aligned} \max_{d(t)} J_s & \\ &= \int_0^{+\infty} \exp(-\rho t) \left[ kp(1-\varepsilon)d(t) - \frac{1}{2}q^2 - \frac{1}{2}\phi e^2 \right] dt, \end{aligned} \quad (7)$$

and

$$\begin{aligned} \max_{d(t)} J_p & \\ &= \int_0^{+\infty} \exp(-\rho t) \left[ kp\varepsilon d(t) - \frac{1}{2}(1-\phi)e^2 \right] dt, \end{aligned} \quad (8)$$

where  $\rho$  is the discount rate.

In the next sections, the optimal product quality and platform advertisement are calculated in two different scenarios. One scenario is that the sponsor and the platform make decisions independently, and the other is that the sponsor and the platform are coordinate as a system to make decisions.

**3.1. The Optimal Decisions in Separation.** Separation means that the sponsor and the platform make decisions independently; the sponsor offers a cost-sharing ratio  $\phi$  firstly to the platform. Then the sponsor decides their product quality and the platform decides their advertisement to maximize their profit. It is reasonable to suppose that the sponsor and the platform make decisions simultaneously since the demand changes over time, and the cost-sharing ratio  $\phi$  is assumed as a fixed value when we calculate the values of decisions about the product quality and platform advertisement. The profit of the two firms will be calculated after we obtain the optimal equilibrium value of decisions, and then the optimal cost-sharing ratio  $\phi$  that maximizes the sponsor's profit is calculated.

To maximize the profit of the sponsor and the platform, the function of Hamiltonian for the sponsor and the platform are given as the follows, respectively,

$$\begin{aligned} H_s &= kp(1-\varepsilon) [m(r-p) + g + \lambda_1 q + \lambda_2 e] - \frac{1}{2}q^2 \\ &\quad - \frac{1}{2}\phi e^2 + \gamma_{1s}(\beta(p-r) + \mu_1 q + \mu_2 e) \\ &\quad + \gamma_{2s}(\theta_1 q + \theta_2 e - \delta g), \end{aligned} \quad (9)$$

and

$$\begin{aligned} H_p &= kp\varepsilon [m(r-p) + g + \lambda_1 q + \lambda_2 e] - \frac{1}{2}(1-\phi)e^2 \\ &\quad + \gamma_{1p}(\beta(p-r) + \mu_1 q + \mu_2 e) \\ &\quad + \gamma_{2p}(\theta_1 q + \theta_2 e - \delta g), \end{aligned} \quad (10)$$

where  $\gamma_{1s}, \gamma_{2s}(\gamma_{1p}, \gamma_{2p})$  represent the costate variables.

Calculating the two Hamiltonian functions, the equilibrium product quality and platform advertisement are given in the following Theorem 1.

**Theorem 1.** *The equilibrium product quality is*

$$\bar{q} = kp(1-\varepsilon) \left( \lambda_1 + \frac{\theta_1}{\rho + \delta} + \frac{m\mu_1}{\rho + \beta} \right), \quad (11)$$

and the equilibrium platform advertisement is

$$\bar{e} = \frac{kp\varepsilon}{1-\phi} \left( \lambda_2 + \frac{\theta_2}{\rho + \delta} + \frac{m\mu_2}{\rho + \beta} \right). \quad (12)$$

Therefore the reference price and accumulated investment desire can also be given as follows:

$$\begin{aligned} r(t) &= A_1 e^{-\beta t} + r_n, \\ g(t) &= A_2 e^{-\delta t} + g_n, \end{aligned} \quad (13)$$

where  $A_1 = r_0 - r_n, r_n = p + (\mu_1 \bar{q} + \mu_2 \bar{e})/\beta, A_2 = g_0 - g_n,$  and  $g_n = (\theta_1 \bar{q} + \theta_2 \bar{e})/\delta.$

The proof of each theorem is illustrated in Appendixes, and some conclusions can be addressed from Theorem 1 as follows:

(i) There are three parts of equilibrium product quality level in (11): the first part  $kp\lambda_1$  represents the product quality influences on the product demand which is the short-term influence on product sales. The second part  $kp\theta_1/(\rho + \delta)$  expresses the long-term effect that means the product quality has the positive influence in accumulate investment desire. The third part  $kpm\mu_1/(\rho + \beta)$  expresses the fact that the product quality has the effect on the reference price with  $\mu_1 \neq 0.$  The structure of (12) is the same as (11), so we do not repeat it here. Since the product quality and platform advertisement have positive effect on the reference price, i.e.,  $\mu_1 > 0$  and  $\mu_2 > 0,$  the sponsor will provide a higher quality product and the platform will put more advertisement in adverting when they take the effect of reference price into consideration. The higher  $m$  means the consumers are more sensitive to the reference price; therefore the decision makers should pay more attention to this impact.

(ii) The reference price and investment desire in (13) will get their steady states  $r_n$  and  $g_n$  when  $t \rightarrow +\infty.$  The steady states, i.e.,  $r_n = p + (\mu_1 \bar{q} + \mu_2 \bar{e})/\beta$  and  $g_n = (\theta_1 \bar{q} + \theta_2 \bar{e})/\delta,$  are mainly influenced by the product price  $p,$  the product quality  $\bar{q},$  and the platform advertisement  $\bar{e}.$  It is shown that the reference price and the investment desire are increasing when the product quality and platform advertisement are taken into account. In addition, investors will have a large reference price when the value of  $\beta$  is small. Analyzing  $g_n$  similarly, the investment desire is positively correlated with the product quality and the platform advertisement.

Based on the above results, the profits of the sponsor and platform are calculated as follows:

$$\begin{aligned} J_s &= kp(1-\varepsilon) \left( \frac{D_1 m}{\rho + \beta} + \frac{D_2}{\rho + \delta} + \frac{m(r_n - p)}{\rho} + \frac{g_n}{\rho} \right. \\ &\quad \left. + \frac{\lambda_1 \bar{q}}{\rho} + \frac{\lambda_2 \bar{e}}{\rho} \right) - \frac{\bar{q}^2}{2\rho} - \frac{\phi \bar{e}^2}{2\rho}, \end{aligned} \quad (14)$$

and

$$J_p = kp\varepsilon \left( \frac{D_1 m}{\rho + \beta} + \frac{D_2}{\rho + \delta} + \frac{m(r_n - p)}{\rho} + \frac{g_n}{\rho} + \frac{\lambda_1 \bar{q}}{p} + \frac{\lambda_2 \bar{e}}{\rho} \right) - \frac{(1 - \phi) \bar{e}^2}{2\rho}. \quad (15)$$

Since sponsors make the decisions about the product price strategy at the beginning of the crowdfunding, they will decide the optimal value of product quality and platform advertisement for any specific price strategy. The optimal results in margin (H) and volume (L) price strategy are analyzed, respectively, based on the research of *Hu* [24]. Therefore, there will be two groups, i.e.,  $(\bar{q}, \bar{e}, H)$ ,  $(\bar{q}, \bar{e}, L)$ , to choose different situations. In the following Theorem 2, the optimal product quality and platform advertisement are given in detail in different price strategy.

**Theorem 2.** (i) When

$$B_2 = (1 - \varepsilon)^2 A_1 + \frac{m\mu_2 A_2 \varepsilon (1 - \varepsilon) + \lambda_2 \beta A_2 \varepsilon (1 - \varepsilon)}{\rho \beta (1 - \phi)} + \frac{\theta_1 A_1 (1 - \varepsilon)^2 (1 - \phi) + \theta_1 A_2 \varepsilon}{\rho \delta (1 - \phi)} + \frac{2\lambda_1 A_1 (1 - \varepsilon)^2 - A_1^2 (1 - \varepsilon)^2}{2\rho} - \frac{A_2^2 \varepsilon^2}{(1 - \phi)^2} > 0, \quad (16)$$

we obtain the following two conclusions:

(a)  $0 < H/l < 1/\alpha^2(1 - \beta)^2$ ; the optimal values of product quality and platform advertisement are

$$\bar{q} = (1 - \varepsilon) L \left( \lambda_1 + \frac{\theta_1}{\rho + \delta} + \frac{m\mu_1}{\rho + \beta} \right), \quad (17)$$

and

$$\bar{e} = \frac{\varepsilon L}{1 - \phi} \left( \lambda_2 + \frac{\theta_2}{\rho + \delta} + \frac{m\mu_2}{\rho + \beta} \right). \quad (18)$$

(b)  $H/l > 1/\alpha^2(1 - \beta)^2$ ; the optimal values of product quality and platform advertisement are

$$\bar{q} = (1 - \varepsilon) H \alpha^2 (1 - \beta)^2 \left( \lambda_1 + \frac{\theta_1}{\rho + \delta} + \frac{m\mu_1}{\rho + \beta} \right), \quad (19)$$

and

$$\bar{e} = \frac{\varepsilon H \alpha^2 (1 - \beta)^2}{1 - \phi} \left( \lambda_2 + \frac{\theta_2}{\rho + \delta} + \frac{m\mu_2}{\rho + \beta} \right). \quad (20)$$

(ii) When

$$B_2 = (1 - \varepsilon)^2 A_1 + \frac{m\mu_2 A_2 \varepsilon (1 - \varepsilon) + \lambda_2 \beta A_2 \varepsilon (1 - \varepsilon)}{\rho \beta (1 - \phi)} + \frac{\theta_1 A_1 (1 - \varepsilon)^2 (1 - \phi) + \theta_1 A_2 \varepsilon}{\rho \delta (1 - \phi)} + \frac{2\lambda_1 A_1 (1 - \varepsilon)^2 - A_1^2 (1 - \varepsilon)^2}{2\rho} - \frac{A_2^2 \varepsilon^2}{(1 - \phi)^2} < 0, \quad (21)$$

the optimal values of product quality and platform advertisement are

$$\bar{q} = \frac{(\varepsilon - 1) B_1}{2B_2} \left( \lambda_1 + \frac{\theta_1}{\rho + \delta} + \frac{m\mu_1}{\rho + \beta} \right), \quad (22)$$

and

$$\bar{e} = \frac{\varepsilon(\varepsilon - 1) B_1}{2B_2(1 - \phi)} \left( \lambda_2 + \frac{\theta_2}{\rho + \delta} + \frac{m\mu_2}{\rho + \beta} \right). \quad (23)$$

Note that the equilibrium value of product quality and platform advertisement will be higher if the sponsor takes the margin price strategy compared to the volume price strategy. It is also clear that both two equilibriums are the increasing function of the consumer's evaluation ( $L$  or  $H$ ). When the evaluation of consumer is higher the sponsor should improve their product quality level and similarly the platform should pay more advertisement to advertise the product. And there are no differences of the sponsor profit in the two price strategies when  $H/l = 1/\alpha^2(1 - \beta)^2$ .

**3.2. The Optimal Decisions in Coordination.** In this section, the sponsor and the platform cooperate as a system to make decisions about product quality and platform advertisement. We calculate the optimal decision of the system to maximize the present value of it. Taking (5) and (6) in to account, the Hamiltonian function for the system is given as follows:

$$H = pk [m(r - p) + g + \lambda_1 q + \lambda_2 e] - \frac{1}{2} q^2 - \frac{1}{2} e^2 + \gamma_1 (\beta(p - r) + \mu_1 q + \mu_2 e) + \gamma_2 (\theta_1 q + \theta_2 e - \delta g), \quad (24)$$

where  $\gamma_1, \gamma_2$  represent the costate variables of the system problem associating the reference price and investment desire with the decisions. Similar to the process of Theorem 1, the equilibrium of the system is shown as follows:

$$\bar{q} = kp \left( \lambda_1 + \frac{\theta_1}{\rho + \delta} + \frac{m\mu_1}{\rho + \beta} \right), \quad (25)$$

and

$$\bar{e} = kp \left( \lambda_2 + \frac{\theta_2}{\rho + \delta} + \frac{m\mu_2}{\rho + \beta} \right). \quad (26)$$

It is clearly shown that the product quality is improved due to the cooperation. In Theorem 3, the optimal cost-sharing ratio is given and the advertisement of platform in two situations is compared.

**Theorem 3.** The optimal cost-sharing ratio  $\phi$  is

$$\phi = \begin{cases} \frac{2 - 3\varepsilon}{2 - \varepsilon} & \text{if } \varepsilon \leq \frac{2}{3}, \\ 0 & \text{else.} \end{cases} \quad (27)$$

Taking (27) into (20), we get

$$\bar{e} = kp \frac{\varepsilon(2 - \varepsilon)}{2\varepsilon} \left( \lambda_2 + \frac{\theta_2}{\rho + \delta} + \frac{m\mu_2}{\rho + \beta} \right), \quad (28)$$

and

$$\check{\varepsilon} - \bar{\varepsilon} = \frac{kp\varepsilon}{2} \left( \lambda_2 + \frac{\theta_2}{\rho + \delta} + \frac{m\mu_2}{\rho + \beta} \right) > 0. \quad (29)$$

For the cooperation system, the reference price and investment desire are also steady similar to the situation of separation as follows:

$$\begin{aligned} r(t) &= E_1 e^{-\beta t} + r_c, \\ g(t) &= E_2 e^{-\delta t} + g_c, \end{aligned} \quad (30)$$

where  $E_1 = r_0 - r_c$ ,  $r_c = p + (\mu_1 \check{q} + \mu_2 \check{\varepsilon})/\beta$ ,  $E_2 = g_0 - g_c$ , and  $g_c = (\theta_1 \check{q} + \theta_2 \check{\varepsilon})/\delta$ . The steady states of the reference price and investment desire are  $r_c$  and  $g_c$ .

From the above results, we can find that both the steady reference price and steady investment desire are higher compared to Theorem 2 since both the product quality and platform advertisement are improved in cooperation system. Substituting (30) into (7), we get the present value of the profit for cooperation system as follows:

$$\begin{aligned} \check{J} &= kp \left( \frac{E_1 m}{\rho + \beta} + \frac{E_2}{\rho + \delta} + \frac{m(r_c - p)}{\rho} + \frac{g_c}{\rho} + \frac{\lambda_1 \check{q}}{p} \right. \\ &\quad \left. + \frac{\lambda_2 \check{\varepsilon}}{\rho} \right) - \frac{\check{q}^2}{2\rho} - \frac{\check{\varepsilon}^2}{2\rho}. \end{aligned} \quad (31)$$

In the next section we will discuss the changes of the results in our model when relevant parameters are in different values. Analyzing these changes does not only build the robustness of our results but also deepen our understanding of this problem.

#### 4. Parameters Sensitivity Analysis

In this section, we analyze the relations between the relevant parameters and the values of profit, product quality, and platform advertisement. Firstly, the impact of parameter  $m$  on the present value of profit is illustrated in Figures 2 and 3. The parameter  $m$  is seen as an item that reflects the degree of the reference price effects on the investor's behavior. There are two firms: the sponsor and the platform in the research, so it is worthwhile to analyze the influence of parameter  $m$  on the profit of them because the effects of  $m$  may be different.

As shown in Figure 2,  $J_s$ ,  $J_p$ , and  $J$  represent the profit of the sponsor, the profit of the platform, and the total profit when the sponsor and the platform chose to cooperate, respectively. The present value of the sponsor's profit will decrease if the investors pay more attention to reference price while the platform's profit is an increasing function of parameter  $m$  when they make decisions independently. It is clear that the profits of the two firms are increased when they cooperate as a system, and the cooperation profit is an increasing function of the parameter  $m$ . Note that parameter  $m$  denotes the degree of the reference price effects on the investor's behavior; the fact implied in Figure 2 shows that the larger the effect of the reference price, the more profit the two items can get when they cooperate as a system.

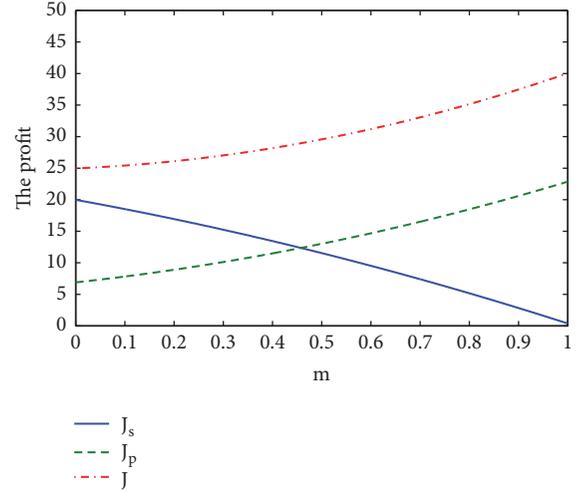


FIGURE 2: Profit changes.

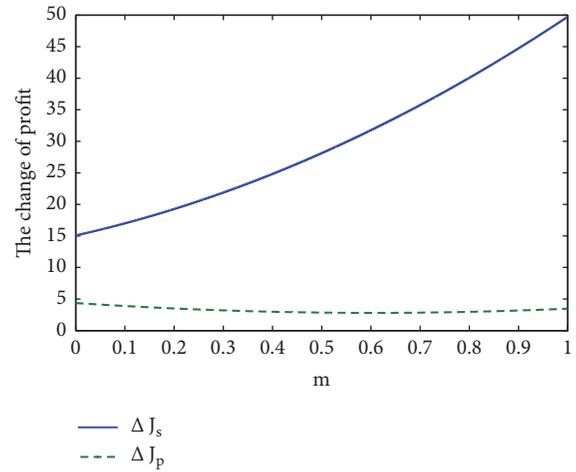


FIGURE 3: Differences of profit in two situations.

From Figure 3, we can find that cooperation is more profitable for the sponsor than the platform and the differences of the two profits are both the increasing function of parameter  $m$ . Therefore, the sponsor should choose to cooperate with the platform when the investors pay more attention to their reference price. And the differences of platform's profit are not obvious in two ways, which means that it has a little difference for platform in two cases.

Comparing Figure 2 with Figure 3, it is found that cooperation with platform not only promotes the profit of sponsor but also can help the sponsor to keep a growing trend of the profit when investors pay more attention to their reference price.

Then, the impact of parameters  $\varepsilon, \phi$  on the values of product quality and platform advertisement are illustrated in Figures 4 and 5.

As shown in Figures 4 and 5, it is clear that the platform will increase their advertisement to improve crowdfunding project with the profit-sharing ratio increasing. And this trend is more obvious when the value of cost-sharing ratio is

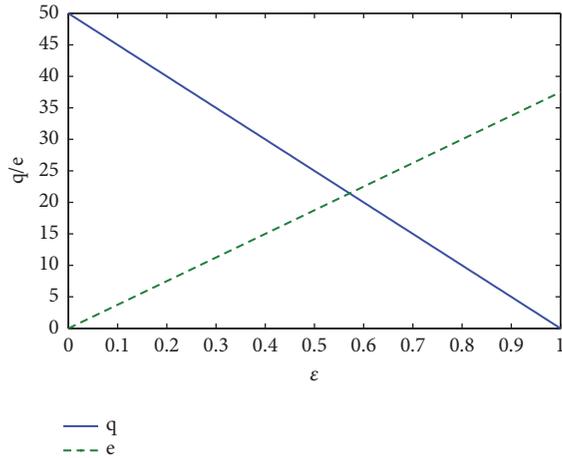


FIGURE 4: The changes of product quality (platform advertisement).

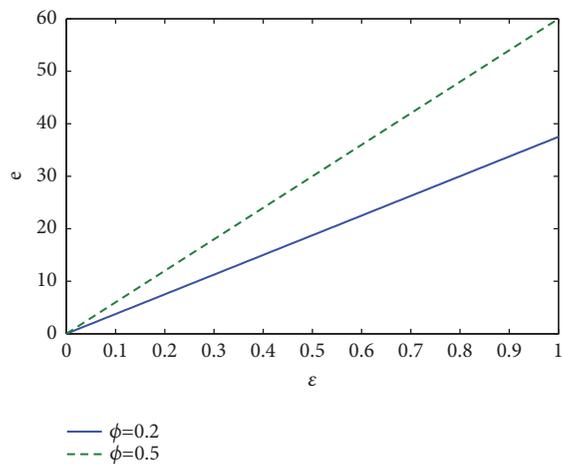


FIGURE 5: The changes of platform advertisement in two situations.

TABLE 2: Impacts of  $\epsilon, \phi$  on the optimal results.

parameters	$J_s$	$J_p$	$e$	$q$
$(\epsilon, \phi)$ (0.03, 0.10)	871.6	906.3	1	48.5
(0.03, 0.20)	2951.5	903.6	1.125	48.5
(0.10, 0.20)	3955.5	781.1	3.75	45
(0.20, 0.30)	8017.5	583.1	8.6	40

higher. It is also clear that the advertisement of platform has a complementary effect with product quality. In Table 2, the impact of  $\epsilon, \phi$  on the optimal results is illustrated in detail.

As shown in Table 2, the present value of the sponsor’s profit is an increasing function of  $\epsilon$  and  $\phi$ , and the trend is more obvious if both the two ratios increase simultaneously. And the advertisement of platform will also increase if the sponsor shares more cost or profit with them. The above results illustrate that sponsor sharing the cost of the platform can not only promote the development of crowdfunding but also improve its own profit.

### 5. Simulation

To demonstrate some external validity for our theoretical model, an empirical study about the impact of product quality and platform advertisement on the success of a crowdfunding project is conducted. Data for our study is derived from publicly available information on the Kickstarter.com website (<https://www.kickstarter.com/>). Kickstarter has grown into one of the best-known and oldest crowdfunding platforms in the world since its inception in April 2009. We write a computer program by using the language of Python to get the information on all projects posted on the platform. Over six months, the information for 1400 projects from Kickstarter.com has been collected. The projects for charity are removed because charity crowdfunding projects are donation-based crowdfunding. And some data for inaccuracies and incomplete information are also removed. After that, the data for analysis purposes included 835 crowdfunding projects in the categories art, design, and technology.

Data about a given project includes the money which have been raised, the number of comments, and the introduction of the project. The introduction of the product includes the number of introduction pictures, the number of videos, and the number of description words. In the present study, the description and introduction of the project are defined as the factors that relate to the quality of the product. And the number of comments is seen as an external influence.

The correlation analysis of the variables is first performed, and results of which are shown in Table 3. It can be seen from the table that correlations between the dependent variable and all independent variables are positive. Therefore, the assumptions that the information about the product quality and platform advertisement will cause some influences on the success of crowdfunding were supported. In Table 4, some results of a regression analysis are illustrated in detail. We choose the signal of introduction picture count to represent the signal of product quality, and the signal of comments count is seen as the signal of external influences. As can be seen from the table, the introduction picture count explains 44.6% of the raised money and the comment count explains 43.5% of the raised money. The F value is 42.462, and the values of VIF ( $<10$ ) for each variable all indicate that there is no multicollinearity in the model. The results of the numerical experiment support the theory and model analyzed in the present study and provide some external validity for our theoretical model.

### 6. Conclusion

Product design allows firms to improve their profit and market competitiveness in crowdfunding market. However, the successful crowdfunding is always hampered by lack of guidance on how to design the product quality to attract investors’ preference. To solve this dilemma, we describe two items: investment desire and reference price to represent investor’s perception that can cause some effects on the success of crowdfunding.

A necessary consideration that investor’s investment desire and reference price have a significant effect on investor

TABLE 3: Correlations of variables.

	1	2	3	4	5
(1) Raised money (\$)	1				
(2) No. of comment	0.6624	1			
(3) Introduction picture count	0.3943	0.4637	1		
(4) Introduction video count	0.3621	0.3889	0.3556	1	
(5) Introduction word count	0.6571	0.4960	0.5173	0.3741	1

TABLE 4: Results of regression analysis.

Variable		Total sample	
			$\beta$
Dependent variable	Raised money (\$)		VIF
Signals of platforms advertisement	No. of comment		0.435
Signals of product quality	Introduction picture count		0.446
	F	42.462	
Summary of the model specified	$R^2$	0.69	
	$\Delta R^2$	0.687	

behavior is shown in our paper. Further, the product quality and platform advertisement which are supposed have some effects on investor's investment desire and reference price. This supposition is different from the previous literature that they assumed that there are only two types of investors: high evaluation and low evaluation investors. Accordingly, a dynamic model about the product quality is investigated by taking the investor's perception into account.

In this paper, the optimal product quality and platform advertisement are calculated in two situations: separation and cooperation, respectively. The results offer some new insights that will be useful to firms to research the product quality decisions in crowdfunding considering investor's perception. First, firms should pay more attention to the potential behavior of investors influenced by online information when they make decisions about the product quality. Then, it is an optimal strategy for firms to cooperate with platform when investors pay more attention to their reference price. Finally, sharing the cost of the platform can not only promote the development of crowdfunding but also improve the profit of firms.

## Appendix

### A. Proof of Theorem 1

In model, the necessary conditions for equilibrium of the sponsor profit are given by

$$\frac{\partial H_s}{\partial q} = 0, \tag{A.1}$$

$$\frac{\partial H_s}{\partial \gamma_{1s}} = \dot{r}(t), \tag{A.2}$$

$$\frac{\partial H_s}{\partial \gamma_{2s}} = \dot{g}(t), \tag{A.3}$$

$$\dot{\gamma}_{1s} = \rho \gamma_{1s} - \frac{\partial H_s}{\partial r}, \tag{A.4}$$

$$\dot{\gamma}_{2s} = \rho \gamma_{2s} - \frac{\partial H_s}{\partial g}. \tag{A.5}$$

From (A.1), we get

$$q = kp\lambda_1 + \mu_1\gamma_{1s} + \theta_1\gamma_{2s}. \tag{A.6}$$

Combining (A.4), (A.5), and (A.6), we can get

$$\dot{q} = (\rho + \beta) \mu_1\gamma_{1s} + (\rho + \delta) \theta_1\gamma_{2s} - kpm\mu_1 - kp\theta_1. \tag{A.7}$$

Substituting  $\mu_1\gamma_{1s}$  into (A.7) by (A.6), we get

$$\dot{q} = (\rho + \beta) q + (\delta - \beta) \theta_1\gamma_{2s} - (\rho + \beta) kp\lambda_1 - kpm\mu_1 - kp\theta_1. \tag{A.8}$$

Similarly, combining (A.6), (A.8), we get the differential of (A.8) as follows:

$$\begin{aligned} \ddot{q} &= (2\rho + \beta + \delta) \dot{q} - (\rho + \delta) (\rho + \beta) q \\ &\quad + (\rho + \delta) (\rho + \beta) kp\lambda_1 + (\rho + \delta) kpm\mu_1 \\ &\quad + (\rho + \beta) kp\theta_1. \end{aligned} \tag{A.9}$$

Calculating (A.9), we get

$$q(t) = D_1 e^{(\rho+\delta)t} + D_2 e^{(\rho+\beta)t} + \bar{q}, \tag{A.10}$$

where  $\bar{q} = kp(\lambda_1 + \theta_1/(\rho + \delta) + m\mu_1/(\rho + \beta))$  and  $D_1, D_2$  are parameters to be determined. The value of  $q$  will be infinite when  $t \rightarrow \infty$  which will obey our rules, so we have  $D_1 = D_2 = 0$ . Therefore, the equilibrium value of  $q(t)$  can expressed as follows:

$$\bar{q} = kp \left( \lambda_1 + \frac{\theta_1}{\rho + \delta} + \frac{m\mu_1}{\rho + \beta} \right). \tag{A.11}$$

Similarly, the equilibrium platform advertisement can be got as follows:

$$\bar{e} = \frac{\varepsilon kp}{1-\phi} \left( \lambda_2 + \frac{\theta_2}{\rho+\delta} + \frac{m\mu_2}{\rho+\beta} \right). \quad (\text{A.12})$$

Taking (A.11), (A.12) into, we get

$$\frac{\partial g(t)}{\partial t} = \theta_1 \bar{q} + \theta_2 \bar{e} - \delta g, \quad (\text{A.13})$$

and

$$\frac{\partial r(t)}{\partial t} = \beta(p-r) + \mu_1 \bar{q} + \mu_2 \bar{e}. \quad (\text{A.14})$$

Solving the differential equations (A.13), (A.14), the general solutions can be got

$$\begin{aligned} r(t) &= E_1 e^{-\beta t} + r_n, \\ g(t) &= E_2 e^{-\delta t} + g_n, \end{aligned} \quad (\text{A.15})$$

where  $E_1 = r_0 - r_n$ ,  $r_n = p + (\mu_1 \bar{q} + \mu_2 \bar{e})/\beta$ ,  $E_2 = g_0 - g_n$ , and  $g_n = (\theta_1 \bar{q} + \theta_2 \bar{e})/\delta$ .

## B. Proof of Theorem 2

Based on the research of *Hu*, with a volume strategy, the sponsor sets the price  $p = L$  which is the lower limit of consumer valuation. Since the price is the lower limits of consumer valuation, we suppose that all consumers will take part in the crowdfunding and the success rate  $k = 1$ . With the margin strategy, the sponsor sets the price  $p = H$ , the high-type consumers will buy the product, and the low-type consumer will decline. The success rate in margin strategy is  $k = \alpha^2(1-\beta)^2$ . Sponsor will make decisions about the product price, so we compare the profit of the sponsor in different price strategy to maximize sponsor's profit as follows. From (7), we can get the fact that the profit of the sponsor is

$$\begin{aligned} J_s &= kp(1-\varepsilon) \left( \frac{D_1 m}{\rho+\beta} + \frac{D_2}{\rho+\delta} + \frac{m(r_n-p)}{\rho} + \frac{g_n}{\rho} \right. \\ &\quad \left. + \frac{\lambda_1 \bar{q}}{\rho} + \frac{\lambda_2 \bar{e}}{\rho} \right) - \frac{\bar{q}^2}{2\rho} - \frac{\phi \bar{e}^2}{2\rho}. \end{aligned} \quad (\text{B.1})$$

Taking (25) and (26) in to (B.1), we can get

$$\begin{aligned} J_s &= kp(1-\varepsilon) B_1 + (kp)^2 \left( (1-\varepsilon)^2 A_1 \right. \\ &\quad \left. + \frac{m\mu_2 A_2 \varepsilon (1-\varepsilon) + \lambda_2 \beta A_2 \varepsilon (1-\varepsilon)}{\rho\beta(1-\phi)} \right. \\ &\quad \left. + \frac{\theta_1 A_1 (1-\varepsilon)^2 (1-\phi) + \theta_1 A_2 \varepsilon}{\rho\delta(1-\phi)} \right. \\ &\quad \left. + \frac{2\lambda_1 A_1 (1-\varepsilon)^2 - A_1^2 (1-\varepsilon)^2}{2\rho} - \frac{A_2^2 \varepsilon^2}{(1-\phi)^2} \right), \end{aligned} \quad (\text{B.2})$$

where  $A_1 = \lambda_1 + \theta_1/(\rho+\delta) + m\mu_1/(\rho+\beta) > 0$ ,  $A_2 = \lambda_2 + \theta_2/(\rho+\delta) + m\mu_2/(\rho+\beta) > 0$ , and  $B_1 = D_1 m/(\rho+\beta) + D_2/(\rho+\delta) > 0$ .

(i) When the coefficient of the quadratic term is positive,

$$\begin{aligned} (1-\varepsilon)^2 A_1 + \frac{m\mu_2 A_2 \varepsilon (1-\varepsilon) + \lambda_2 \beta A_2 \varepsilon (1-\varepsilon)}{\rho\beta(1-\phi)} \\ + \frac{\theta_1 A_1 (1-\varepsilon)^2 (1-\phi) + \theta_1 A_2 \varepsilon}{\rho\delta(1-\phi)} \\ + \frac{2\lambda_1 A_1 (1-\varepsilon)^2 - A_1^2 (1-\varepsilon)^2}{2\rho} - \frac{A_2^2 \varepsilon^2}{(1-\phi)^2} \\ > 0. \end{aligned} \quad (\text{B.3})$$

The value of inequity (B.3) is positive which means that the function of (B.2) is the monotonically increasing function of the item  $kp$  when  $kp > 0$ , so the following can be got easily:

(a) When  $0 < H/l < 1/\alpha^2(1-\beta)^2$ , we get  $L > H\alpha^2(1-\beta)^2$ , so the volume strategy is better than margin strategy for the sponsor and the price is  $p = L$ . Taking  $p = L$ ,  $k = 1$  into (19) and (20), we can get

$$\bar{q} = (1-\varepsilon)L \left( \lambda_1 + \frac{\theta_1}{\rho+\delta} + \frac{m\mu_1}{\rho+\beta} \right), \quad (\text{B.4})$$

and

$$\bar{e} = \frac{\varepsilon L}{1-\phi} \left( \lambda_2 + \frac{\theta_2}{\rho+\delta} + \frac{m\mu_2}{\rho+\beta} \right). \quad (\text{B.5})$$

(b) When  $H/l > 1/\alpha^2(1-\beta)^2$ , we get  $L < H\alpha^2(1-\beta)^2$ , and the margin strategy is better than the volume strategy. Similarly, the optimal values of product quality and platform advertisement are

$$\bar{q} = (1-\varepsilon)H\alpha^2(1-\beta)^2 \left( \lambda_1 + \frac{\theta_1}{\rho+\delta} + \frac{m\mu_1}{\rho+\beta} \right), \quad (\text{B.6})$$

and

$$\bar{e} = \frac{\varepsilon H\alpha^2(1-\beta)^2}{1-\phi} \left( \lambda_2 + \frac{\theta_2}{\rho+\delta} + \frac{m\mu_2}{\rho+\beta} \right). \quad (\text{B.7})$$

(ii) When the coefficient of the quadratic term is negative, that is,

$$\begin{aligned} B_2 &= (1-\varepsilon)^2 A_1 \\ &\quad + \frac{m\mu_2 A_2 \varepsilon (1-\varepsilon) + \lambda_2 \beta A_2 \varepsilon (1-\varepsilon)}{\rho\beta(1-\phi)} \\ &\quad + \frac{\theta_1 A_1 (1-\varepsilon)^2 (1-\phi) + \theta_1 A_2 \varepsilon}{\rho\delta(1-\phi)} \\ &\quad + \frac{2\lambda_1 A_1 (1-\varepsilon)^2 - A_1^2 (1-\varepsilon)^2}{2\rho} - \frac{A_2^2 \varepsilon^2}{(1-\phi)^2} \\ &< 0, \end{aligned} \quad (\text{B.8})$$

the optimal price is constant:

$$kp = \frac{(\varepsilon-1)B_1}{2B_2}, \quad (\text{B.9})$$

where  $B_1 = D_1 m / (\rho + \beta) + D_2 / (\rho + \delta) > 0$ . Taking (B.4) into (11) and (12), we get

$$\bar{q} = \frac{(\varepsilon - 1) B_1}{2B_2} \left( \lambda_1 + \frac{\theta_1}{\rho + \delta} + \frac{m\mu_1}{\rho + \beta} \right), \quad (\text{B.10})$$

and

$$\bar{e} = \frac{\varepsilon(\varepsilon - 1) B_1}{2B_2(1 - \phi)} \left( \lambda_2 + \frac{\theta_2}{\rho + \delta} + \frac{m\mu_2}{\rho + \beta} \right). \quad (\text{B.11})$$

### C. Proof of Theorem 3

Since the sponsor decides the optimal  $\phi$  to get the present value of his profit, by the first-order optimal condition

$$\frac{dJ_s}{d\phi} = 0 \quad (\text{C.1})$$

we get

$$\phi = \frac{2 - 3\varepsilon}{2 - \varepsilon}, \quad (\text{C.2})$$

and combining  $\phi \in [0, 1]$ , we get

$$\phi = \begin{cases} \frac{2 - 3\varepsilon}{2 - \varepsilon} & \text{if } \varepsilon \leq \frac{2}{3}, \\ 0 & \text{else.} \end{cases} \quad (\text{C.3})$$

### Data Availability

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

### Conflicts of Interest

The authors declare no conflicts of interest.

### Acknowledgments

The authors gratefully acknowledge the support from the National Social Science Foundation of China (17BGL083).

### References

- [1] S. Bi, Z. Liu, and K. Usman, "The influence of online information on investing decisions of reward-based crowdfunding," *Journal of Business Research*, vol. 71, pp. 10–18, 2017.
- [2] A. Agrawal, C. Catalini, and A. Goldfarb, "Some simple economics of crowdfunding," *Innovation Policy and the Economy*, vol. 14, no. 1, pp. 63–97, 2014.
- [3] P. Belleflamme, T. Lambert, and A. Schwienbacher, "Crowdfunding: Tapping the right crowd," *Journal of Business Venturing*, vol. 29, no. 5, pp. 585–609, 2014.
- [4] M. Cholakova and B. Clarysse, "Does the Possibility to Make Equity Investments in Crowdfunding Projects Crowd Out Reward-Based Investments?" *Entrepreneurship Theory and Practice*, vol. 39, no. 1, pp. 145–172, 2015.
- [5] E. Mollick, "The dynamics of crowdfunding: An exploratory study," *Journal of Business Venturing*, vol. 29, no. 1, pp. 1–16, 2014.
- [6] S. P. Anderson and L. Celik, "Product line design," *Journal of Economic Theory*, vol. 157, pp. 517–526, 2015.
- [7] L. Bottazzi and M. Da Rin, "Venture capital in Europe and the financing of innovative companies," *Economic Policy*, no. 34, pp. 229–269, 2002.
- [8] D. Irwin and J. M. Scott, "Barriers faced by SMEs in raising bank finance," *International Journal of Entrepreneurial Behavior and Research*, vol. 16, no. 3, pp. 245–259, 2010.
- [9] A. Schwienbacher and B. Larralde, "Crowdfunding of Small Entrepreneurial Ventures," *SSRN Electronic Journal*, 2010.
- [10] F. Kleeman, G. G. Voss, and K. Rieder, "Under paid innovators: The commercial utilization of consumer work through crowdsourcing Science," *Science, Technology and Innovation Studies*, vol. 4, no. 1, pp. 5–26, 2011.
- [11] B. Armendáriz and J. Morduch, *The Economics of Microfinance*, vol. 1, Mit Press Books, 2nd edition, 2010.
- [12] J. M. Leimeister, "Crowdsourcing: Crowdfunding, crowdvoting, crowdcreation," *Zeitschrift für Controlling und Management*, vol. 56, pp. 388–392, 2012.
- [13] A. Tomczak and A. Brem, "A conceptualized investment model of crowdfunding," *Venture Capital*, vol. 15, no. 4, pp. 335–359, 2013.
- [14] M. Cho and G. Kim, "A cross-cultural comparative analysis of crowdfunding projects in the United States and South Korea," *Computers in Human Behavior*, vol. 72, pp. 312–320, 2017.
- [15] A. Lukkarinen, J. E. Teich, H. Wallenius, and J. Wallenius, "Success drivers of online equity crowdfunding campaigns," *Decision Support Systems*, vol. 87, pp. 26–38, 2016.
- [16] T. H. Allison, B. C. Davis, J. C. Short, and J. W. Webb, "Crowdfunding in a prosocial microlending environment: Examining the role of intrinsic versus extrinsic cues," *Entrepreneurship Theory and Practice*, vol. 39, no. 1, pp. 53–73, 2015.
- [17] D. Frydrych, A. J. Bock, T. Kinder, and B. Koeck, "Exploring entrepreneurial legitimacy in reward-based crowdfunding," *Venture Capital*, vol. 16, no. 3, pp. 247–269, 2014.
- [18] K. Hyeog-In, S. Lee, and Y. Na, "Success and failure case analysis of crowdfunding: Focused on community art project," *The Journal of the Korea Contents Association*, vol. 14, no. 7, pp. 125–136, 2014.
- [19] V. Alfredo, "Crowdfunding in Arts and Culture: from funding to engagement," *Economia della Cultura*, vol. 1, pp. 121–130, 2017.
- [20] M. Beier and K. Wagner, "Crowdfunding Success of Tourism Projects - Evidence from Switzerland," *SSRN Electronic Journal*, 2014.
- [21] S. Bernstein, A. Korteweg, and K. Laws, "Attracting Early-Stage Investors: Evidence from a Randomized Field Experiment," *Journal of Finance*, vol. 72, no. 2, pp. 509–538, 2017.
- [22] K. Sharma, "Mutual fund purchases by high net worth individuals in India," *Journal of Management Accounting Research*, vol. 6, no. 2, pp. 59–71, 2006.
- [23] J. Joshi and N. N. Patel, "Investors perception towards the financial information disclosed in annual reports of listed companies of national stock exchange with special reference vadodara city," *Optics Express*, vol. 36, no. 1, p. 43, 2013.
- [24] M. Hu, X. Li, and M. Shi, "Product and pricing decisions in crowdfunding," *Marketing Science*, vol. 34, no. 3, pp. 331–345, 2015.
- [25] Y. Chen, R. Zhang, and B. Liu, "Joint Decisions on Production and Pricing with Strategic Consumers for Green Crowdfunding Products," *International Journal of Environmental Research and Public Health*, vol. 14, no. 9, 2017.

- [26] Z. Li and D.-q. Tan, “Two-stage dynamic pricing and advertising strategies for online video services,” *Discrete Dynamics in Nature and Society*, vol. 2017, Article ID 1349315, 8 pages, 2017.
- [27] S. Jørgensen, S. Taboubi, and G. Zaccour, “Cooperative advertising in a marketing channel,” *Journal of Optimization Theory and Applications*, vol. 110, no. 1, pp. 145–158, 2001.

## Research Article

# Goodwill and System Dynamics Modeling for Film Investment Decision by Interactive Efforts

Jian Feng and Bin Liu 

*School of Economics and Management, Shanghai Maritime University, Shanghai 201306, China*

Correspondence should be addressed to Bin Liu; [liubin@shmtu.edu.cn](mailto:liubin@shmtu.edu.cn)

Received 8 May 2018; Accepted 26 August 2018; Published 13 September 2018

Academic Editor: Lu Zhen

Copyright © 2018 Jian Feng and Bin Liu. 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.

Academic research pertaining to the marketing of film industry has identified advertising, film-making, and star power as the important factors influencing a movie's market performance. Prior research, however, has not investigated the joint influences of these factors. The current study has extended previous research by analyzing the investment decision of studios or investors. In order to analyze the optimal film investment decision in advertising, film-making, and stars power, this paper develops a goodwill model and system dynamic (SD) model, which allow us to disentangle the effects of advertising, film-making, and star power on film market performance. The results show that the film producer should increasingly lay emphasis on investing in advertising to absorb moviegoers' attention. Then the film producer should focus on investing in film-making when film quality has a great impact on the movie's reputation and audience's viewing decision. Furthermore, the film producer should pay more attention to the higher cost-performance stars who have more reasonable remuneration, better acting skills, and bigger box-office guarantee. Moreover, the numerical analysis reveals that rational audience contribute more than fans to a movie's box-office and bankable stars contribute more than high-profile stars to a movie's returns. Through SD simulation analysis, the film series yields higher profits than new theme movies although the cost of investment is the same.

## 1. Introduction

In the film industry, the investors are expected to make the best decisions involved in lots of funds in the shortest possible time [1]. According to Schwartz [2], "Industry estimates reveal that 60 percent or more of movies produced each year are box-office flops, those that do resonate with viewers can generate a pretty penny for investors". There are some cases where promising movies show poor box-office performance, and inconspicuous movies show unexpectedly good results [3]. "Make no mistake" shows that investing in films is a risky business [4–8]. If investment decisions fail, the costs of these mistakes for consumers are only the ticket price and an opportunity cost, but the costs for investors or producers are highly expensive [9]. Success (or mere survivability) largely depends on quickly aligning the organizational resources, like genre, director, super stars, advertising, technical effects, release time, etc. [1, 10]. The main factors affecting the financial success of a movie are of great use in making investment decisions. The film industry uses different hallmarks

of quality in an attempt to control the level of uncertainty. We will categorize the factors contributing to film investment observed in the documents into three main effects: the advertising effect, film-making effect, and the star power effect.

The advertising effect refers to the fact that studios promote their own movies to increase the number of moviegoers and the box-office revenues through advertising channels, such as the release conference, the titbits, the trailers, and show tours. Nowadays, there are multiple channels, both online and offline, for advertising delivery. Typical examples are Weibo (microblog), WeChat, variety show, and outdoor billboard [11]. Frequent new product introductions and rapid turnover lead movie studios to advertise heavily. Movie advertising is an effective measure to influence consumers' quality perception [12]. It has been confirmed by previous scholars that advertising plays a crucial role in a movie's success. [13–15]. During the movie's prerelease period, advertising would inform viewers of the movie's characteristics and signal potential studio profit ability to investors. What is

more, effective movie advertising can not only lead to success (or failure) of a single movie at the box-office, but also result in an increase (or decrease) of the releasing studio's market value. [16]. Therefore, advertising effect is the first important factor contributing to film investment.

The film-making effect refers to factors including an initial story or idea through screenwriting, casting, shooting, sound recording and reproduction, editing, and screening the finished product before a movie showing. A quantitative approach using a combination of screenwriting domain knowledge, natural-language processing techniques, and modern statistical learning methods will help studios evaluate scripts and improve a movie's gross return on investment [17, 18]. Compared with advertising and star power, the film-making gives more direct impact on a movie's word-of-mouth in the eyes of the audience. For example, the top three movies—Star Wars, E.T. the Extra-Terrestrial, and Titanic—had no stars, but it is the movie itself—not the star—that makes the success [19]. The word-of-mouth, known as the passing of information from person to person, is a method used by consumers to express their experiences and feelings about a brand, product, or service without intending to take part in marketing and promotion [20]. Since movie is a kind of good experience, potential consumers may have difficulty in evaluating film-making quality and usually search for former consumer opinions. In this paper, the word-of-mouth refers to the moviegoer's reviews and movie ratings on the level of film-making [21–23]. Industry experts agree that word-of-mouth is a critical factor underlying a movie's staying power, which remains a powerful source for movie consumers and leads to its ultimate financial success [14, 24]. Duan et al. also found that both a movie's box-office revenue and word-of-mouth valence significantly influence word-of-mouth volume. Word-of-mouth volume in turn leads to higher box-office performance [25, 26]. Additionally, we often observe that some movies produced within much smaller budgets have achieved great success by utilizing viral marketing through word-of-mouth (e.g., *The Hangover*, *Wolf Warriors II*). Consequently, film-making effect is the second vital factor involved with film investment. The experience of a producer is a considerable factor, which is as powerful as director and actor [27].

The star (e.g., award winning actor and director) power effect on film success has yielded somewhat mixed results. Some papers have reported a positive relationship between the presence of a star and box-office revenues [28, 29]. However, another scholars found no relationship, even negative effects, between stars and theatrical revenues [30–34]. Even in the study of star power, the director and actor have different degrees of importance to a movie success. Most studies have indicated that director power is not very significant and at times is much weaker than an actor power [3, 35]. However, in China, a director plays a bigger role than actor regarding the box-office [36]. Meanwhile, the high-profile stars are able to command high fees, which itself contributed to the high budgets of their films, often disproportionately so [32]. That is to say, high-profile stars are not always equal to method actors. Last but not least, the star (actor and director) power effect is the third crucial factor in regard to film investment.

There is very little literature which focus on above factors' impact on financial performance of the film from the perspective of investment decisions. Both the film's reputation and box-office success are not solely influenced by individual movie attributes such as advertising, production, director, actor, distributor, release season, and screen number, which also depended on the causal relationship among all attributes related to movies [3]. Our paper departs from the recent literature and focuses on joint influences of these interactive factors, not just one single factor. Interestingly, this paper considers the investment decision of studios or investors and develops a goodwill model and system dynamic (SD) model, which allow us to disentangle the comprehensive effects of advertising, film-making, and star power.

The rest of the paper is organized as follows. In Section 2, we will present an expanded Nerlove-Arrow model, demand function, and profit functions by using differential game. In Section 3, based on the expanded Nerlove-Arrow model, we will derive the game equilibrium solution of film investors and introduce four star selection strategies. Then, according to the four star selection scenarios, we will discuss the cost-benefit of film investment by taking spectator heterogeneity and film heterogeneity into account separately. In Section 4, we will develop a SD model of film goodwill to reflect the dynamic impact of advertising, film-making, and star power on moviegoers and box-office; then we will simulate box-office trend in the film's release cycle. Concluding remarks and managerial implications are discussed in the last section.

## 2. Model Development and Notations

A movie is considered in this paper. The film producers make investment decisions in order to attract the majority of moviegoers who focus on the advertising, film-making, and star power in a movie. We introduce the notations in this paper as shown in Table 1.

We assume that the change of film goodwill follows the Nerlove-Arrow framework [37, 38]; i.e.,

$$\dot{G}(t) = \alpha A(t) + \beta M(t) + \eta S(t) - \delta G(t) \quad (1)$$

We suppose that advertising and star power have direct effects on the number of moviegoers. The film-making effort indirectly influences moviegoers through word-of-mouth, that is, film's goodwill. The number of moviegoers satisfies the following equation:

$$N(t) = \sigma S + \gamma A + \varepsilon G \quad (2)$$

Similar to the previous literature such as Feng and Liu [38], the advertising cost, film-making cost, and star remuneration functions are quadratic with marketing efforts, respectively; namely,  $C(A) = (\mu_A/2)A^2(t)$ ,  $C(M) = (\mu_M/2)M^2(t)$  and  $C(S) = (\mu_S/2)S^2(t)$ .

Furthermore, we assume that the film's investment funds are not constrained and have a common positive discount rate  $\lambda$ . For the convenience of calculation, we suppose that the distribution cost of the film is zero. The film producers

TABLE 1: Variables and definitions used in the model.

$t$	Time $t, t \geq 0$
$G(t)$	The accumulated film's goodwill at time $t$ .
$A(t)$	The film's advertising effort at time $t$ .
$M(t)$	The film-making effort at time $t$ .
$S(t)$	The star power at time $t$ .
$N(t)$	The number of moviegoers along time $t$ .
$C(A)$	The film's advertising cost.
$C(M)$	The film-making cost.
$C(S)$	The star remuneration.
$\mu_A, \mu_M, \mu_S$	Constants
$p$	The film ticket price.
$J$	The objective profit functions of the film.
$\alpha > 0$	Positive coefficient measuring the impact of film's advertising.
$\beta > 0$	Positive coefficient measuring the impact of film-making.
$\eta > 0$	Positive coefficient measuring the impact of star power.
$\delta > 0$	The decay rate of the film goodwill.
$\sigma > 0$	Positive constant representing the effect of star power on the number of moviegoers.
$\gamma > 0$	Positive constant representing the effect of film's advertising on the number of moviegoers.
$\varepsilon > 0$	Positive constant representing the effect of film's goodwill on the number of moviegoers.
$\lambda > 0$	The discount rate.

and investors strive to maximize their profit. Then, the film's objective function is

$$J = \int_0^{\infty} e^{-\lambda t} \left[ p \times N(t) - \frac{\mu_A}{2} A^2(t) - \frac{\mu_M}{2} M^2(t) - \frac{\mu_S}{2} S^2(t) \right] dt \quad (3)$$

### 3. Equilibrium Solutions and Numerical Analysis

#### 3.1. Equilibrium Solutions

**Proposition 1.** *Film investors' equilibrium strategies are given as follows.*

(1) *The film advertising effort is*

$$A^* = \frac{\varepsilon \alpha p}{(\delta + \lambda) \mu_A} + \frac{\gamma p}{\mu_A} \quad (4)$$

(2) *The film-making effort is*

$$M^* = \frac{\varepsilon \beta p}{(\delta + \lambda) \mu_M} \quad (5)$$

(3) *The film star power effort is*

$$S^* = \frac{\varepsilon \eta p}{(\delta + \lambda) \mu_S} + \frac{\sigma p}{\mu_S} \quad (6)$$

(4) *The film investor's profit under the equilibrium condition is*

$$\begin{aligned} V^*(G) = & \frac{\varepsilon p}{\delta + \lambda} G + \frac{\varepsilon p}{\delta + \lambda} \left[ \frac{\alpha p}{\mu_A} + \frac{\alpha^2 \varepsilon p}{(\delta + \lambda) \mu_A} \right. \\ & + \frac{\beta^2 \varepsilon p}{(\delta + \lambda) \mu_M} + \frac{\eta p}{\mu_S} + \left. \frac{\eta^2 \varepsilon p}{(\delta + \lambda) \mu_S} \right] - \frac{\mu_S}{2} \left[ \frac{\sigma p}{\mu_S} \right. \\ & + \left. \frac{\eta \varepsilon p}{(\delta + \lambda) \mu_S} \right]^2 + p \left[ \frac{\gamma p}{\mu_A} + \frac{\alpha \varepsilon p}{(\delta + \lambda) \mu_A} + \frac{\sigma p}{\mu_S} \right. \\ & + \left. \frac{\varepsilon \eta p}{(\delta + \lambda) \mu_S} \right] - \frac{\mu_A}{2} \left[ \frac{\gamma p}{\mu_A} + \frac{\alpha \varepsilon p}{(\delta + \lambda) \mu_A} \right]^2 \\ & - \frac{\beta^2 \varepsilon^2 p^2}{2(\delta + \lambda)^2 \mu_M} \quad (7) \end{aligned}$$

*Proof.* See the Appendix. □

Proposition 1 illustrates the following insights. (i) The film producer increasingly lays emphasis on investing in film advertising, such as releasing conference, titbits, trailers, and show tours, so as to absorb moviegoers' attention. Due to Proposition 1, there are three positive correlations as follows. One is the relationship between investment in film advertising and advertising contribution to film goodwill. The other is the relationship between investment in film advertising and advertising contribution to the number of moviegoers. The last one is the relationship between investment in film advertising and film goodwill's contribution to the number of moviegoers. (ii) The investment in film-making has positive correlation with both film-making contribution to film

TABLE 2: Movie star selection strategies.

	Box-office guarantee	No box-office guarantee
High movie remuneration	Strategy 1 (H, G)	Strategy 2 (H, NG)
Low movie remuneration	Strategy 3 (L, G)	Strategy 4 (L, NG)

TABLE 3: Parameters assignment of rational audiences.

	$\alpha$	$\beta$	$\eta$	$\sigma$	$\gamma$	$\epsilon$
Strategy 1 (H, G)	0.1	0.5	0.4	0.5	0.1	0.4
Strategy 2 (H, NG)	0.1	0.5	0.1	0.1	0.1	0.4
Strategy 3 (L, G)	0.1	0.5	0.4	0.5	0.1	0.4
Strategy 4 (L, NG)	0.1	0.5	0.1	0.1	0.1	0.4

TABLE 4: Parameters assignment of fans.

	$\alpha$	$\beta$	$\eta$	$\sigma$	$\gamma$	$\epsilon$
Strategy 1 (H, G)	0.1	0.2	0.7	0.7	0.1	0.2
Strategy 2 (H, NG)	0.1	0.2	0.7	0.1	0.1	0.2
Strategy 3 (L, G)	0.1	0.2	0.7	0.7	0.1	0.2
Strategy 4 (L, NG)	0.1	0.2	0.1	0.1	0.1	0.2

goodwill and its impact on audience's viewing decision. The more attention that film producer pays to investing in film-making, the more importance moviegoers attach to the film's quality. On the contrary, the film producer will reduce film-making investment if the audience pursue actor or director in a movie. So there are lots of shoddy movies which focus on the star popularity but ignore film-making quality. (iii) The investment in star has positive correlation with both star contribution to film goodwill and its impact on audience's viewing decision. When the star popularity hugely influences film goodwill and potential audiences, the film investors would like to invite well-known directors or famous actors whose cost performance is higher to participate in a movie.

### 3.2. Numerical Analysis and Strategic Comparative Analysis

#### 3.2.1. Star Selection and Spectator Heterogeneity

(1) *Star Selection.* The star selection, as a critical component of the film investment, has a huge influence on the outcome of box-office, which has been largely ignored in past academic research [39]. The selection of movie stars by film investor can be divided into four strategies (in Table 2) according to the remuneration and box-office guarantee. The higher the remuneration that movie stars earn, the larger the  $\mu_S$ , and vice versa. The value of  $\alpha$  and  $\gamma$  is large if movie stars can bring high box-office, and vice versa.

(2) *Spectator Heterogeneity.* Anecdotal evidence suggests that the presence of superstars is not always a guarantee of success and, hence, a deeper study is required to analyze the potential audience of a movie [7]. Therefore, this paper considers the film audience to be heterogeneous. We only consider spectator heterogeneity: rational audiences and fans. The former's proportion is  $\omega$  and the latter's share is  $1-\omega$ .

It is also assumed that the film ticket price is constant,  $p=1$ . Let  $\mu_A = \mu_M = 1$ ,  $\lambda=\delta=0.05$ . We assume that advertising has weak influence on film goodwill and moviegoers, so the range of  $\alpha$  and  $\gamma$  is small.

(1) *Rational audiences.* They lay more attention on film-making and film goodwill. Therefore,  $\beta$  and  $\epsilon$  have higher values, see Table 3.

(2) *Fans.* They lay emphasis on famous actor or director. Therefore,  $\eta$  has a higher value, but  $\beta$  and  $\epsilon$  have lower values; see Table 4.

(3) *Investment Analysis of Spectator Heterogeneity.* On the basis of equilibrium solutions and numerical values, we draw bar charts shown in Figure 1. The film's investment strategy varies as the rational audience and fans have different viewing requirements for the film. Studios give high investment in film-making in the face of rational audiences. But the investment in star is high in order to cater to the fans. Firstly, the studios' investment in film advertising is the same whether for the rational audiences or fans. Secondly, because the film-making level has a great impact on word-of-mouth, studios increase investment in film-making to ingratiate rational audiences but economize the film-making costs in the face of fans. Thirdly, among these 4 strategies of star selection, the film investment in strategy 3 is the highest and significantly higher than the sum of investments in advertising and film-making. Meanwhile, the film investment in strategy 3 is several times higher than the cost of actors in other strategies. The investment in star of strategy 1 is the second and strategy 2 gives lowest investment in star when facing with the rational audiences. But, for fans, strategy 4 has the second investment to star, and strategy 2 gives lowest investment in star. No matter rational audiences or fans, the star selection in strategy 3 is the most cost-effective while the star cost-effectiveness of strategy 2 is minimum.

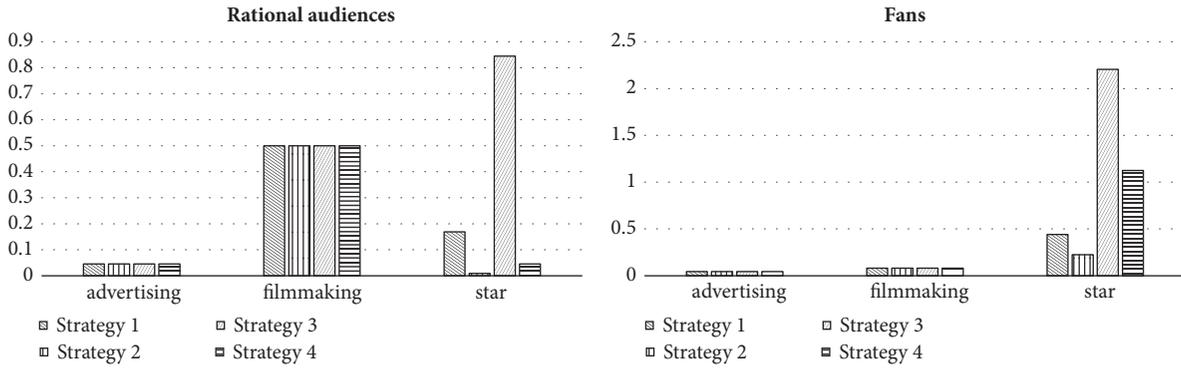


FIGURE 1: Investment analysis of spectator heterogeneity.

In fact, above analysis is precisely the case of investment made by film producers. For example, when selecting actors and actresses, film producers comprehensively take the combination of male and female protagonists into account. In Hong Kong movies, for instance, “award winning actors from Hong Kong usually work together with Chinese mainland second-tier actresses” or “Chinese mainland first-tire actresses usually work together with second and third-tier actors from Hong Kong and Taiwan”. While considering the number of fans, they would arrange some well-known second- and third-tier stars with reasonable remuneration.

**3.2.2. Film Heterogeneity and Numerical Analysis.** This paper divides the film genre into new theme film and film series (e.g., Star Trek, Pirates of the Caribbean, or Fast & Furious). Compared with new theme films, films series refer to continuous, multiset, serial movies. The numerical analysis of film investment performance is as follows.

*(1) New Theme Film. (1) Cost analysis.* The movie cost under different star selection strategies is shown in Figure 2 (the vertical axis is dimensionless). The investment cost under different strategies increases with the proportion of rational audience ( $\omega$  is the proportion of rational audience,  $\omega \in [0, 1]$ ). The cost increment of strategy 3 is the highest, and that of strategy 4 is the lowest. On the one hand, the cost of strategies 3 and 4 is higher than that of strategies 1 and 2 when fans are the main spectator group. On the other hand, when rational audience accounts for a large proportion, the total cost of strategy 1 is higher than that of strategy 4, but the cost of strategy 2 and strategy 4 tends to approach each other. Overall, the investment cost of strategy 3 is the highest and that of strategy 2 is the lowest. No matter what kind of movies, the investment cost of rational audience is higher than that of fans.

*(2) The number of moviegoers and the box-office analysis.* Figure 3 (the vertical axis is dimensionless) shows that high-paid stars contribute less to the number of moviegoers than the low-paid ones. The reason is that over-high cost limits film producers’ investment in high-profile stars. Although the high popularity of stars makes a huge contribution to the number of moviegoers, the over-high remuneration lowers stars’ cost performance. At the same time, the higher cost

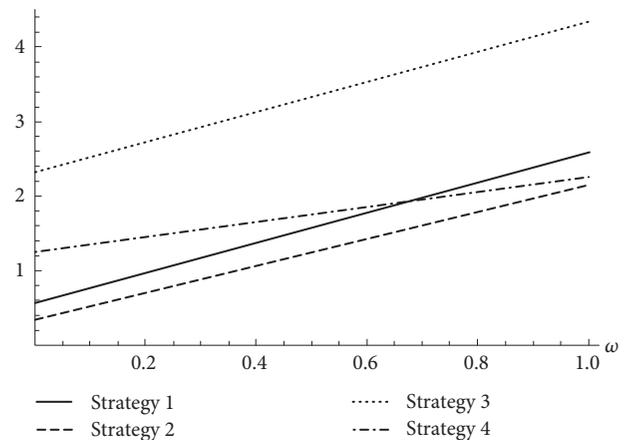


FIGURE 2: Total movie cost under different star selection strategies.

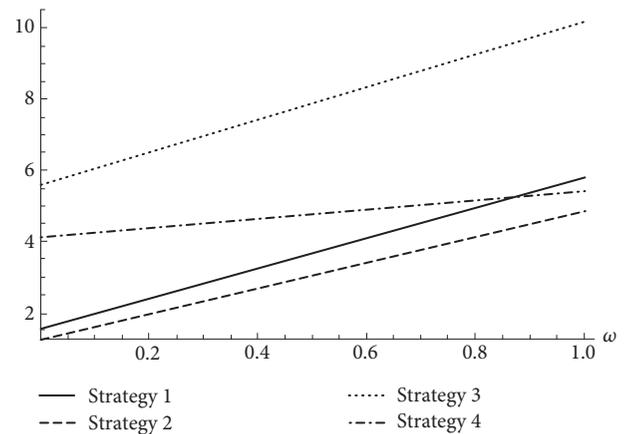


FIGURE 3: Moviegoers of new theme film under different star selection strategies.

performance makes producers increase the star investment and also increase the number of moviegoers. Bankable stars truly guarantee the box-office and the influence of such stars on the number of moviegoers is great. As the proportion of rational audience increases, this influence is more significant. It can be interpreted that rational audience always pay more attention to actors’ acting skills and film-making quality. And

box-office guarantee means getting the audience's approval in the aspects of directors, acting skills, script selection, etc. The audience loyalty to bankable stars is higher, and then the number of moviegoers is more.

In Figure 3, the biggest number of moviegoers is in strategy 3 which has the participation of the low-paid but bankable stars, so-called low-profile stars with real strength. These stars own good acting skills, loyal audience, relatively reasonable remuneration, and high cost performance. Therefore, strategy 3 has a relatively strong influence on both fans and the rational audience and earns more expected profits than other strategies. The smallest number of moviegoers is in strategy 2 which has the participation of the high-paid but with no box-office guaranteed actors. Despite such stars having high popularity, their acting skills have been questioned and criticized by moviegoers. Generally, the movie they are involved in has no guarantee at the box-office, and its online word-of-mouth is also poor.

It is generally recognized that fans are an important criterion for film producers to choose actors. But Figure 5 shows that rational audiences are the main guarantee of box-office rather than fans. Due to that film producers' investment in rational audiences is higher than that in fans; the market performance of rational audiences is better than that of fans. The number of moviegoers is directly proportional to the ratio of rational audience and is inversely proportional to the ratio of fans. So, it is easy to understand the poor box-office performance of movies starring high-paid "little fresh meat" who have a huge number of fans but no box-office guarantee. For actors who have no box-office guarantee, as the number of fans decreases, moviegoers in strategy 2 with high-paid actors will outnumber strategy 4 with low-paid actors, which is mainly because of the impact of star popularity on the movie's reputation.

Meanwhile, with the development of network, film producers pay more and more attention to moviegoers' opinions. For example, during the preparation of many movies, as a result of fans' doubts and dissatisfaction with the star lineup, the incident of changing actors occurs frequently. This also shows film producers' expectations for moviegoers and box-office, and their consideration about stars' impact on film reputation when making investment decisions. At the same time, with the changes of online word-of-mouth after the movie is released, the reputation of stars changes accordingly. Some stars may surge in popularity because of their excellent acting skills, which further enhances the reputation of the movie; some actors are questioned because of poor acting skills, resulting in damage to the movie reputation and box-office slump.

(3) *Revenue analysis.* The final revenue of a movie is affected by box-office and investment costs. This paper focuses on film investment decision-making, and the ultimate goal is to maximize the movie's revenue. We assume that the revenues of all strategies are positive and the vertical axis is dimensionless. Figure 4 shows that the higher the rational audience ratio, the higher the movie revenue. Because when the film investor expects a movie to be positive and supposes that the rational audience is influenced by the word-of-mouth, with the increase of moviegoers, the accumulation

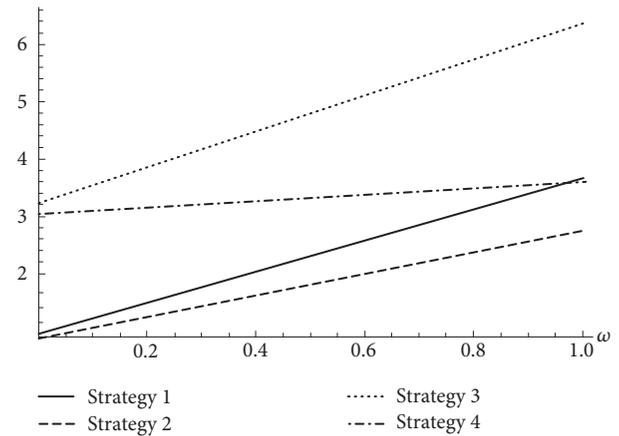


FIGURE 4: Revenue of new theme film under different strategies.

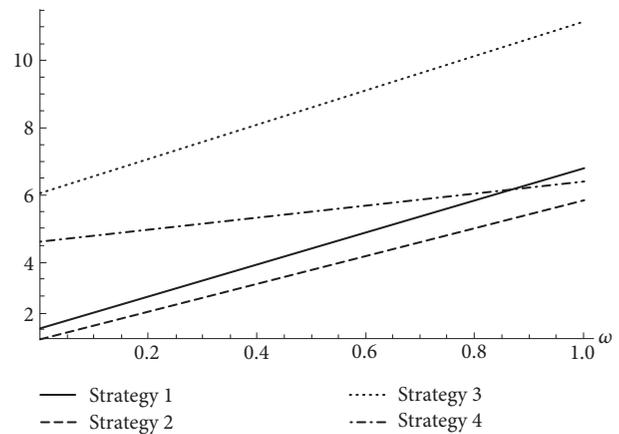


FIGURE 5: Moviegoers of film series under different star selection strategies.

of online word-of-mouth will increase the possibility of audience's going to cinema, thus increasing the film revenue. It is easy to understand that, corresponding to the number of moviegoers, revenues of movies played by bankable stars are higher than that with no box-office guaranteed stars. In addition, under cost's influence, revenues of movies played by low-paid actors are higher than that of high-paid actors.

(2) *Film Series.* Film series have a high reputation in the early stage of film release due to the accumulation of early series. High film reputation does not change the investment strategy, so the investment cost of film series is consistent with that of a new theme movie. As shown in Figures 5 and 6 (the vertical axis is dimensionless), influenced by accumulated popularity, the number of film series' moviegoers and its revenues are higher than that of a new theme movie. That is why film producers prefer giving investment in sequel movies rather than new theme movies when there is a good response to the serial movies. Therefore, there are always movies based on TV dramas, variety shows, and popular online novels. Besides, it is also easy to understand that the same story (e.g., *Journey to the West*) is reshot over and over again. All the above are

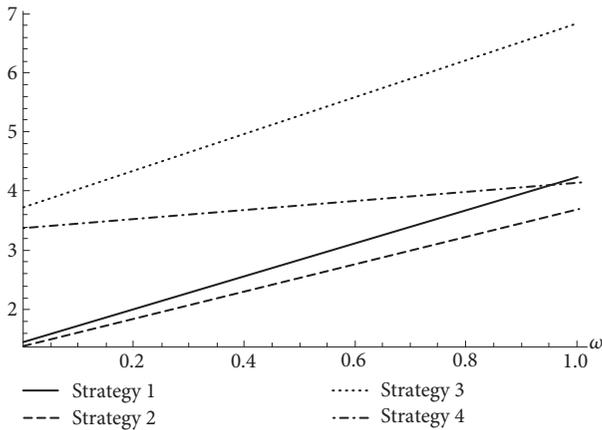


FIGURE 6: Revenue of film series under different star selection strategies.

the cases in which film producers take advantage of the high reputation accumulated before.

Similarly, the audience keeping high loyalty to some actor or director (being a fan of the star or director), on the one hand, is the recognition of the director's directing skills or actor's acting skills; and on the other hand, the director or actor's reputation that accumulated over a long period of time gives them a box-office guarantee in the mind of spectator. Therefore, the relationship between actors, directors, and movies should be mutual promotion: when an actor's popularity and acting skills or director's directing ability have high-performance in online word-of-mouth and box-office, then the movie will further enhance and consolidate the actor's or the director's popularity.

#### 4. System Dynamics Analysis of Film Goodwill and Box-Office

*4.1. The System Dynamics Model.* System dynamics is a well-established methodology to model and understand the behavior of complex systems, and it has been extensively used for modeling the dynamic behavior of complex nonlinear systems [40, 41]. Based on the above film's goodwill model, the SD model of film goodwill reflects the dynamic impact of advertising, film-making, and star power on film performance. The advertising and star power factors have direct effects on the box-office, and the film-making effort indirectly influences the box-office through film goodwill. Then the box-office in turn affects the three factors. Figure 7 shows the stock and flow diagram for the system dynamics model of film goodwill by SD software-Vensim, version 7.1.

*4.2. The System Dynamics Analysis.* According to Chen and Yin [42], movies have relatively short life cycles, and the average exposure time of a movie is approximately one month. Based on SD model in Section 4.1 and equilibrium solutions in Section 3, we will analyze four star selection strategies of new theme film by SD simulation. Then we get the box-office trend of rational audiences and fans during the release period, as shown in Figures 8 and 9 (the vertical

axis of box-office is dimensionless). In Figures 3 and 8, for rational moviegoers, strategy 3 has a higher investment in stars who are reasonably paid with box-office guarantee, and the film-making is excellent. Compared with other movies, strategy 3 can have a good performance at the early stage of release time and the potential development is great. As a result, it has considerable returns on investment. Movies (Strategies 1, 2, and 4) are ordinary in performance. Except for movies (Strategy 1) played by high-paid, bankable stars whose performance is slightly better, the box-office of other movies (Strategies 2 and 4) is close. In Figures 2 and 9, for fans, even though their idols act in the movies, they do not pay much for the low-quality film-making. Particularly in strategies 1 and 2, due to bad word-of-mouth caused by low investment in film-making, movies with high popularity stars cannot achieve good performance if just relying on fans effect. For movies with high investment in film-making, obviously, bankable stars are far more attractive to audience than that with no box-office guarantee. But, generally, the performance of movies with high investment in film-making is better than that with low investment in film-making.

In summary, in order to maximize the profit from Proposition 1, film producers choose different types of stars according to different needs of moviegoers, and thus make different investment decisions. However, returns of film investment follow the principle that high investment gets high returns. Among film investment strategies, strategy 3 has the highest investment, most moviegoers, and highest returns, while strategy 2 has the lowest investment, least moviegoers, and least returns.

Compared with new theme film, what film series reflect in the SD simulation is nothing but the fact that initial value of simulated box-office is higher. Other performances of film series in the release period are consistent with new theme film. So the SD analysis of film series is omitted in this paper. Because of the accumulated high reputation, the box-office of film series is always better than that of new theme film.

#### 5. Concluding Remarks and Managerial Implications

This paper considers the joint effects of advertising, film-making, and star power on film reputation, the number of moviegoers, and box-office. We consider the investment decision of a movie made by studios or investors and develop a goodwill model and system dynamic model which allow us to disentangle advertising, film-making, and star power effects. The results are proved as follows.

Firstly, the film producers' investment in advertising, film-making, and star power are positively correlated with their contribution to film reputation and the number of moviegoers, but negatively correlated with the cost coefficient of three factors. The film producer will increasingly lay emphasis on investing in advertising, such as the release conference, titbits, trailers, and show tours, to absorb moviegoers' attention. The film producer focuses on investing in film-making when film quality has a great impact on the movie's reputation and the audience's viewing decision. For investment in stars, the film producer pays more attention to higher

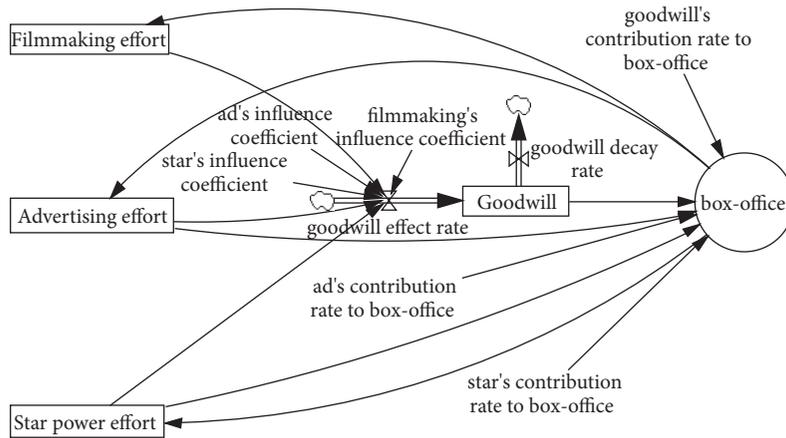


FIGURE 7: Stock and flow diagram for film goodwill.

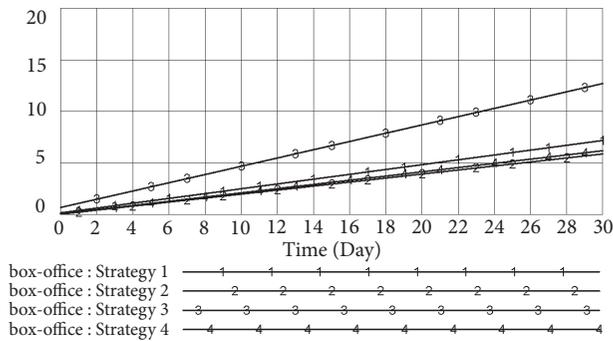


FIGURE 8: Box-office trend of rational audience against film release time.

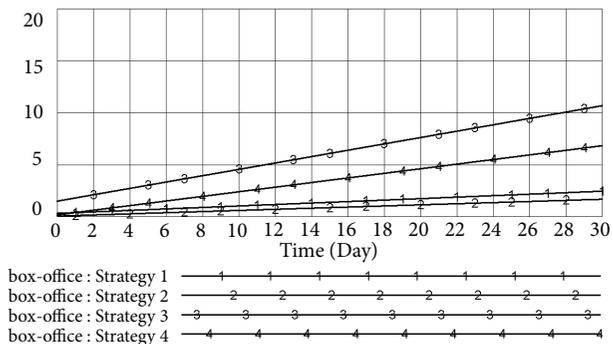


FIGURE 9: Box-office trend of fans against film release time.

cost-performance stars who have reasonable remuneration, good acting skills, and big box-office guarantee.

Secondly, the film producer makes different decisions to meet different audiences' viewing requirements. For rational audience, more attention is paid to film-making investment; but for fans, more attention is paid to investment in stars. Whether it is a new theme film or film series, the film investment cost for rational audience is higher than that for fans. Correspondingly, rational audience also contributes more box-office returns.

Thirdly, by the principle that high investment yields high returns, movies played by stars with reasonable remuneration and good acting skills have the highest investment cost. Correspondingly, these movies have the largest number of moviegoers and box-office returns. When movies played by stars with over-high remuneration and poor acting skills have the lowest investment, correspondingly, those movies get the least moviegoers and box-office returns. When rational audiences outnumber fans, despite the high remuneration, both moviegoers and box-office of movies acted by bankable stars are more than that acted by stars with no box-office guarantees.

Fourthly, the box-office will increase gradually over film release time. For the movies acted by bankable stars who are not well-known with low remuneration, they have excellent film-making quality and show the best performance. Therefore, they are most worthy of investment. For those movies acted by well-known stars but with no box-office guarantees, they show the worst performance and are least worthy of investment. Accordingly, such poor quality movies may yield disappointing box-office and reputation.

Finally, compared with new theme film, film series enable a better performance (e.g., number of moviegoers, box-office, revenues) because of high reputation. This is the fundamental reason why the film producers love serial movie.

The research findings about the film investment decision are based on the goal of maximizing film producers' profits. In reality, there are many random factors except for advertising, film-making, and star power affecting audience's viewing decision. However, this paper can provide instructions for making investment decisions in film industry from a managerial perspective. Compared with fans, the rational audience contributes more to a movie's box-office. Compared with popularity, bankable stars contribute more to a movie's returns. And compared with new theme film, film series yields higher profits.

## Appendix

We solve the optimal control question and obtain feedback equilibrium solutions using backwards induction. We

suppose that the optimal profit function of a movie is  $V(G)$ . Then the Hamilton-Jacobi-Bellman (HJB) equation must be satisfied.

$$\lambda V(G) = \max \left[ p(\sigma S + \gamma A + \varepsilon G) - \frac{\mu_A}{2} A^2 - \frac{\mu_M}{2} M^2 - \frac{\mu_S}{2} S^2 + V'(G)(\alpha A + \beta M + \eta S - \delta G) \right] \quad (\text{A.1})$$

where  $V'(G) = dV(G)/dG$ . Maximization of the right-hand side of the HJB equation yields

$$A^*(V') = \frac{\gamma p}{\mu_A} + \frac{\alpha V'}{\mu_A} \quad (\text{A.2})$$

$$M^*(V') = \frac{\beta V'}{\mu_M} \quad (\text{A.3})$$

$$S^*(V') = \frac{\sigma p}{\mu_S} + \frac{\eta V'}{\mu_S} \quad (\text{A.4})$$

It is easy to know Hessian Matrix of (A.1) is negative-definite. Equations (A.2)-(A.4) are the optimum solutions to maximize the profit in (A.1). Substituting (A.2)-(A.4) into (A.1), we obtain

$$\begin{aligned} \lambda V(G) &= p \left( \varepsilon G + \frac{\alpha V'}{\mu_A} + \frac{\gamma p}{\mu_A} + \frac{\eta V'}{\mu_S} + \frac{\sigma p}{\mu_S} \right) \\ &+ V' \left( \frac{\alpha^2 V'}{\mu_A} + \frac{\alpha p}{\mu_A} + \frac{\beta^2 V'}{\mu_M} + \frac{\eta^2 V'}{\mu_S} + \frac{\eta p}{\mu_S} - \delta G \right) \\ &- \frac{\mu_A}{2} \left( \frac{\alpha V'}{\mu_A} + \frac{\gamma p}{\mu_A} \right)^2 - \frac{\beta^2 V'}{2\mu_M} \\ &- \frac{\mu_S}{2} \left( \frac{\eta V'}{\mu_S} + \frac{\sigma p}{\mu_S} \right)^2 \end{aligned} \quad (\text{A.5})$$

We shall show that linear optimal value functions satisfy (A.5). Hence, we define

$$V(G) = e_1 G + e_2 \quad (\text{A.6})$$

where  $e_1$  and  $e_2$  are constants.

By using the method of undetermined coefficients, we obtain

$$e_1 = \frac{\varepsilon p}{\delta + \lambda} \quad (\text{A.7})$$

$$\begin{aligned} e_2 &= \frac{\varepsilon p}{\delta + \lambda} \left[ \frac{\alpha p}{\mu_A} + \frac{\alpha^2 \varepsilon p}{(\delta + \lambda) \mu_A} + \frac{\beta^2 \varepsilon p}{(\delta + \lambda) \mu_M} + \frac{\eta p}{\mu_S} \right. \\ &+ \left. \frac{\eta^2 \varepsilon p}{(\delta + \lambda) \mu_S} \right] + p \left[ \frac{\gamma p}{\mu_A} + \frac{\alpha \varepsilon p}{(\delta + \lambda) \mu_A} + \frac{\sigma p}{\mu_S} \right. \\ &+ \left. \frac{\varepsilon \eta p}{(\delta + \lambda) \mu_S} \right] - \frac{\mu_A}{2} \left[ \frac{\gamma p}{\mu_A} + \frac{\alpha \varepsilon p}{(\delta + \lambda) \mu_A} \right]^2 \\ &- \frac{\beta^2 \varepsilon^2 p^2}{2(\delta + \lambda)^2 \mu_M} - \frac{\mu_S}{2} \left[ \frac{\sigma p}{\mu_S} + \frac{\eta \varepsilon p}{(\delta + \lambda) \mu_S} \right]^2 \end{aligned} \quad (\text{A.8})$$

Substituting (A.7)-(A.8) into (A.6), we get

$$V' = \frac{\varepsilon p}{\delta + \lambda} \quad (\text{A.9})$$

Then substituting (A.9) into (A.2)-(A.4), we also obtain (4)-(6). Finally, substituting (4)-(6) into (A.1), and get (7).

This concludes the proof.

## Data Availability

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

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors gratefully acknowledge support from National Natural Science Foundation of China (Project No. 71571117) and Innovation Method Fund of China (Project No. 2016IM010400).

## References

- [1] D. Delen, R. Sharda, and P. Kumar, "Movie forecast Guru: A Web-based DSS for Hollywood managers," *Decision Support Systems*, vol. 43, no. 4, pp. 1151-1170, 2007.
- [2] S. K. Schwartz, Investing in the big screen can be a profitable story, <https://www.cnbc.com/id/39342145>, 2010.
- [3] K. J. Lee and W. Chang, "Bayesian belief network for box-office performance: A case study on Korean movies," *Expert Systems with Applications*, vol. 36, no. 1, pp. 280-291, 2009.
- [4] A. De Vany and W. D. Walls, "Bose-einstein dynamics and adaptive contracting in the motion picture industry," *Economic Journal*, vol. 106, no. 439, pp. 1493-1514, 1996.
- [5] A. De Vany and W. D. Walls, "Does Hollywood Make Too Many R-Rated Movies? Risk, Stochastic Dominance, and the Illusion of Expectation," *Journal of Business*, vol. 75, no. 3, pp. 425-451, 2002.
- [6] A. De Vany and W. D. Walls, "Motion picture profit, the stable Paretian hypothesis, and the curse of the superstar," *Journal of Economic Dynamics & Control*, vol. 28, no. 6, pp. 1035-1057, 2004.
- [7] E. Montañés, A. Suárez-Vázquez, and J. R. Quevedo, "Ordinal classification/regression for analyzing the influence of superstars on spectators in cinema marketing," *Expert Systems with Applications*, vol. 41, no. 18, pp. 8101-8111, 2014.
- [8] W. D. Walls, "Chapter 8—bestsellers and blockbusters: movies, music, and books," in *Handbook of the Economics of Art & Culture*, vol. 2, pp. 185-213, 2014.
- [9] J. Eliashberg, A. Elberse, and M. A. A. M. Leenders, "The motion picture industry: Critical issues in practice, current research, and new research directions," *Marketing Science*, vol. 25, no. 6, pp. 638-661, 2006.
- [10] L. Zhang, J. Luo, and S. Yang, "Forecasting box office revenue of movies with BP neural network," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6580-6587, 2009.

- [11] J. Xiao, X. Li, S. Chen, X. Zhao, and M. Xu, "An inside look into the complexity of box-office revenue prediction in china," *International Journal of Distributed Sensor Networks*, vol. 13, no. 1, pp. 1–14, 2017.
- [12] T. Hennig-Thurau, M. B. Houston, and G. Walsh, "Determinants of motion picture box office and profitability: an interrelationship approach," *Review of Managerial Science*, vol. 1, no. 1, pp. 65–92, 2007.
- [13] F. S. Zufryden, "Linking advertising to box office performance of new film releases - A marketing planning model," *Journal of Advertising Research*, vol. 36, no. 4, pp. 29–41, 1996.
- [14] A. Elberse and J. Eliashberg, "Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures," *Marketing Science*, vol. 22, no. 3, pp. 329–354, 2003.
- [15] A. Elberse and B. Anand, "The effectiveness of pre-release advertising for motion pictures: an empirical investigation using a simulated market," *Information Economics & Policy*, vol. 19, no. 3–4, pp. 319–343, 2007.
- [16] A. M. Joshi and D. M. Hanssens, "Movie advertising and the stock market valuation of studios: a case of great expectations?" *Science*, vol. 28, no. 2, pp. 239–250, 2009.
- [17] W. D. Walls, "Modeling Movie Success When 'Nobody Knows Anything': Conditional Stable-Distribution Analysis Of Film Returns," *Journal of Cultural Economics*, vol. 29, no. 3, pp. 177–190, 2005.
- [18] J. Eliashberg, S. K. Hui, and Z. J. Zhang, "From story line to box office: A new approach for green-lighting movie scripts," *Management Science*, vol. 53, no. 6, pp. 881–893, 2007.
- [19] A. Elberse, "The power of stars: Do star actors drive the success of movies?" *Journal of Marketing*, vol. 71, no. 4, pp. 102–120, 2007.
- [20] C. Hung, "Word of mouth quality classification based on contextual sentiment lexicons," *Information Processing & Management*, vol. 53, no. 4, pp. 751–763, 2017.
- [21] S. Basuroy, S. Chatterjee, and S. Abraham Ravid, "How Critical are Critical Reviews? The Box Office Effects of Film Critics, Star Power, and Budgets," *Journal of Marketing*, vol. 67, no. 4, pp. 103–117, 2003.
- [22] C. C. Moul, "Measuring word of mouth's impact on theatrical movie admissions," *Journal of Economics & Management Strategy*, vol. 16, no. 4, pp. 859–892, 2007.
- [23] P. Winoto and T. Y. Tang, "The role of user mood in movie recommendations," *Expert Systems with Applications*, vol. 37, no. 8, pp. 6086–6092, 2010.
- [24] Y. Liu, "Word of mouth for movies: its dynamics and impact on box office revenue," *Journal of Marketing*, vol. 70, no. 3, pp. 74–89, 2006.
- [25] W. D. Walls, "Increasing returns to information: Evidence from the Hong Kong movie market," *Applied Economics Letters*, vol. 4, no. 5, pp. 287–290, 1997.
- [26] W. Duan, B. Gu, and A. B. Whinston, "The dynamics of online word-of-mouth and product sales - an empirical investigation of the movie industry," *Journal of Retailing*, vol. 84, no. 2, pp. 233–242, 2008.
- [27] M. Bagella and L. Becchetti, "The determinants of motion picture box office performance: Evidence from movies produced in Italy," *Journal of Cultural Economics*, vol. 23, no. 4, pp. 237–256, 1999.
- [28] B. R. Litman and L. S. Kohl, "Predicting financial success of motion pictures: the '80s experience," *Journal of Media Economics*, vol. 2, no. 2, pp. 35–50, 1989.
- [29] S. Albert, "Movie stars and the distribution of financially successful films in the motion picture industry," *Journal of Cultural Economics*, vol. 22, no. 4, pp. 249–270, 1998.
- [30] B. R. Litman, "Predicting success of theatrical movies: an empirical study," *The Journal of Popular Culture*, vol. 16, no. 4, pp. 159–175, 1983.
- [31] S. P. Smith and V. K. Smith, "Successful movies: a preliminary empirical analysis," *Applied Economics*, vol. 18, no. 5, pp. 501–507, 1986.
- [32] J. Sedgwick and M. Pokorny, "Movie stars and the distribution of financially successful films in the motion picture industry," *Journal of Cultural Economics*, vol. 23, no. 4, pp. 319–323, 1999.
- [33] A. De Vany and W. D. Walls, "Uncertainty in the movie industry: does star power reduce the terror of the box office?" *Journal of Cultural Economics*, vol. 23, no. 4, pp. 285–318, 1999.
- [34] W. D. Walls, "Screen wars, star wars, and sequels," *Empirical Economics*, vol. 37, no. 2, pp. 447–461, 2009.
- [35] R. A. Nelson and R. Glotfelty, "Movie stars and box office revenues: an empirical analysis," *Journal of Cultural Economics*, vol. 36, no. 2, pp. 141–166, 2012.
- [36] J. Zhou, Y. Yuan, and P. Tu, "Analysis of the collaboration network of Chinese film directors and actors," *Complex Systems and Complexity Science*, vol. 13, no. 3, pp. 69–75, 2016.
- [37] M. Nerlove and K. J. Arrow, "Optimal advertising policy under dynamic conditions," *Economica*, vol. 29, no. 114, pp. 129–142, 1962.
- [38] J. Feng and B. Liu, "Dynamic impact of online word-of-mouth and advertising on supply chain performance," *International Journal of Environmental Research and Public Health*, vol. 15, no. 1, p. 69, 2018.
- [39] A. Liu, T. Mazumdar, and B. Li, "Counterfactual decomposition of movie star effects with star selection," *Science*, vol. 61, no. 7, pp. 1704–1721, 2015.
- [40] J. Forrester, "Industrial dynamics: A major breakthrough for decision makers," *Harvard Business Review*, vol. 36, no. 4, pp. 37–66, 1958.
- [41] R. R. P. Langroodi and M. Amiri, "A system dynamics modeling approach for a multi-level, multi-product, multi-region supply chain under demand uncertainty," *Expert Systems with Applications*, vol. 51, pp. 231–244, 2016.
- [42] K. Chen and J. Yin, "Information competition in product launch: evidence from the movie industry," *Electronic Commerce Research Applications*, vol. 26, pp. 81–88, 2017.

## Research Article

# An Incentive Dynamic Programming Method for the Optimization of Scholarship Assignment

Di Huang,<sup>1</sup> Yu Gu ,<sup>1</sup> Hans Wang,<sup>2</sup> Zhiyuan Liu ,<sup>1</sup> and Jun Chen <sup>1</sup>

<sup>1</sup>Jiangsu Key Laboratory of Urban ITS, Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, School of Transportation, Southeast University, China

<sup>2</sup>Department of Logistics & Maritime Studies, The Hong Kong Polytechnic University, Kowloon, Hong Kong

Correspondence should be addressed to Zhiyuan Liu; zhiyuanl@seu.edu.cn

Received 8 June 2018; Accepted 31 July 2018; Published 12 August 2018

Academic Editor: Xinchang Wang

Copyright © 2018 Di Huang 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.

Scholarship assignment is an operations management problem confronting university administrators, which is traditionally solved based on administrators' personal experiences. This paper proposes an incentive method inspired by dynamic programming to replace the traditional decision-making process in the scholarship assignment. The objective is to find the optimal scholarship assignment scheme with the highest equity while accounting for both the practical constraints and the equity requirement. Moreover, with the proposed method, the scholarship assignment avoids time- and energy-consuming application processes conducted by students. A solution algorithm is used to find feasible assignment schemes by iteratively solving a series of knapsack subproblems based on dynamic programming and adjusting the monetary value of a unit score. The optimal assignment scheme can then be screened out by applying the Gini coefficient for quantifying the equity of each feasible scheme. A numerical case is investigated to illustrate the applicability of the proposed method and solution algorithm. The results indicate that the proposed method is an efficient tool to assign scholarships to students with consideration of the equity.

## 1. Introduction

Scholarships act as an important incentive and plays a crucial role in encouraging undergraduate and graduate students in universities to work hard, excel, and aim for higher education in their academic careers. It is also considered as an important component from the teaching perspective. Many researches have proved that scholarships have positive effects on students' performances in college and encourage students to seek for further education [1–3]. Different scholarships are funded by different sponsors, such as entrepreneurs, institutions, individuals, etc., who are in charge of determining the number of awardees and amount of money awarded to each recipient. Departments of universities are responsible for assigning scholarships to candidates by ranking them based on their performances in departments. However, the university administrators have no right to further divide the existing scholarships into new scholarships with smaller amounts of monetary awards. The performances of students

are evaluated by the score they gained each year. The final score of a student is the summation of the following two parts: (1) basic score, the grade of the student's coursework, and (2) research score, the equivalent score converted by the student's academic achievements (e.g., winning competitions, publishing papers and patents).

In practice, there are many factors to be considered in scholarship assignment. First, a scholarship from the same provider should not be assigned to the same student more than once, as sponsors require to cover as many deserving students as they can. Second, a scholarship cannot be further divided, as the amount of each scholarship is determined by its sponsor. Third, all scholarships must be awarded to students. Fourth, scholarships should be assigned to students based on their performances (scores gained from coursework and research), stating that the student with a higher score should receive an amount of scholarship no less than that received by students with lower scores. Fifth, the amount of scholarship assigned to each student should be strictly

in line with his/her score, in order to guarantee equity of scholarship assignment. For instance, when two students get similar scores but the difference between the amounts of the scholarships awarded to these students is too large to ignore, such discrimination may lead to discontent as the students may feel unfairly treated.

Our personal experiences show that almost all scholarships are not directly assigned to students but awarded to the best applicants who apply for the scholarships. Students can separate their research scores into smaller parts and use different parts to apply for different scholarships. Then the applicants for each scholarship will be ranked by the summations of their basic scores and chosen research scores. The applicants with the highest rankings have higher chances of winning the scholarships. As we all know, the assignment mechanism is not only dependent on applicants' performances, but also strongly influenced by applicants' strategies. One can receive a higher amount of scholarships than students with higher scores by adopting a better strategy for dividing his/her research scores and applying for appropriate scholarships with fewer competitors. Moreover, such mechanism is both time- and energy-consuming for students as they tend to focus on collecting information of their competitors and developing appropriate application strategies at the expense of their studies and research, thus, exerts an overall negative impact on their academic performance. In this regard, it is imperative and of vital significance for university administrators to develop a systematic tool to assign scholarships in an efficient and equitable way.

*1.1. Literature Review.* It is reported in a diversity of researches that, in spite of providing financial assistance, scholarships will help students in various aspects including motivating studying, improving confidence, and even changing their future plans [4–6]. The assignment of scholarships thus becomes an important issue that might highly affect living and learning of graduate students. To the best of our knowledge, few researches have investigated the scholarship assignment problem, except the research conducted in [7]. Amorós et al. admitted that the best student should receive the most prestigious scholarship and proposed an assignment mechanism in Nash Equilibrium to achieve a socially optimal outcome. However, they fail to consider the equity norm in their assignment mechanism.

The equity norm has gained increasing attention in higher education [8–11] and is a pressing concern in the scholarship assignment as much as in other policy making [12–14] and resource or reward allocation problems [15–19]. Equity, unlike equality which implies that all members should be assigned to equal amount of resource, requires that the allocated resource should be proportionate to the magnitude of each member's contribution [20, 21]. In a group, the members who perform better often require equity when assigning resources. It is found that people who outperform their peers intend to ask for equity in resource allocation [22]. Such mechanism implies that, in the scholarship assignment, students with higher scores might possess greater demands for equity. Therefore, it is important to consider the equity norm in scholarship assignment.

As few efforts have focused on the equity concern of scholarship assignment problem, this research aims to develop a scholarship assignment model to optimize the quality of scholarship assignment scheme by considering the equity norm. The model can guarantee a basic requirement for equity; i.e., students receive a higher amount of scholarships only when they have higher scores. Further, this proposed model can quantify the equity of all feasible assignment schemes satisfying the above requirement and help select the scheme with the highest equity.

*1.2. Objectives and Contributions.* The objective of this research is to develop an efficient tool to generate the assignment scheme of scholarships with the highest equity. The contribution of the paper is threefold. First, we develop a new approach that allows the university administrators to assign scholarships to students according to the ranking of coursework and research performances. This is because our method prioritizes students with higher scores in scholarship assignment. As a result, the equity of scholarship assignment scheme can be guaranteed by meeting the requirement that a student with a higher score must receive an amount of scholarships no less than students with lower scores. Such equity requirement, however, is hard to satisfy when students take the leading role in scholarship assignment process, as if scholarships should be applied by students. On the contrary, the proposed method is led by university administrators, which contributes to helping motivate and encourage the students to work hard in their study and research, as better performances are strictly associated with higher monetary rewards.

Second, as this assignment is done by university administrators instead of students themselves, who are naturally busy with studying, researching, and other commitments, our method can save the valuable time for students.

Third, the method can assess the equity of different feasible assignment schemes. Various schemes can meet the requirement that better students receive more scholarships, while the qualities of schemes differ from each other. One may find that another student who performs little better than s/he can receive much more scholarships, which is another form of inequity. This concern can be addressed by introducing the Gini coefficient to measure the inequity of each feasible scholarship assignment scheme.

The remainder of this paper is organized as follows. The description of considerations in the scholarship assignment problem is first introduced. The method and solution algorithm to solve the proposed scholarship assignment problem are presented in detail, followed by a small example to illustrate the application of the proposed method. Then the results and findings of the numerical example are reported. Finally, the conclusion is drawn.

## 2. The Scholarship Assignment Problem

This study investigates the assignment mechanism of scholarships at a department. In a department, different kinds of scholarships are offered to graduate students who have done excellent jobs in academic research. Different types

of scholarships may differ from each other in two aspects: (1) the number of awardees and (2) the amount of money an awardee can receive. Students are free to apply for different scholarships. The results of scholarship appraisal are dependent on the students' scores they have obtained in the previous year. As mentioned above, the total score of a student is composed of the basic score obtained from coursework and additional research scores based on his academic achievements such as published journal papers. In order to simplify the evaluation of students' works, the basic score is measured in the hundred-mark system, while different kinds of academic achievements are measured in the same weight. Additionally, a standardization process is applied to convert a specific work into standard scores. For instance, the score of a published paper can be calculated in terms of two components: journal's impact factor (IF) ranking and author's order. A student who has published a paper in the journal which occupies in Q1 [23] and is also the first author of that journal will achieve additional 10 points. The score decreases linearly with the slip of IF or author's order.

In this study, we assume that the department is in charge of the assignment of scholarships and the scholarship is allocated according to the ranking of students' total scores. To avoid the unfairness in assigning scholarships, we include the constraint that the student with a higher score should be offered at least no less amount of scholarship than students with lower scores. Meanwhile, other practical constraints are considered in the study. For instance, a single scholarship cannot be shared by different students; a student can only be awarded by a certain type of scholarship no more than once; a student can either fail to win any scholarship or win multiple scholarships; all the scholarships should be assigned to students. The optimal assignment should be the scheme that best guarantee the amount of scholarship received by a student is proportionate to the magnitude of his/her total score.

The notations used in the method are listed below:

#### Sets

$J$ : Set of scholarships

$I$ : Set of students

$J_i$ : Set of remaining scholarships that can be assigned to student  $i$

#### Indices

$j$ : An index  $j \in J$  that refers to a particular scholarship

$i, k$ : Indices  $i, k \in I$  that refer to particular students

#### Parameters and Variables

$m_j$ : The amount of scholarship  $j$  (\$)

$c_j$ : The number of scholarship  $j$

$e_j$ : Equivalent score of scholarship  $j$

$s_i$ : The score of student  $i$

$v$ : Monetary value of a unit score

$G$ : Gini coefficient

$n$ : Number of students

$W_i$ : Welfare of student  $i$

$\bar{W}$ : Average welfare of individual students

$W_{\min}$ : Minimum welfare of students

$\kappa$ : A constant used to adjust individual welfare

$W_i^s$ : Adjusted welfare of student  $i$

$\bar{W}^s$ : Adjusted average welfare of individual students

$\delta_{ij}$ : A binary decision variable which equals 1 if scholarship  $j$  is assigned to student  $i$ , and 0 otherwise

### 3. The Scholarship Assignment Mechanism

The scholarship assignment problem discussed above can be deemed as a combination of two subproblems. The first subproblem is to search for all feasible assignment schemes which meet the practical constraints and the basic equity requirement simultaneously. The second subproblem is to quantify the equity of each feasible scheme and select the optimal scheme with the highest equity.

To address the first subproblem, a method composed of knapsack subproblems is developed to assign scholarships to students in sequence. In the knapsack subproblem, the total score gained by each student is deemed as the knapsack and the equivalent score of each scholarship is regarded as the item to be put in the knapsack. Given a certain value of a unit score, the amounts of all scholarships can be converted to equivalent scores. Assigning scholarships to each student by solving knapsack problems can guarantee that the amount of scholarships (equivalent score) awarding to a student is the maximum one that can be allocated to him/her among remaining scholarships. Hence, the students who have priorities in such scholarship assignment process can receive more scholarships. Through ranking students according to their scores and assigning scholarships in sequence, the proposed method thus certainly satisfies the basic equity requirement that students with higher scores should receive higher amounts of scholarships.

To address the second subproblem, the Gini coefficient is introduced to measure the inequity of each feasible solution. The Gini coefficient can quantify the equity of distribution of welfare (difference between the actual and ideal amount of scholarship gained by each individual student). The Gini coefficient equals 0 when students receive the same welfare; i.e., there exists no difference between the actual and ideal amounts of scholarship gained by each student. A higher Gini coefficient is associated with a higher inequity of the investigated assignment scheme. The proposed method modifies the original calculation of Gini coefficient used in the field of economics, making it applicable to the scholarship assignment problem. The feasible assignment scheme with the lowest value of Gini coefficient is selected as the optimal assignment scheme.

**3.1. Searching for Feasible Solutions.** To satisfy the requirement that students with higher scores can obtain an amount

of scholarships no less than students with lower scores, the scholarship assignment problem is converted to a series of one-dimensional 0-1 knapsack problems [24, 25].

Firstly, the initial monetary value  $v$  of a unit score can be calculated by

$$v = \frac{\sum_{j \in J} m_j \cdot c_j}{\sum_{i \in I} s_i}. \quad (1)$$

According to the monetary value of a unit score, the amount of scholarships,  $m_j$ , can then be converted to equivalent scores,  $e_j$ . The process can be expressed as

$$e_j = \frac{m_j}{v}, \quad \forall j \in J. \quad (2)$$

The sequence of assigning scholarships to students is in accordance with the ranking of students' scores. In each iteration, the scholarships are allocated to the student with the highest score among unsigned students, and the amount of scholarships to be allocated to the student is derived through solving a knapsack subproblem. Let the score of a student  $i$ ,  $s_i$ , be the capacity of the knapsack, and the equivalent score of scholarship  $j$ ,  $e_j$ , be the items to be put into the knapsack. Thus the scholarship assignment scheme of the chosen student  $i$  can be derived from

$$\max \sum_{j \in J_i} e_j \cdot \delta_{ij} \quad (3)$$

$$\text{s.t.} \quad \sum_{j \in J_i} e_j \cdot \delta_{ij} \leq s_i. \quad (4)$$

By iteratively solving knapsack subproblems for each chosen student and assigning scholarships in accordance to the solutions, the students with higher scores, who possess priority in the scholarship assignment, can certainly obtain greater amounts of scholarships. Hence, given a certain monetary value of the score  $v$ , we can derive a corresponding feasible solution set (scholarship assignment scheme) using the above method. By iteratively adjusting the monetary value of a unit score, we can enumerate all the feasible solution sets.

However, the qualities of feasible solution sets derived by the above method cannot be judged. As this paper aims to assign scholarships to students considering the equity norm, the finally selected assignment scheme must be the one with the highest equity. The following section introduces the method to quantify equity of each feasible solution.

### 3.2. Measuring Equity of Scholarship Assignment Scheme.

Note that perfect equity means that every student can receive a proportion of scholarship same as the proportion of his/her score in the total score. To quantify the equity of scholarship assignment schemes which aim to assigning different scholarships to a group of students with different scores, a suggestion is associated with modifying an existing metric of income inequity, i.e., Gini coefficient. In the transportation field, the Gini coefficient is widely utilized to analyze inequity in accessibility [26–28] and effect of transport policies [29, 30]. Herein, the Gini coefficient is adopted to measure the social

welfare (for instance, the accessibility to certain opportunities or the welfare of toll pricing scheme) achieved by transport users. In scholarship assignment, the Gini coefficient  $G$  can express an inequality metric that evaluates the departure from perfect equity, which can be expressed as

$$G = \frac{1}{2n^2\bar{W}} \sum_{i \neq k \in I} |W_i - W_k| \quad (5)$$

where  $W_i$  denotes the benefit that student  $i$  receives from the actual assignment scheme,  $\bar{W}$  denotes the average individual welfare, and  $n$  is the number of students. Considering the practical constraints, the ideal assignment result with perfect equity is hard to reach. An individual student might gain either positive benefit (receive more scholarship than s/he deserves) or negative benefit (receive less scholarship than s/he deserves) in certain scholarship assignment scheme. To measure the benefit each student receives from the assignment scheme, we define the welfare of student  $i$ ,  $W_i$ , as the difference between the student's actual scholarship and ideal scholarship, which can be derived by

$$W_i = s_i^a - s_i^p \quad (6)$$

where  $s_i^a$  and  $s_i^p$  denote the actual and ideal results of scholarship assignment with regards to student  $i$  respectively. The average individual welfare  $\bar{W}$  can be expressed as

$$\bar{W} = \frac{\sum_{i \in I} W_i}{n}. \quad (7)$$

It is evident that the average individual welfare  $\bar{W}$  equals 0, and the individual welfare can be negative. These properties prohibit the usage of original Gini coefficient in that the value of average and individual income used in the original Gini coefficient evaluation can only be positive. Therefore, a rescaling process is necessary to implement (5). The re-scaled individual welfare  $W_i^s$  can be derived by

$$W_i^s = W_i + |W_{\min}| + \kappa \quad (8)$$

where  $W_{\min}$  denotes the minimum individual welfare and  $\kappa$  is a given constant.

By using  $W_i^s$  to replace  $W_i$  in (5) and (7), the Gini coefficient  $G$  of each solution can be obtained:

$$G = \frac{1}{2n^2\bar{W}^s} \sum_{i \neq k \in I} |W_i^s - W_k^s|. \quad (9)$$

## 4. Solution Algorithm

As discussed in Section 3, the task of searching for feasible assignment scheme can be converted to a series of one-dimensional 0/1 knapsack problem, which can be exactly solved by dynamic programming algorithms [31–36]. In this section, a solution algorithm for the knapsack subproblem based on dynamic programming is used to find the optimal scholarship assignment for each student in sequence. On the basis of it, a solution algorithm for finding feasible assignment schemes and selecting the optimal scheme is proposed.

**4.1. Algorithm 1: Algorithm for One-Dimensional 0-1 Knapsack Subproblem.** Assuming there are totally  $p$  types of scholarships, let stage  $h$  denote the first  $h$  types of scholarships that are considered in scholarship assignment;  $s$  denote the equivalent score that is currently assigned to the selected student; and  $f_h(s)$  denote the maximum equivalent score the selected student can obtain at stage  $h$ .

*Step 1* (initialization). Set  $f_0(s) = 0, \forall s \leq s_i, h = 1, s = 0, j = 1$ .

*Step 2.* Update  $f_h(s)$ .

*Step 2.1.*  $f_h(s) = \max\{f_{h-1}(s), f_{h-1}(s - m_i) + m_j\}$ , where  $m_i$  is the equivalent score of the scholarship previously assigned to the selected student.

*Step 2.2.*  $j = j + 1$ . If  $j > h$ , go to Step 2.3; otherwise, go to Step 2.1

*Step 2.3.*  $s = s + 1$ . If  $s > s_i$ , go to Step 3; otherwise, go to Step 2.1.

*Step 3.*  $h = h + 1$ . If  $h \leq p$ , go to Step 2; otherwise, return  $f_{h-1}(s_i)$  and stop the algorithm.

**4.2. Algorithm 2: Derive Feasible Schemes and Select Optimal Scheme**

*Step 1.* Initialize the monetary value of a unit score  $v$  using (1); set the value of step size  $d$  and upper bound of monetary score value  $v_{\max}$ .

*Step 2.* Obtain the equivalent score of each scholarship  $m_j$  using (2).

*Step 3.* Derive feasible scholarship assignment.

*Step 3.1.* Rank students according to their scores and put students into set  $R\{1, 2, \dots, i, \dots\}$ .

*Step 3.2.* Assign scholarships to the first student  $i$  in set  $R$  using Algorithm 1.

*Step 3.3.* Delete student  $i$  from set  $R$ . If set  $R$  is not empty,  $i = i + 1$ , go to Step 3.2; otherwise, go to Step 4.

*Step 4.* Store assignment result of Step 3. Increase monetary score value  $v = v + d$ ; if  $v < v_{\max}$ , go to Step 2; otherwise, go to Step 5.

*Step 5.* Using (9) to calculate the Gini coefficient of each assignment scheme stored in Step 4, return the optimal assignment scheme with lowest Gini coefficient value.

## 5. An Illustrative Example

To illustrate the proposed method, we propose a small case with 3 types of scholarships and 5 candidates as an example. The unit of scholarship is US dollar (\$) in Tables 1, 3, and 4.

TABLE 1: Details of scholarships.

ID	Number of awardees	Amount of scholarship
1	2	8000
2	3	5000
3	1	10000

TABLE 2: Data of candidates.

ID	Score
1	25
2	20
3	30
4	35
5	15

TABLE 3: Ideal assignment results.

ID	Amount of scholarship
1	8200
2	6560
3	9840
4	11480
5	4920

Tables 1 and 2 show the details of scholarships and scholarship candidates.

It can be seen that the number of awardees of each scholarship is separately 2, 3, and 1; the monetary awards provided by each scholarship are, respectively, \$8000, 5000, and 10000. The total scores of each student are 25, 20, 30, 35, and 15. Following the equity norm that the ideal amount of scholarship received by each student is proportionate to his/her score, the ideal assignment result neglecting the practical constraints is shown in Table 3.

Using the method discussed in Section 3.1, four feasible schemes can be obtained by iteratively solving the knapsack subproblem for each student and then adjusting the monetary value of the score. Note that the derived schemes only satisfy the basic equity requirement that a student with higher score must receive an amount of scholarship no less than that received by students with comparatively lower scores. The details regarding the 4 schemes are shown in Table 4.

Through the method described in Section 3.2, we can then compute the Gini coefficient of each scheme, which is shown in Table 5.

As per the Gini coefficients of each scheme, it can be found that, among 4 feasible schemes, scheme 1 is the one with the highest equity, as the Gini coefficient of scheme 1 is the lowest. Therefore, scheme 1 is chosen as the final scholarship assignment scheme in this numerical example.

From this example, it can be concluded that the proposed method can enumerate several feasible scholarship assignment schemes and then choose the optimal one with the minimum value of Gini coefficient. This demonstrates the practical relevance of the proposed method.

TABLE 4: Derived feasible schemes of scholarship assignment.

Scheme 1				Scheme 2			
ID	Score	Amount of scholarship	Type of scholarship	ID	Score	Amount of scholarship	Type of scholarship
1	25	8000	1	1	25	8000	1
2	20	5000	2	2	20	5000	2
3	30	10000	3	3	30	13000	1,2
4	35	13000	1,2	4	35	15000	1,3
5	15	5000	2	5	15	0	\

Scheme 3				Scheme 4			
ID	Score	Amount of scholarship	Type of scholarship	ID	Score	Amount of scholarship	Type of scholarship
1	25	5000	2	1	25	5000	2
2	20	5000	2	2	20	0	\
3	30	13000	1,2	3	30	13000	1,2
4	35	18000	1,3	4	35	23000	1,2,3
5	15	0	\	5	15	0	\

TABLE 5: Gini coefficients of each scheme.

Scheme	Gini coefficient
1	0.334
2	0.351
3	0.475
4	0.540

**6. Conclusion**

This paper develops a method to derive the optimal scholarship assignment scheme with the highest equity for university administrators. The method is applicable as (1) it meets the equity requirement that students who perform better ought to receive scholarships equal to or more than those received by less-achieving students; (2) the scholarship assignment eliminates the need of students to manually apply for specific scholarships, which is a time- and energy-consuming process; and (3) the equity of derived assignment schemes can be quantified through the Gini coefficient, and the scheme with the maximum equity (minimum Gini coefficient value) can be identified and selected. An illustrative example is adopted to show the applicability of the proposed method. The results indicate that the proposed method can obtain a set of feasible assignment schemes efficiently and then derive the optimal one with the highest equity.

**Data Availability**

The hypothetical data used to support the findings of this study are included within the article. No external data were used to support this study.

**Conflicts of Interest**

The authors declare that they have no conflicts of interest.

**Acknowledgments**

This study is supported by the Top-Notch Academic Programs Project of Jiangsu Higher Education Institutions (PPZY2015B148) and the Project of Educational Reform and Practice in Southeast University (2017-071).

**References**

- [1] P. E. Peterson, D. Myers, and W. G. Howell, *An Evaluation of the New York City School Choice Scholarships Program: The First Year*, Mathematica Policy Research, Washington, DC, USA, 1998.
- [2] S. L. Desjardins, “The impact of Washington State Achievers Scholarship on student outcomes,” in *Annual meeting of Association for the Study of Higher Education*, Vancouver, BC, Canada, 2009.
- [3] S. L. DesJardins and B. P. McCall, “The impact of the Gates Millennium Scholars Program on college and post-college related choices of high ability, low-income minority students,” *Economics of Education Review*, vol. 38, no. 1, pp. 124–138, 2014.
- [4] M. R. Anderson-Rowland, “Evaluation of a ten year life planning assignment for an academic scholarship success class,” in *Proceedings of the 41st Annual Frontiers in Education Conference: Celebrating 41 Years of Monumental Innovations from Around the World, FIE 2011*, pp. 1–7, 2012.
- [5] A. Porter, R. Yang, J. Hwang, J. McMaken, and J. Rorison, “The Effects of Scholarship Amount on Yield and Success for Master’s Students in Education,” *Journal of Research on Educational Effectiveness*, vol. 7, no. 2, pp. 166–182, 2014.
- [6] H. Forsyth and A. Cairnduff, “A scholarship of social inclusion in higher education: why we need it and what it should look like,” *Higher Education Research & Development*, vol. 34, no. 1, pp. 219–222, 2015.
- [7] P. Amorós, L. C. Corchón, and B. Moreno, “The scholarship assignment problem,” *Games & Economic Behavior*, vol. 38, no. 1, pp. 1–18, 2002.
- [8] G. Goastellec, “Globalization and implementation of an equity norm in higher education: Admission processes and funding

- framework under scrutiny,” *Peabody Journal of Education*, vol. 83, no. 1, pp. 71–85, 2008.
- [9] W. O. Lee and M. Manzon, “The issue of equity and quality of education in Hong Kong,” *The Asia-Pacific Education Researcher*, vol. 23, no. 4, pp. 823–833, 2014.
- [10] T. Pitman, “Unlocking the gates to the peasants: are policies of ‘fairness’ or ‘inclusion’ more important for equity in higher education?” *Cambridge Journal of Education*, vol. 45, no. 2, pp. 281–293, 2015.
- [11] X. Qu, W. Yi, T. Wang, S. Wang, L. Xiao, and Z. Liu, “Mixed-integer linear programming models for teaching assistant assignment and extensions,” *Scientific Programming*, vol. 2017, Article ID 9057947, 7 pages, 2017.
- [12] Z. Liu, S. Wang, and Q. Meng, “Optimal joint distance and time toll for cordon-based congestion pricing,” *Transportation Research Part B: Methodological*, vol. 69, pp. 81–97, 2014.
- [13] Z. Liu, S. Wang, W. Chen, and Y. Zheng, “Willingness to board: a novel concept for modeling queuing up passengers,” *Transportation Research Part B: Methodological*, vol. 90, pp. 70–82, 2016.
- [14] Z. Liu, S. Wang, B. Zhou, and Q. Cheng, “Robust optimization of distance-based tolls in a network considering stochastic day to day dynamics,” *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 58–72, 2017.
- [15] E. A. Mannix, M. A. Neale, and G. B. Northcraft, “Equity, equality, or need? The effects of organizational culture on the allocation of benefits and burdens,” *Organizational Behavior and Human Decision Processes*, vol. 63, no. 3, pp. 276–286, 1995.
- [16] S. Wang, “Efficiency and equity of speed limits in transportation networks,” *Transportation Research Part C: Emerging Technologies*, vol. 32, pp. 61–75, 2013.
- [17] S. Wang, Q. Meng, and Z. Liu, “Containership scheduling with transit-time-sensitive container shipment demand,” *Transportation Research Part B: Methodological*, vol. 54, pp. 68–83, 2013.
- [18] L. Anselmi, M. Lagarde, and K. Hanson, “Going beyond horizontal equity: An analysis of health expenditure allocation across geographic areas in Mozambique,” *Social Science & Medicine*, vol. 130, pp. 216–224, 2015.
- [19] S. Caleo, “When distributive justice and gender stereotypes coincide: Reactions to equity and equality violations,” *Journal of Applied Social Psychology*, vol. 48, no. 5, pp. 257–268, 2018.
- [20] J. S. Adams, “Inequity in social exchange,” *Advances in Experimental Social Psychology*, vol. 2, no. 4, pp. 267–299, 1965.
- [21] E. Walster, G. Walster, and E. Berscheid, *Equity: Theory and Research*, Allyn & Bacon, Boston, MA, USA, 1978.
- [22] D. M. Messick and K. P. Sentis, “Fairness and preference,” *Journal of Experimental Social Psychology*, vol. 15, no. 4, pp. 418–434, 1979.
- [23] ISI (Institute for Scientific Information), “Data from: Journal Citation Reports” (Dataset). Clarivate Analytics. Accessed July 20, 2018, <http://jcr.incites.thomsonreuters.com/JCRJournalHomeAction.action?>
- [24] S. Martello and P. Toth, “Algorithms for Knapsack problems,” *North-Holland Mathematics Studies*, vol. 132, pp. 213–257, 1987.
- [25] H. M. Salkin and K. C. A. De, “The knapsack problem: A survey,” *Naval Research Logistics*, vol. 22, no. 1, pp. 127–144, 2010.
- [26] T. Neutens, “Accessibility, equity and health care: Review and research directions for transport geographers,” *Journal of Transport Geography*, vol. 43, pp. 14–27, 2015.
- [27] H. R. Waters, “Measuring equity in access to health care,” *Social Science & Medicine*, vol. 51, no. 4, pp. 599–612, 2000.
- [28] T. F. Welch and S. Mishra, “A measure of equity for public transit connectivity,” *Journal of Transport Geography*, vol. 33, no. 33, pp. 29–41, 2013.
- [29] A. Sumalee, “Optimal toll ring design with equity constraint: An evolutionary approach,” *Journal of Eastern Asia Society for Transportation Studies*, vol. 5, pp. 1813–1828, 2003.
- [30] X. Sun, Z. Liu, and S. Chen, “The Equity Issue for Cordon-Based Congestion Pricing with Distance Toll,” in *Proceedings of the 2nd International Conference on Vulnerability and Risk Analysis and Management, ICVRAM 2014*, pp. 2310–2319, July 2014.
- [31] S. Martello, D. Pisinger, and P. Toth, “Dynamic programming and strong bounds for the 0-1 Knapsack Problem,” *Management Science*, vol. 45, no. 3, pp. 414–424, 1999.
- [32] P. Toth, “Dynamic programming algorithms for the Zero-One Knapsack problem,” *Computing*, vol. 25, no. 1, pp. 29–45, 1980.
- [33] D. Pisinger, “A minimal algorithm for the 0-1 knapsack problem,” *Operations Research*, vol. 45, no. 5, pp. 758–767, 1997.
- [34] F. Li, Z. L. Chen, and L. Tang, “Integrated Production, Inventory and Delivery Problems: Complexity and Algorithms,” *INFORMS Journal on Computing*, vol. 29, no. 2, pp. 232–250, 2017.
- [35] H. Liu and D. Z. W. Wang, “Locating multiple types of charging facilities for battery electric vehicles,” *Transportation Research Part B: Methodological*, vol. 103, pp. 30–55, 2017.
- [36] J. Chen, S. Wang, Z. Liu, and Y. Guo, “Network-based optimization modeling of manhole setting for pipeline transportation,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 113, pp. 38–55, 2018.

## Research Article

# Optimizing Price of Credit Default Swaps for Dynamic Project System of Public-Private Partnership

Ming Wu,<sup>1</sup> Wenya Lv,<sup>2</sup> and Qiuji Sun <sup>2</sup>

<sup>1</sup>*School of Management, University of Science and Technology of China, Anhui, China*

<sup>2</sup>*School of Management, Shanghai University, Shanghai, China*

Correspondence should be addressed to Qiuji Sun; [qiujisun@outlook.com](mailto:qiujisun@outlook.com)

Received 13 May 2018; Revised 25 June 2018; Accepted 8 July 2018; Published 16 July 2018

Academic Editor: Xinchang Wang

Copyright © 2018 Ming Wu 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.

Most project operations management belongs to the type of public-private partnership (PPP), which is usually dynamic. This paper aims to propose a method for optimizing the price of credit default swaps (CDS) for the dynamic PPP system. This study investigates the credit risk measurement of PPP project financing and the pricing of risk mitigation instruments which are widely used in the case of immature markets in the early stage of China's PPP development. Based on the credit risk measurement theory of the corporate and debt ratings, this paper considers the differences in various credit enhancement methods in the equity-like debt agreement and determines the credit rating of the equity-like debt in PPP projects. Some optimization methods are also proposed to derive the probability of default, so as to determine the price of the credit risk mitigation instrument of CDS which is based on the equity-like debt.

## 1. Introduction

As a long-term partnership of public sector (government) and private sector (for-profit or nonprofit companies) which is established by signing formal agreements, public-private partnership (PPP) has the advantages of revitalizing existing assets, transforming government functions, and improving the quality of public services in the provision of public goods or services. In recent years, the system established for PPP has been very mature and complete in many developed countries. Compared with developed countries, developing countries have greater needs in the area of large-scale infrastructure construction, and governments are more likely to fall into the embarrassing situation of insufficient funds. Therefore, PPP related policies have been intensively published in China since 2013 and PPP has been widely used in various fields of China's infrastructure and public service construction. The World Bank [1] comprehensively considered asset ownership, operating rights, investment relations, business risks, and contract terms and then classified the PPP contracts into six types: service contract, management contract, Lease, Concession, BOT/BOO, and Divestiture. Due to the long franchise period of the PPP project, the huge amount of investment, the

different participation purposes of the project participants, and the complicated financing structure, the use of PPP to build public infrastructure and provide services will face multiple risks. Therefore, it is necessary to correctly identify risks. Only after identifying risks effectively can the risk be reasonably distributed. In the identification and evaluation of PPP project risks and reasonable distribution, a large number of scholars have conducted relevant research.

In the area of PPP risk identification, readers can refer to the relevant literatures given by Hastak and Shaked [2], Grimsey and Lewis [3], Li et al. [4], Shen et al. [5], Zhang [6], Ameyaw and Chan [7], and Ghorbani et al. [8]. Among them, Li et al. [4] divided the risk into macrolevels, medium levels, and microlevels. Zhang [6] has classified PPP risks in more detail: political, economic, and social environment, natural environment, third-party infringement, engineering decision-making and preparation, payment, supervision, completion and handover process, coordination, and relations. This helps participants who are not familiar with the project identify risks quickly. Ameyaw and Chan [7] analyzed six cases of the PPP project and classified 40 risk factors of the PPP project into eight categories, the common risks that practitioners are concerned with include the weak regulatory

and monitoring regime, financing, absence of risk allocation mechanisms, inexperience in PPPs, public opposition, and delayed and nonpayment of bills. Ghorbani et al. [8] divided the risk into the development phase, the operation phase and the entire life cycle. Each phase has different risk categories and factors.

In the PPP risk assessment, representative researches are shown as follows. Gumming and Macintosh [9] created the general theory of venture capital exit and considered that the best exit timing of social capital is that the marginal cost of investment is equal to the marginal return. The investor must determine the equity exit mechanism within the project duration. Khallaf et al. [10] applied game theory to analyze the interactions between parties to understand dynamic outcome of the associated risks in PPP projects. A case study of the railway project in Tanzania was also presented. Kumar et al. [11] applied a standard risk-analysis model of NPV-at-risk tool supported by Monte Carlo Simulation to the real-world PPP based highway infrastructure projects, which is able to identify which source of uncertainty has the most influence on the project's financial returns and what is the actual relationship between the critical influencing parameters and associated NPV. Wu et al. [12] combined the 2-dimension linguistic information of fuzziness and randomness with the cloud model to put forward a risk assessment framework of PPP waste-to-energy incineration projects.

In terms of PPP risk allocation, representative researches are shown as follows. Ewijk [13] identified, evaluated, and allocated risks by establishing a camp model for the operation of a PPP project. Carbonara et al. [14] established a win-win model of fair risk sharing during the franchise period of the PPP project by adopting the NPV method. Nasirzadeh et al. [15] calculated the proportion of risk sharing between the owner and the contractor based on system dynamics, fuzzy logic, and other methods. Almarri and Blackwell [16] studied and proposed two methods to improve the risk sharing and investment evaluation of PPP of large-scale greening projects. The first is to review the contract of the PPP project; the second is to improve the risk simulation method by improving input variables and characteristics.

However, the development of PPP is not mature in China, due to the monopoly of state-owned capital in the area of infrastructure construction, social capital is relatively unfamiliar to the field and lacks the relevant experience of investment, financing, and operational management. The enthusiasm for participating in PPP projects is limited. Local governments usually disguise their liabilities in the name of PPP to meet the large amount of funds and long-term capital requirements. The phenomenon of "equity-like debt" is ubiquitous in the process of project construction. The equity-like debt in PPP projects refers to the fact that, in a public-private partnership construction project, the investor nominally obtains the equity of the project company, while the financing side provides the investor with a revenue commitment and guarantees the mature repurchase, which is actually an indirect debt financing. Therefore, studying the credit risk of debt financing of PPP projects has become a major issue for each investor in the PPP projects. Therefore, Credit Default Swaps (CDS), which are subject to PPP

projects, have become important financial instruments for identifying and hedging credit risk.

The common CDS pricing models are mainly divided into structured models and parsimonious models. The parsimony model considers that there is no clear correlation between the default condition of the reference entity and its value. The default phenomenon of the reference entity is random and the market information is used to directly calculate the risk-neutral probability of default under probability model. Shuai Li et al. [17] presented a general approach to PPP debt financing and CDS pricing in mature markets. In the typical nonmature Chinese market, due to the lack of the secondary market price of the project company's equity, the complete disclosure of operating and financial information, and the related complete financing service system, we combine the corporate credit risk measurement with the measurement of supporting credit of third-party's forward repurchase equity and making up principal and interest shortfalls, based on the bond yield curve of the Chinese market, we use the Jarrow-Turnbul binary tree model to derive its default probability and determine the price of CDS which is based on the debt.

## 2. Credit Risk Measurement of Equity-Like Debt in PPPs

Since the underlying rights and obligations (such as equity financing and the committed repurchasing of nominal equity, making up of investment principal and interest) have been signed by both parties and are subject to the "Contract Law," the transaction structures can affect the credit rating of the company's debts. Although project financing is in the form of equity, the guarantee terms for investment principal and interest, such as repurchases and making up shortfalls, make the financing risk essentially have the property of debt financing. Therefore, the default risk of equity-like debt needs to be considered from the relevant corporate credit risk and the characteristics of transaction structures; then its corresponding credit risk rating can be determined.

In addition to considering the default risk of debt entities, when determining the credit rating of equity-like debt, we also need to take into account trading structure risks (such as the risk of equity pricing, insufficient capital, project control risk, repurchase risk, and supporting credit loss), participating institution risks, and overall systemic market risks. The debt rating is determined based on a combination of the above factors. Therefore, in order to simplify the quantification and summarize its rating method, we take the support rating of the most critical credit enhancement measures as an example to give a corresponding debt rating. All other factors can be considered similarly; then a comprehensive rating of the debt will be determined.

*2.1. Corporate and Debt Rating Methods.* We proceed from the general framework of the corporate credit rating and consider the credit rating of the project company based on the default probability model, and the corresponding credit rating will be determined by considering the factor of equity-like debt. For the basic framework of the corporate credit rating in this paper, we quote from the market generally

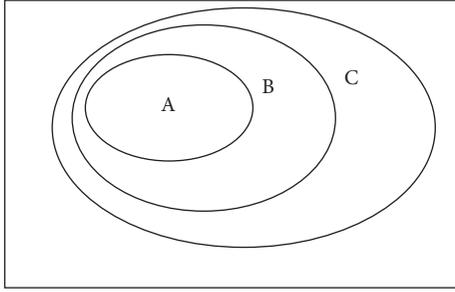


FIGURE 1: The relationship between macroeconomy, industry, and corporation.

accepted frameworks and variables which have been published by major international rating companies such as S&P, Moody's, and Fitch. The definition of default in the default probability model refers to the basic definition of default in the Basel Capital Accord. Among the existing default probability models, we use the logistic regression model which has been widely used by considering the accuracy, robustness, and performance.

The general framework of the credit rating model mainly refers to which factors have been considered to analyze the corporate credit risk and the role and interaction of these risk factors in the occurrence of default probability. The corporate credit rating of the bond market generally follows the principles of independence, objectivity, consistency, and reasonable prudence. The basic framework of corporate credit analysis includes three aspects: the macroeconomic environment, the industry conditions, and the operating financial status of the issuer. The issuer's financial status is the result of the entity operating in a specific industry under the macroeconomic environment, which can measure the credit status of the issuer to a large extent. However, the financial data has a certain lag, which cannot reflect the macroeconomic environment and industry conditions. When determining the credit rating, it is required that the credits of different industries, the different subject credits in the industry, and the issuer's own time series under the macroindustry are comparable. Therefore, the corporate credit analysis method in this paper is based on the fundamentals of the issuing entity such as operational and financial information. Meanwhile, the adjustment effects of the industry and the macroeconomic environment are also needed to be considered.

The deterioration of the macroeconomic environment is denoted by A, the deterioration of the industry is denoted by B, and the occurrence of corporate default is denoted by C. The relationship of the macroeconomy, industry, and corporation is shown in Figure 1.

So we obtained  $P(B | A) = 1$ ,  $P(C | B) = 1$ .

Based on the full probability formula,  $P(C)$  can be calculated by the following formula:

$$\begin{aligned} P(C) &= P(C | B) P(B) + P(C | \bar{B}) P(\bar{B}) \\ &= P(B) + P(C | \bar{B}) P(\bar{B}) \end{aligned}$$

$$\begin{aligned} &= P(B | A) P(A) + P(B | \bar{A}) P(\bar{A}) \\ &\quad + P(C | \bar{B}) P(\bar{B}) \\ &= P(A) + P(B | \bar{A}) P(\bar{A}) + P(C | \bar{B}) P(\bar{B}) \\ &= P(A) + P(B | \bar{A}) (1 - P(A)) \\ &\quad + P(C | \bar{B}) (1 - P(B)) \end{aligned} \tag{1}$$

In the default probability model, we assume that if there is a problem with the macroeconomy, the corporate entity must have credit risks. Therefore, for the corporate entity and the industry, the probability of default risk arising from the macroeconomic environment and the probability of default risk when the macroeconomy is not at risk have additivity, so the probability of default risk arising from the macroeconomy will be used as an adjustment factor to affect the upper limit of the industry and corporate's rating.

The assumption applies to the relationship between the industry and the corporate. The probability of default risk arising from the industry will be used as an adjustment factor to affect the upper limit of the corporate's rating.

Measuring the probability of the corporate credit default risk  $P(C | \bar{B})$  with the entity's operational and financial information will be the key to the corporate rating. For the aspect of the operational information, the basic conditions of the business entity (such as capital injection size, owner nature, and social influence), the level of managers, the ability to control the company's operations and resources, and the decentralized operations are generally considered when determining the credit risk. For the aspect of financial information, some factors are often considered, such as the profitability of the entity, the existing debt structure, especially the financial leverage, the financial status of the company, and the protection of the debt. In the following section, we will set the indicators of these influencing factors to be the explanatory variable X of the generalized linear model, which affect the value of the corporate default probability  $= P(C | \bar{B})$ .

Before estimating the default probability model, we need to introduce the definition of breach of contract. According to the Basel Accord, when borrowers are not willing to fulfill their credit obligations in full for over 90 days, they can be categorized into breach of contract. In addition, Standard & Poor's generally define that if borrowers have not paid the bond within a grace period of 10 to 30 days, it will be considered as a default event. European credit institutions generally considers that if borrowers fail to pay the bond overdue 120 days, it will be considered as a default event. Moody's defines a default event if it is overdue for more than one day. In summary, this paper defines that if the project company is unwilling or unable to perform its debt payment obligations (including principal, interest, and fees stipulated in related agreements) in full more than 90 overdue days and is still unable to perform its debt payment obligations in full

TABLE 1: Credit ratings with the corresponding default probability of Moody, S&amp;P, and domestic commercial banks.

Rating Symbols	Default probability interval of Moody		Default probability interval of S&P		Default probability interval of domestic commercial banks	
AAA	0.00%	0.02%	0.00%	0.02%	0.00%	0.01%
AA	0.02%	0.05%	0.02%	0.04%	0.01%	0.5%
A	0.05%	0.15%	0.04%	0.11%	0.5%	1%
BBB	0.15%	0.50%	0.11%	0.27%	1%	2%
BB	0.50%	2.00%	0.27%	0.88%	2%	5%
B	2.00%	7.00%	0.88%	4.86%	5%	10%
CCC	7.00%	15.00%	4.86%	15.94%	10%	20%
CC	15.00%	20.00%	15.94%	17.61%	20%	100%
C	20.00%	100.00%	17.61%	19.10%	100%	100%
D		100.00%	19.10%	20%	100%	100%

within the grace period ( $t$  days), it will be considered as PPP's breach of contract.

We use a logistic regression model to describe the default probability  $Y = P(C | \bar{B})$  of the corporate entity in the absence of macroeconomic and industry risk events. The parameters of the model are estimated and  $P(C | \bar{B})$  is obtained by substituting the sample data and applying the maximum likelihood estimation method.

Similarly, we can use the same method to describe the probability  $Z = P(B | \bar{A})$  of general default of the industry in the absence of a macroeconomic risk event and obtained the probability of industry risk  $P(B | \bar{A})$ . Then we substitute  $P(B | \bar{A})$  into (1) and calculate the corporate default probability  $P(C)$ .

After obtaining the probability of default, to obtain the corresponding credit rating of the corporate, it is necessary to correspond with the main scale of the credit rating. At present, the main scale of the rating agencies in the financial market is based on a large amount of historical data and years of experience in rating, which has both an explanatory and statistically significant economic meaning. The corporate credit rating is determined by referring to the main scale of rating agencies.

We compared the credit ratings of Moody's and Standard & Poor's and domestic commercial banks' internal rating methods with their corresponding default probability, which is shown in Table 1.

Considering that the credit risk management system of domestic commercial banks is becoming increasingly standard, and it can reflect the actual situation of domestic companies' breach of contract accurately, we will refer to the main scale of the domestic commercial banks to determine the credit rating symbol which corresponds to the default probability of the PPP project company.

*2.2. Credit Rating after the Credit Enhancement under Various Financing Structures of Equity-Like Debt.* In some PPP projects, the major shareholders of the project company and the local government often support the project company's operations and financing in many ways, such as jointing guarantees and making up shortfalls of principal and interest.

Therefore, for this type of project, besides the corporation's individual rating, the support rating is also needed to be considered.

Support rating refers to the promotion of individual ratings by considering supporting factors. To determine the degree of individual rating improvement, it is necessary to evaluate the status and strategy of individuals or local governments and the possible manifestations of support for the individual default (such as a predefault support and a postdefault bailout). For state-owned enterprise groups, the status and role of the group in the local state-controlled enterprise system are also needed to be considered when assessing the degree of support given by the local government and determining the government support rating.

Due to the nonstandardized characteristics of the financial instruments of equity-like debt, on the basis of identifying the corporate credit risk, it is necessary to determine the support rating from the following three aspects. The first is the comprehensiveness and rigorousness of the repurchase arrangement or the shortfall replenishment measure on the debt agreement, which is used to ensure that the contract that contains the payment obligation becomes a debt of the supporter. The second is to compare the repurchase or difference between complemented debts of the agreement with the ordinary debts of the supporter and set that whether the right holders of repurchase or difference complement can claim the creditor's rights declaration to be the reference for the priority or inferior layer of the debts. The third is to inspect the financial quality, economic value, and social value of the project company and pay attention to the value of relief in extremely unfavorable situations. It is used to assess the financial risk, the sustainability of its operations, and the strength of obtaining external support of the project.

The above three aspects are the key aspects of measuring the support rating of equity-like debt. In the following section, we will analyze the effectiveness of the credit enhancement measures by combining several major credit enhancement methods. Since the process of generating the credit rating symbol ultimately needs to be combined with the experience of rating analysts, the descriptions of several major credit enhancement methods are given in Table 2.

TABLE 2: The description of four main credit enhancement methods.

Credit enhancement methods	Description
Project company's full premiums repurchase + third-party's joint guarantees	The project company repurchase the equity (beneficial right or fund share) held by the financial instruments of equity-like debt (including trust plans, asset management plans, fund shares, etc.) according to the principal and income premium agreed in the investment contract, and the third-party agency (such as sponsoring group companies or local governments) undertakes joint guarantees on the repurchase obligations of the project company. The project company assumes the obligation to pay the debt, and the third-party institution which acts as the trustworthy entity will be the target of the rating. This is a guarantee security method, and the credit enhancement effect is based on the credit quality of the project company and the third-party organization. Any party with a high credit rating is sufficient to support the high level of rating of financing instruments. Therefore, the final credit rating is determined based on the higher rating of the two parties.
Project company's full premiums repurchase + third-party's making up principal and interest shortfalls	In this case, the guarantee effect of the third-party institution's commitment letter must be first determined. If the commitment letter of making up principal and interest shortfalls has strict terms, and the triggering mechanism is conducive to implementation, it can be defined that the supplemental commitments have substantial safeguard effect, but the nature of the liability is the supplementary responsibility instead of the repayment liability, so the replenishment obligor has the counterplead right of telling first and cannot constitute joint guarantees. The "Enterprise Bankruptcy Law" stipulates that the debtor's unexpired guarantee liability is deemed to have been expired and needs to be paid off in advance. However, in the case of the bankruptcy of the third party, the guarantor who bears the supplemental responsibilities still has the problem of handling the counterplead right of telling first. In order to protect the fairness of transactions and achieve the original intention of setting rights, if the guarantor's counterplead right is cancelled, the creditor may first get the repayment from the guarantor. Judging from the legal practice, as the supplement number of debts is difficult to define, the creditor's claims can be denied by the bankruptcy administrator. Therefore, under this situation, if there is a substantial guarantee effect, and the terms in the agreement substantially cancel the guarantor's counterplead right, the higher credit rating of the third party can support the debt rating to the corresponding level. If there is only a substantial guarantee effect and the credit rating of the third party institution is higher than that of the project company, it can support the debt which is higher than the project company level but not higher than the third party institution's credit rating. In other cases, third-party agencies are not considered, and the credit rating of the debt is determined only based on credit rating of the project company.
Project company's principal repurchase + third-party's making up principal and interest shortfalls	Project company repurchase the equity, income right and fund share of the financing instrument, The third party agency has the liability for the repurchase shortfalls of the project company and the interest of the financing instrument. However, as the project company acts as the debtor and its repurchase obligation cannot cover the principal and interest of the financing instrument, its corporate credit has little impact on the credit rating of the financing instrument. The main consideration is the substance of the third-party's replenishment commitment shortfalls. If the credit rating of the third-party company is high, the debt rating can be appropriately close to its credit rating.
Project company's principal repurchase + profit supplement or project company's full premiums repurchase of principal and interest	The project company acts as the debtor and assumes all payment obligations. Therefore, the credit rating of the debt is largely determined by the rating of the project company.

For ease of use, it is assumed that the rating of the project company is  $R_M$ , and the rating of the third-party credit agency is  $R_N, M, N = 1, 2, \dots, 10, M \geq N$ , and the corresponding credit rating symbol of  $R_M$  and  $R_N$  is shown in Table 3.

The support ratings of the above four typical credit enhancement methods are shown in Table 4.

After obtaining the credit rating of the project company and the third party through the method described in Section 2.1, the final credit rating (denoted by  $R_k$ ) of the debt can basically be obtained by referring to Table 4.

### 3. Analysis of the Price of Credit Risk Mitigation Instrument of CDS

After obtaining the value of credit rating, we consider the actual situation of domestic debt default, based on the yield curve of the debenture bond, and we derive the default probability of different maturities by applying the Jarrow-Turnbull model.

The Jarrow-Turnbull model is also known as the Jarrow-Turnbull binary tree model. The model considers the impact of interest rate fluctuation processes and default processes on

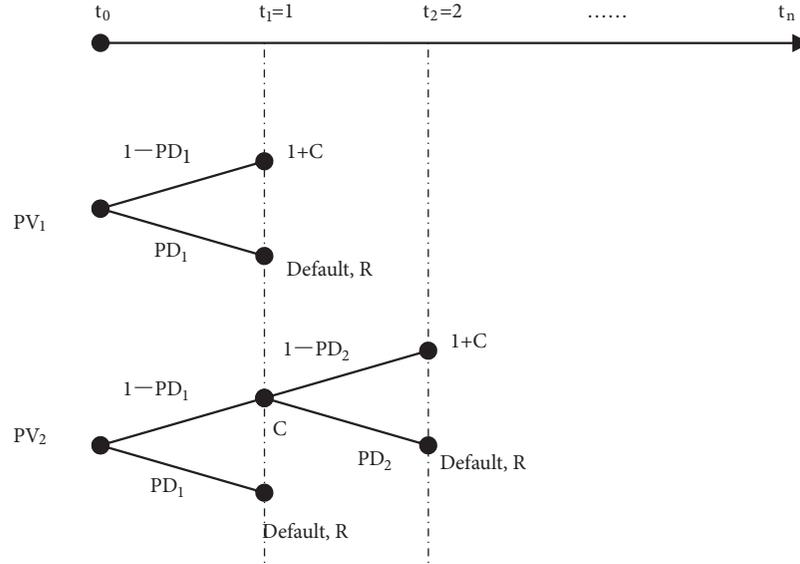


FIGURE 2: Binary tree-solution of default probability.

TABLE 3: List of credit rating symbols.

Ranking	Rating Symbols
1	AAA
2	AA
3	A
4	BBB
5	BB
6	B
7	CCC
8	CC
9	C
10	D

bond prices. The model assumes that when the breach event occurs, the seller only pays compensation on the agreed fixed trading day.

The bond pricing formula in the Jarrow-Turnbull model [18] is shown as follows:

$$\begin{aligned}
 V_{JT}(0, t_n) = & \sum_{i=1}^n P(0, t_i) Q(0, t_i) C \Delta t \\
 & + P(0, t_n) Q(0, t_n) + R \\
 & \cdot P(0, t_n) [1 - Q(0, t_n)]
 \end{aligned} \quad (2)$$

In formula (2),  $V_{JT}(0, t_n)$  represents the present value of the bond at time  $t_n$ ;  $P(0, t_i)$  represents the discount rate at time  $t_i$ ,  $P(t_i, t_i) = 1$ ;  $Q(0, t_i)$  denotes the survival rate of bond at time  $t_i$ ;  $C$  denotes the nominal interest rate of bond;  $\Delta t$  denotes the interval between adjacent payment days;  $R$  represents the rate that the bond can be recovered when a credit event occurs.

The idea of using this model to price the CDS is to derivate the default probability curve through the credit spreads of the government bond and the debenture bond of the same rating, and then the CDS price is calculated by using the back-off default probability.

**3.1. Calculating Probability of Default.** For the sake of simplicity, the face value of the bond is set to be 1; the coupon rate of the bond is set to be  $C$  (annual interest payment); the rate that the bond can be recovered when a credit event occurs is denoted by  $R$ ; the relationship between  $R$  and the default loss rate ( $LGD$ ) is  $R = 1 - LGD$ .

We mainly calculate default probabilities of debenture bonds with different maturities through variables such as present value of bonds  $PV$ , default recovery rate  $R$ , and risk-free interest rate  $r$ . We use the Chinese government bond yield curve as a risk-free rate. The binary tree method is used to analyze the future cash flow of credit bonds, as shown in Figure 2.

In Figure 2,  $PD_1$  indicates the probability of a default occurring before time  $t_1$ ;  $PD_2$  indicates the probability of no default occurring before time  $t_1$  and a default occurring between times  $t_1$  and  $t_2$ . If the current price of the debenture bond is set as the present value of the future cash flow income, the price of the 1-year debenture bond can be calculated as

$$PV_1 = \frac{(1 + C)(1 - PD_1) + R \cdot PD_1}{1 + r_1} \quad (3)$$

We can calculate the value of  $PD_1$  by solving the equation

$$PD_1 = \frac{(1 + C) - PV_1(1 + r_1)}{1 + C - R} \quad (4)$$

TABLE 4: Credit rating after the credit enhancement of various financing structures of equity-like debt.

Credit Enhancement	Conditions	Support Rating
Project company's full premiums repurchase + third-party's joint guarantees	-	$R_N$
Project company's full premiums repurchase + third-party's making up principal and interest shortfalls	(a) substantial guarantee effect, and the terms in the agreement substantially cancel the guarantor's counterplead right	$R_N$
	(b) only a substantial guarantee effect	$R_k, N \leq k \leq M$
	(c) no guarantee effect	$R_M$
Project company's principal repurchase + third-party's making up principal and interest shortfalls	(a) only a substantial guarantee effect	$R_k, N \leq k$
	(b) no guarantee effect	-
Project company's principal repurchase +profit supplement, or project company's full premiums repurchase of principal and interest	-	$R_M$

Similarly, the price of a 2-year debenture bond can be calculated as follows:

$$PV_2 = \frac{C(1 - PD_1) + R \cdot PD_1}{1 + r_1} + \frac{[(1 + C)(1 - PD_2) + R \cdot PD_2](1 - PD_1)}{(1 + r_2)^2} \quad (5)$$

The value of  $PD_1$  is obtained by solving the (4), after getting the value of  $PD_1$ , and the value of  $PD_2$  is obtained by solving the (5). By analogy, we can use the 3-year debenture bond price to calculate the value of  $PD_3$ . Similarly, the value of  $PD_4, PD_5$ , and  $PD_n$  ( $n > 5$ ) can be calculated by using the same method. In the above derivation process,  $r_1$  is the 1-year risk-free interest rate,  $r_2$  is the 2-year risk-free interest rate, and  $r_n$  is the n-year risk-free interest rate.

Based on the above process, we can obtain the default probability curve of the same credit rating bond.

Since the trading volume of bonds in the secondary market of China bonds (including interbank and exchange-traded markets) is limited, all transaction data are concentrated within a period of 0.5 to 5 years from the expiration date of the bonds, and the transaction data for more than 5 years are extremely limited. The present value of  $PV_1, PV_2, \dots, PV_n$  corresponding to the  $R_k$  rating is often not exactly reached at the condition of an integer year of maturity period. Even if there is exactly a trading price for the whole year, due to the small trading volume, there are often a large number of liquidity premiums that cause prices to deviate, so the present value of a bond that is taken from the market price tends to deviate greatly. Therefore, we use the China Bond Yield Curve to estimate the present value of the bond price.

The China Bond Yield Curve is a yield curve published by the Central Government Securities Depository Trust & Clearing Corporation, the curve is formed based on bond transactions, quotations, and institutional valuation data in the Chinese market, it is used to characterize the various maturity structures of Chinese bonds. The data includes the interbank bond market settlement data, transaction data for exchange bonds, bilateral quotation at the counter, bilateral quotes for the interbank bond market and the profitability

valuation data of some of the core members. In the current Chinese bond market, the price discovery mechanism is still incomplete (especially when bilateral interbank quotation does not exert its price discovery function properly), the Chinese bond yield curve mainly refers to the profitability valuation data of some of the core members. Therefore, whether in the maturity stage of better liquidity or in the maturity period of poor liquidity, the China Bond Yield Curve mostly truly reflect the general value of bonds and market conditions, and it has become an important basis for the basic yield curve of the Chinese bond market.

We use the debt credit rating  $R_k$  of the PPP fake equity real debt which is obtained in Section 2.2 to corresponds to the spot rate yield curve of corporate bonds rated  $R_k$  in the China Bond Yield Curve and get yield rate  $y_{k,i}$  in which the maturity period  $i$  is an integer ( $i = 1, 2 \dots, n$ ), and then  $PV_i$  is calculated as follows:

$$PV_i = \frac{1}{(1 + y_{k,i})^i} + \sum_{j=1}^i \frac{C}{(1 + y_{k,j})^j} \quad (1 \leq j \leq i) \quad (6)$$

The corresponding default probability sequence  $PD_i$  can be obtained by substituting the present value of the bond obtained by (6) into (4) and (5).

### 3.2. CDS Price in the Case of One-Time Payment of Premiums.

In this case, we assume that the CDS buyer pays the CDS premium once at the beginning of the transaction, and the prepaid premium rate at time  $t_0$  is denoted by  $S$ . Once a credit event occurs, the delivery will be made in cash on the next future premium payment date of the credit event, and the CDS will be terminated. We simulate the future cash flow during the CDS duration by using the binary tree method, which is shown in Figure 3.

On the other hand, at the beginning of the transaction, the current CDS value is zero for the buyer and the seller, and the CDS rate  $S_1$  is determined by

$$S_1 = \frac{PD_1(1 - R)}{1 + r_1} \quad (7)$$

So  $S_1$  is the quote for 1-year CDS based on the above bond.

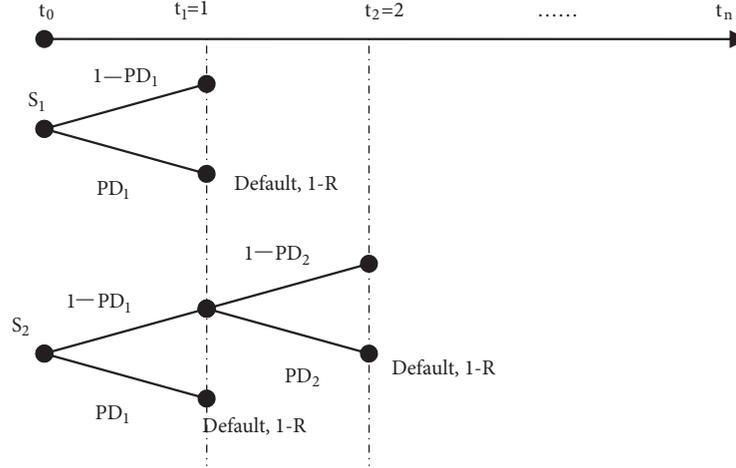


FIGURE 3: Binary tree-CDS pricing solution (one-time payment of premiums).

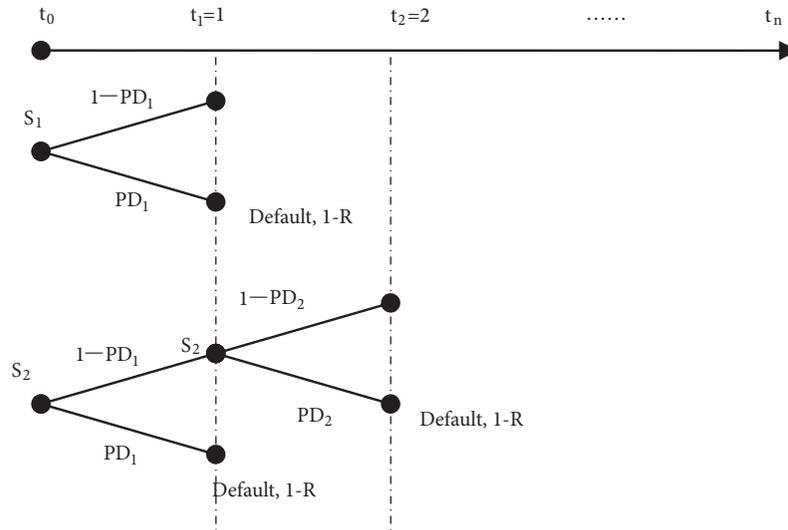


FIGURE 4: Binary tree-CDS pricing solution (premium payment in installments).

Similarly, the net future cash flow for a 2-year CDS can be described as follows:

$$S_2 = \left[ \frac{PD_1(1-R)}{1+r_1} + \frac{(1-PD_1)PD_2(1-R)}{(1+r_2)^2} \right] \quad (8)$$

$S_2$  can be calculated by the

$$S_2 = \frac{PD_1(1-R)}{1+r_1} + \frac{(1-PD_1)PD_2(1-R)}{(1+r_2)^2} \quad (9)$$

By analogy, we can obtain the prices of 3-year, 4-year, and  $n$ -year CDS and calculate the value of  $S_3$ ,  $S_4$ ,  $S_n$ .

**3.3. CDS Price in Case of Premium Payment in Installments.** Assuming that CDS premiums are paid in installments, they are paid at the beginning of each payment interval, and the payment interval is equal to the bond interest payment interval. Here we assume that the interest payment interval

for the bond is one year, so for a CDS with a contract term of  $n$  years, the buyer pays premiums at time  $t_0 = 0$ ,  $t_1 = 1, \dots, t_{n-1} = n-1$  and the interval of the adjacent payment time is  $\Delta t = t_{i+1} - t_i = 1$ , and the premium rate is denoted by  $S$ , which means that the CDS buyer needs to pay  $S$  times the face value of the bond of the CDS seller's assets each year. The future cash flow during the CDS duration is shown in Figure 4.

Considering the 1-year CDS, the net present value of the future cash flow of the CDS can be described as follows:

$$S_1 = \frac{PD_1(1-R)}{1+r_1} \quad (10)$$

$S_1$  can be calculated by

$$S_1 = \frac{PD_1(1-R)}{1+r_1} \quad (11)$$

Similarly, the net future cash flow for a 2-year CDS can be determined as follows:

$$\left[ S_2 + \frac{S_2(1 - PD_1)}{(1 + r_2)^2} \right] - \left[ \frac{PD_1(1 - R)}{1 + r_1} + \frac{(1 - PD_1)PD_2(1 - R)}{(1 + r_2)^2} \right] \quad (12)$$

$S_2$  can be calculated by the formula

$$S_2 = \frac{[PD_1(1 - R)/(1 + r_1) + (1 - PD_1)PD_2(1 - R)/(1 + r_2)^2]}{[1 + (1 - PD_1)/(1 + r_2)^2]} \quad (13)$$

By analogy, we can obtain the prices of 3-year, 4-year, and  $n$ -year CDS and calculate the value of  $S_3, S_4, S_n$ .

#### 4. Conclusions

For the widely adopted PPP mode in the project operations management, how to design some mechanism to improve the operation efficiency is one of the important decision problems for project managers. Since the PPP project is a dynamic system, the decision is usually very complex. This paper proposes a method for optimizing the price of credit default swaps (CDS) for the dynamic PPP system. This study investigates the credit risk measurement of PPP project financing and the pricing of risk mitigation instruments which are widely used in the case of immature markets in the early stage of China's PPP development. Based on the credit risk measurement theory of the corporate and debt ratings and the differences in various credit enhancement methods in the equity-like debt agreement, this paper proposes the credit rating of the equity-like debt in PPP project. Some optimization methods are also proposed to derive the probability of default, so as to determine the price of the credit risk mitigation instrument of CDS which is based on the equity-like debt.

#### 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 paper is sponsored by the National Natural Science Foundation of China (71422007 and 71671107).

#### References

- [1] The World Bank, *Selecting An Option for Private Sector Participation*, 1997.

- [2] M. Hastak and A. Shaked, "ICRAM-1: model for international construction risk assessment," *Journal of Management in Engineering*, vol. 16, no. 1, pp. 59–69, 2000.
- [3] D. Grimsey and M. K. Lewis, "Evaluating the risks of public private partnerships for infrastructure projects," *International Journal of Project Management*, vol. 20, no. 2, pp. 107–118, 2002.
- [4] B. Li, A. Akintoye, P. J. Edwards, and C. Hardcastle, "Critical success factors for PPP/PFI projects in the UK construction industry," *Construction Management and Economics*, vol. 23, no. 5, pp. 459–471, 2005.
- [5] L.-Y. Shen, A. Platten, and X. P. Deng, "Role of public private partnerships to manage risks in public sector projects in Hong Kong," *International Journal of Project Management*, vol. 24, no. 7, pp. 587–594, 2006.
- [6] W. Zhang and Y. Bai, "Risk allocation and dispute resolution mechanisms in chinese public projects: an empirical study," in *Proceedings of the ICCREM 2013: Construction and Operation in the Context of Sustainability*, pp. 637–649, 2013.
- [7] E. E. Ameyaw and A. P. C. Chan, "Identifying public-private partnership (PPP) risks in managing water supply projects in Ghana," *Journal of Facilities Management*, vol. 11, no. 2, pp. 152–182, 2013.
- [8] A. Ghorbani, M. Ravanshadnia, and M. B. Nobakht, "A survey of risks in public private partnership highway projects in Iran," in *Proceedings of the ICCREM 2014: Smart Construction and Management in the Context of New Technology*, pp. 482–492, September 2014.
- [9] D. J. Cumming and J. G. MacIntosh, "Venture capital investment duration in Canada and the United States," *Journal of Multinational Financial Management*, vol. 11, no. 4–5, pp. 445–463, 2001.
- [10] R. Khallaf, N. Naderpajouh, and M. Hastak, "Interactional analysis for two-party risk assessment in Public Private Partnerships (PPP)," in *Proceedings of the World Construction Symposium 2016: Greening Environment, Eco Innovations & Entrepreneurship*, 2016.
- [11] L. Kumar, A. Jindal, and N. R. Velaga, "Financial risk assessment and modelling of PPP based Indian highway infrastructure projects," *Transport Policy*, vol. 62, pp. 2–11, 2017.
- [12] Y. Wu, C. Xu, L. Li, Y. Wang, K. Chen, and R. Xu, "A risk assessment framework of PPP waste-to-energy incineration projects in China under 2-dimension linguistic environment," *Journal of Cleaner Production*, vol. 183, pp. 602–617, 2018.
- [13] C. van Ewijk and P. J. G. Tang, "How to price the risk of public investment?" *Economist*, vol. 151, no. 3, pp. 317–328, 2003.
- [14] N. Carbonara, N. Costantino, and R. Pellegrino, "Concession period for PPPs: A win-win model for a fair risk sharing," *International Journal of Project Management*, vol. 32, no. 7, pp. 1223–1232, 2014.
- [15] F. Nasirzadeh, M. Khanzadi, and M. Rezaie, "Dynamic modeling of the quantitative risk allocation in construction projects," *International Journal of Project Management*, vol. 32, no. 3, pp. 442–451, 2014.
- [16] K. Almarri and P. Blackwell, "Improving Risk Sharing and Investment Appraisal for PPP Procurement Success in Large Green Projects," *Procedia - Social and Behavioral Sciences*, vol. 119, pp. 847–856, 2014.
- [17] S. Li, D. Abraham, and H. Cai, "Infrastructure financing with project bond and credit default swap under public-private

partnerships,” *International Journal of Project Management*, vol. 35, no. 3, pp. 406–419, 2017.

- [18] R. A. Jarrow and S. M. Turnbull, “Pricing derivatives on financial securities subject to credit risk,” *Journal of Finance*, vol. 50, pp. 53–85, 1995.

## Research Article

# Complexity Analysis of Dynamic Cooperative Game Models for Supply Chain with the Remanufactured Products

Jianwei Chang<sup>1</sup> and Liuwei Zhao <sup>2</sup>

<sup>1</sup>School of Finance and Economics, Jiangsu University, Zhenjiang, Jiangsu 212013, China

<sup>2</sup>Computational Experiment Center for Social Science, School of Management, Jiangsu University, Zhenjiang, Jiangsu 212013, China

Correspondence should be addressed to Liuwei Zhao; [136901672@qq.com](mailto:136901672@qq.com)

Received 2 April 2018; Revised 23 May 2018; Accepted 4 June 2018; Published 3 July 2018

Academic Editor: Lu Zhen

Copyright © 2018 Jianwei Chang and Liuwei Zhao. 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 focused on the supply chain within the remanufacturing production system, which is composed of one manufacturer and one retailer. In this case, the manufacturer is responsible for the production of new products and the remanufactured products. This system can be regarded as a coupling dynamics of the forward supply chain of Stackelberg game model. Based on the analysis, some dynamic phenomena such as bifurcation and chaos were found. Numerical simulations and the maximum Lyapunov exponent were therefore utilized to provide experimental evidence for the complicated behaviors of the system evolution. The findings of the study revealed that the equilibrium of the system can lose stability via flip bifurcation or Neimark-Sacker bifurcation and that time-delayed feedback control is appropriate for stabilizing the chaotic behaviors of the system.

## 1. Introduction

In recent years, numerous people are paying high attention to the environment and sustainable development. This accentuates the continuous development and advancement of mechanisms towards protecting and safeguarding the environment. Recycling and remanufacturing has therefore become a key tool by which many enterprises enhance their core competitiveness [1–3].

Economic and environmental sustainability contributes to the reduction in production cost and enhances the profits of enterprises; hence more enterprises are embracing the remanufactured product supply chain [4, 5]. Sustainable recycling of resources has increasingly become the focus of attention by governments, enterprises, and society. At present, governments around the world through the legislative process are enacting policies to protect the ecological environment and safeguard the natural resources. New measures for energy conservation and emission reduction are constantly being rolled out by governments and key stakeholders. One of the key measures for energy conservation and emission reduction is remanufacturing. Remanufacturing

has consequently been acknowledged for the superior benefits it offers. In comparison with other approaches, remanufacturing is more economically and environmentally sustainable; it reduces the cost of production for enterprises, increases the profits of enterprises, and offers several opportunities for more companies to join product remanufacturing [4].

In recent years, the competition strategy and power structure of the remanufacturing supply chain have further attracted widespread attention from relevant scholars. For instance, Majumder and Groenevelt studied the manufacturer's pricing and remanufacturing strategy facing remanufacturer competition [6]. Bulmus et al.'s research focuses not only on the competition between the original manufacturer and the remanufacturer on the product side, but also on the competition between the two in the product recycling market [7]. Adem et al. took into full consideration the impact of competition in quality and quantity between the original manufacturer and the remanufacturer on profits, consumer surplus, and total social welfare [8]. Gao Juhong et al. comprehensively studied the influence of different market power structures between manufacturers and retailers on the

optimal decision-making and market performance of closed-loop supply chains [9]. Cao Xiaogang and others established the Stackelberg game, manufacturer and retailer Stackelberg competition game model led by manufacturers, obtained the equilibrium solutions under different decision modes, and analyzed parameters such as consumer preference level, remanufacturing cost saving, and decision and profit impact [10].

The above research has greatly enriched the theory of remanufacturing supply chain competition and optimization design. Under certain market conditions, the supply chain system can also be seen as a game between manufacturers and retailers. Therefore, the system dynamics method is an important research perspective. These researches enrich the competition theory and optimal design of the remanufacturing supply chain greatly. These researches augment the competition theory and optimal design of the remanufacturing supply chain greatly. Under certain market conditions, the supply chain system can be regarded as a game between manufacturers and retailers, so system dynamics is an important perspective of research. Vlachos et al. gave a simulation model based on the principles of the system dynamics methodology to evaluate alternative long-term capacity planning policies using total supply chain profit as a measure of policy effectiveness [11]. Hammond and Beullens expanded the work dealing with oligopolistic supply chains to the field of supply chain. They elucidated a supply chain network model consisting of manufacturers and consumer markets with perfect information [12]. They pointed out that supply chain equilibrium occurred when all players agreed on volumes shipped and prices charged. Chen and Chang dealt with the strategic issue of supply chain with remanufacturing by developing analytic models under cooperative and competitive settings to investigate the conditions under which an original equipment manufacturer may take a cooperative approach by participating in remanufacturing [13]. Their analyses revealed that the strategic decisions depended critically on the costs of remanufacturing and the competition intensity. Guo and Ma employed the numerical simulation, the complex dynamic phenomena, such as bifurcations, chaos, and continuous power spectrum, in analyzing the complexity of retailer collecting and selling in closed-loop supply chain [14]. Wang Wenbin et al. built a variational inequality optimization model to analyze the impact of two strategies (government penalty for producers and government subsidy for recyclers) on the closed-loop supply chain. Their findings revealed that government subsidies for recyclers can improve the total recovery as well as increase the demands for new and old materials [15]. Jin Changfei et al. built a Stackelberg model of game between producers and retailers to analyze the sales volumes of green products and ordinary products in the following circumstances, respectively: producers not participating in recycling and remanufacturing; producers participating in recycling and remanufacturing; the government adopting an incentive strategy based on total recovery subsidy; and the government adopting an incentive strategy based on recovery rate subsidy [16]. Chang Xiangyun et al. introduced a system dynamic model in analyzing the impact of carbon tax constraint

and subsidy incentive on the auto parts remanufacturing decision-making and proposed that the combination of tax policy and subsidy incentive is more conducive to higher carbon reduction efficiency and higher economic efficiency of auto parts remanufacturing [17]. Several other researchers such as Guide and Van Wassenhove [18], Hatcher et al. [19], and Souza [20] have extensively reviewed literature covering all aspects of remanufacturing. Regarding research on remanufacturing and subsidies, Tukker et al. [21] and Vercaere et al. [22] studied the impact of demand uncertainty on government subsidies for green technology adoption and appraised trade-in remanufacturing with strategic customer behavior and government subsidies [23–25], etc.

From the extant literature, it is lucidly clear that most researchers that studied remanufacturing supply chains system employed models that consider only one aspect of the remanufacturing supply chain. There is therefore the urgent need to explore newer models to discuss the impact of different pricing strategies on the stability of supply chain systems. This current paper therefore extends the frontlines of recent research as it proposes a noncooperative game model approach in which manufacturers manufacture new products and remanufactured products at the same time. Through numerical simulation, the dynamics of the supply chain and the influence of decision parameters on the complex nonlinear dynamic behavior are analyzed.

This rest of the paper is organized as follows. In Section 2, the assumptions and notations are described. In Section 3, we present the Stackelberg game model while using numerical simulations to analyze its dynamic phenomena. In Section 4, the time-delayed feedback control is used to stabilize the chaotic behaviors of the system. Finally, conclusions are drawn in Section 5.

## 2. Model Assumptions and Notations

*2.1. Assumptions.* The following assumptions are made to develop the model for the supply chain system.

(i) The supply chain consists of a manufacturer and a retailer as shown in Figure 1. In this case, the manufacturer engages in the production of new products and remanufactured products. The retailer sells new products and remanufactured products to consumers. The supply chain system is an integration of the forward supply chain. The manufacturer and the retailer are independent decision-makers, and their goal is to maximize returns in discrete time period as shown in Figure 1.

(ii) The demand for new products in the market is greater than the market supply for new products, so remanufactured products occupy a certain market share.

(iii) There is no difference between remanufactured products and new products which are made of raw materials, but the cost of remanufactured products and the accessories in the production process are lesser than those of new products, so they are listed in the market, forming a core competitiveness advantage.

(iv) Consumers have personal preference for purchasing new products or remanufacturing products.

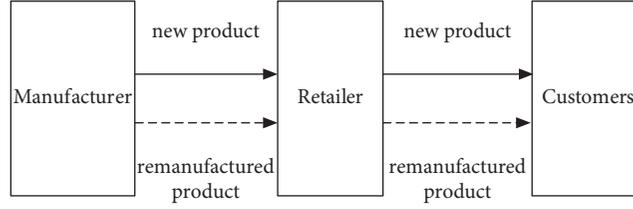


FIGURE 1: The supply chain system.

**2.2. Notations.** We considered the demand for products in the market as a linear function of product price. The demand function of new products manufactured with raw materials is computed as  $D_n = a_1 - b_1 p_{Rn}$ ; the demand function of remanufactured products is computed as  $D_r = a_2 - b_2 p_{Rr}$ , where  $a_i > 0 (i = 1, 2)$  is the market scale and  $b_i > 0 (i = 1, 2)$  is the price sensitivity of consumers. The parameter  $b_i$  is determined by the property of the products, such as similar degree. In general, the higher (lower) the degree of substitutability, the higher (lower) the value of  $b_i$ . Similarly the importance of goods determines the value of  $b_i$ . The  $b_i$  is lower for basic necessities of life, while the value of  $b_i$  is higher for luxury goods that are nonessential; and as for the adjustment time of consumers to demand, in general, the shorter (longer) the adjustment time, the lower (higher) the value of  $b_i$ .

The wholesale price of the manufacturer, wholesaling to the retailer, is defined as the cost of the remanufactured products and the new products manufactured with raw materials. This is computed mathematically as follows:  $p_n(t) = p_r(t)(1 + k_n)$ .

The retailer's unit price for selling the new products and the remanufactured products are estimated, respectively, as follows:  $p_{Rn}(t) = p_n(t)(1 + k_1)$  and  $p_{Rr}(t) = p_r(t)(1 + k_2)$ .

Hence, the profit function of the manufacturer is computed as follows:

$$\begin{aligned} \pi_m(t) &= \pi_{mn}(t) + \pi_{mr}(t) \\ &= [p_n(t) - c_n] D_n + [p_r(t) - c_r] D_r. \end{aligned} \quad (1)$$

The profit function of the retailer is as follows:

$$\begin{aligned} \pi_R(t) &= \pi_{Rn}(t) + \pi_{Rr}(t) \\ &= [p_{Rn}(t) - p_n(t)] D_r + [p_{Rr}(t) - p_r(t)] D_r. \end{aligned} \quad (2)$$

Consequently, the overall profit of the supply chain is computed as follows:

$$\pi(t) = \pi_m(t) + \pi_R(t). \quad (3)$$

The manufacturer chooses the remanufactured products price  $p_r(t)$  and profitability  $k_n$  as decision variables, while the retailer chooses selling the new products at a profit  $k_1$  and the remanufactured products at a profit  $k_2$ . The three profit functions are expressed as follows:

$$\begin{aligned} \pi_m(t) &= ((1 + k_n) p_r(t) - c_n) (a_1 \\ &\quad - b_1 (1 + k_1) (1 + k_n) p_r(t)) + (p_r(t) - c_r) (a_2 \\ &\quad - b_2 p_r(t) (1 + k_2)) \end{aligned}$$

$$\pi_R(t) = p_r(t)$$

$$\begin{aligned} &\cdot ((1 + k_1) (1 + k_n) (a_1 - b_1 (1 + k_1) (1 + k_n) p_r(t)) \\ &\quad - k_2 (b_2 (1 + k_2) p_r(t) - a_2)) \end{aligned}$$

$$\pi(t) = p_r(t)$$

$$\begin{aligned} &\cdot (b_1 (1 + k_1) (1 + k_n) (c_n + k_1 (1 + k_n) p_r(t)) \\ &\quad + b_2 (1 + k_2) (c_r + (k_2 - 1) p_r(t))) - a_2 (c_r \\ &\quad + (k_2 - 1) p_r(t)) - a_1 (c_n + k_1 (1 + k_n) p_r(t)). \end{aligned} \quad (4)$$

### 3. Stackelberg Game Model

Suppose that, in the supply chain, the manufacturer and the retailer are all in pursuit of maximum profits and make decisions independently as independent subjects where the manufacturer is the leader and the retailer is the follower. Then this relationship implies the dominance of the manufacturer over the retailer. They form a Stackelberg game, and the game equilibrium is referred to as the Stackelberg Equilibrium [26]. In this game, the manufacturer makes decisions for the determination of the wholesale price of the remanufactured products and profitability according to the market information. The retailer, on the other hand, makes decisions of the profitability of the new products and the remanufactured products according to the decision-making of the manufacturer.

**3.1. Model and Analysis.** The retailer's selling of the new products and the remanufactured products profitability are decided by solving the following optimization problem:

$$(k_1, k_2) = \arg \max \pi_R. \quad (5)$$

This can be adduced by the first-order conditions of  $\pi_R$ :

$$\begin{aligned} \frac{\partial \pi_R(t)}{\partial k_1} &= (1 + k_n) p_r(t) (a_1 - 2b_1 (1 + k_1) (1 + k_n) p_r(t)) \end{aligned} \quad (6)$$

$$\frac{\partial \pi_R(t)}{\partial k_2} = p_r(t) (a_2 - b_2 (1 + 2k_2) p_r(t)).$$

The retailer's best reply functions is derived by solving (6) as follows:

$$\begin{aligned} k_1 &= \frac{a_1 - 2b_1 p_r(t) - 2b_1 k_n p_r(t)}{2b_1(1+k_n)p_r(t)} \\ k_2 &= \frac{a_2 - b_2 p_r(t)}{2b_2 p_r(t)}. \end{aligned} \quad (7)$$

Similarly, we can get the manufacturer's marginal profit as follows:

$$\begin{aligned} \frac{\partial \pi_m(t)}{\partial p_r(t)} &= a_2 + b_2(1+k_2)(c_r - p_r(t)) - b_2(1+k_2)p_r(t) \\ &\quad - b_1(1+k_1)(1+k_n)((1+k_n)p_r(t) - c_n) \\ &\quad + (1+k_n)(a_1 - b_1(1+k_1)(1+k_n)p_r(t)) \\ \frac{\partial \pi_m(t)}{\partial k_n} &= p_r(t)(a_1 + b_1(1+k_1)(c_n - 2(1+k_n)p_r(t))). \end{aligned} \quad (8)$$

Considering the fact that agents of the manufacturer can never get the whole market information, their decisions are considered not completely rational. The managers always hope to gain more profit through the active managerial behavior. This paper will therefore adjust the remanufactured products price  $p_r(t+1)$  and the profitability  $k_n(t+1)$  on the bases of the marginal profits. Accordingly, the dynamic supply chain system is constructed as follows:

$$\begin{aligned} k_1(t) &= \frac{a_1 - 2b_1 p_r(t) - 2b_1 k_n(t) p_r(t)}{2b_1(1+k_n(t))p_r(t)} \\ k_2(t) &= \frac{a_2 - b_2 p_r(t)}{2b_2 p_r(t)} \\ p_r(t+1) &= p_r(t) + \alpha_1 p_r(t)(a_2 + b_2(1+k_2(t)) \\ &\quad \cdot (c_r - p_r(t)) - b_2(1+k_2(t))p_r(t) - b_1(1+k_1(t)) \\ &\quad \cdot (1+k_n(t))((1+k_n(t))p_r(t) - c_n) + (1+k_n(t)) \\ &\quad \cdot (a_1 - b_1(1+k_1(t))(1+k_n(t))p_r(t))) \\ k_n(t+1) &= k_n(t) + \alpha_2 k_n(t)(p_r(t)(a_1 + b_1(1+k_1(t)) \\ &\quad \cdot (c_n - 2(1+k_n(t))p_r(t))), \end{aligned} \quad (9)$$

where  $\alpha_i > 0 (i = 1, 2)$  are the profit modification speed parameter, which reflect the company active managerial behavior.

**3.2. Numerical Simulation.** Our study sets the following parameters:  $a_1 = 15$ ;  $a_2 = 10$ ;  $b_1 = 1.2$ ;  $b_2 = 1$ ;  $c_n = 4$ ;  $c_r = 3$  and the initial values as  $p_r(0) = 5$ ;  $k_n(0) = 0.4$  in the system as shown in (9).

The influence of decision parameters to the supply chain system is shown in (9). Figures 2 and 3, respectively, show the

bifurcation diagrams and the largest Lyapunov exponent plot of the supply chain system with  $\alpha_2 = 0.1$  and  $\alpha_1$  varying from 0 to 0.7. We can conclude that the supply chain system is stable when  $\alpha_1 < 0.435$ . When  $\alpha_1 = 0.435$ , the supply chain system (9) causes the first bifurcation,  $\alpha_1 < 0.538$  causes the second bifurcation, and the equilibrium of the supply chain system can lose stability via flip bifurcation and gradually enter the Neimark-Sacker bifurcation and chaotic state. The influences of  $\alpha_1$  on profits of the manufacturer, retailer, and supply chain are shown in Figures 2(h), 2(i), and 2(j). In stable state, the profits of the manufacturer, the retailer, and supply chain are, respectively, 38.5, 30.02, and 68.52. The retailer's unit selling price for the new products and the remanufactured products price are shown in Figures 2(f) and 2(g). The variation of largest Lyapunov exponent is closely related to the change of the price adjustment rate of remanufactured products ( $\alpha_1$ ), which can be seen in Figure 3. This figure proves that when the price adjustment rate of remanufactured products ( $\alpha_1$ ) changes, the stability of the supply chain system becomes unstable, and the characteristics of mixed dynamics after bifurcation evolve.

Figure 4 denotes the cycle and chaotic diagram of the supply chain system with the adjustment speeds  $\alpha_1, \alpha_2$  at the same time change. The equivalent period graph is shown in Figure 4(a). In the figure, from the lower-left corner, the district of different color depth converges once in 1 period, 2 periods, 4 periods, and 8 periods, which is made of parameters  $(\alpha_1, \alpha_2)$ . The economic meaning of  $p$  period is that the parameter randomly chooses values in the point set, as the manufacturer chooses one initial remanufactured products price and profitability; after limited time games, the remanufactured products prices represent the periodicity changing of  $p$ , which is  $M_i^p(p_r^{t_0}, k_n^{t_0}) = p_i^{t_0}, M_i^k(p_r^{t_0}, k_n^{t_0}) \neq p_i^{t_0}, i = 1, 2$ . All random integers are set up for  $k < p$ .  $t_0$  represents arbitrary integer greater than random natural number;  $M_i$  represents mapping in the system (9). The green district at the top right corner through iteration makes parameters on the value of enterprise production strategy become negative, which is also referred to as nonfeasible parameter set, from the path of parameter set. It describes the collapse of quality cost and profit system evolution. Based on Sarkovskii periodic orbit theorem, period 1, period 2, and period 4 all account for approximately 84.6%. It can be inferred that the appeal of low cycle track has high possibility in the quality cost and profit dynamic evolution system.

In chaotic map, Figure 4(b), the dark area represents the chaotic state in the model of largest Lyapunov index meaning. It does not consider the fact that the production has negative value situation. It revealed that 15.6% is chaotic. This implies that the periodic phenomena are more common, also suitable for the realistic economic behavior. When  $\alpha_1, \alpha_2$  increase to the certain degree, it has chaotic phenomenon. Thus, when enterprise makes production decision, the speed of production adjustment becomes faster; the market responding speed becomes faster, making the market become unstable. It has many different paths from period state to chaotic state as shown in Figure 4. Figure 4 fully explains the complexity of system evolution and the chaotic phenomenon.

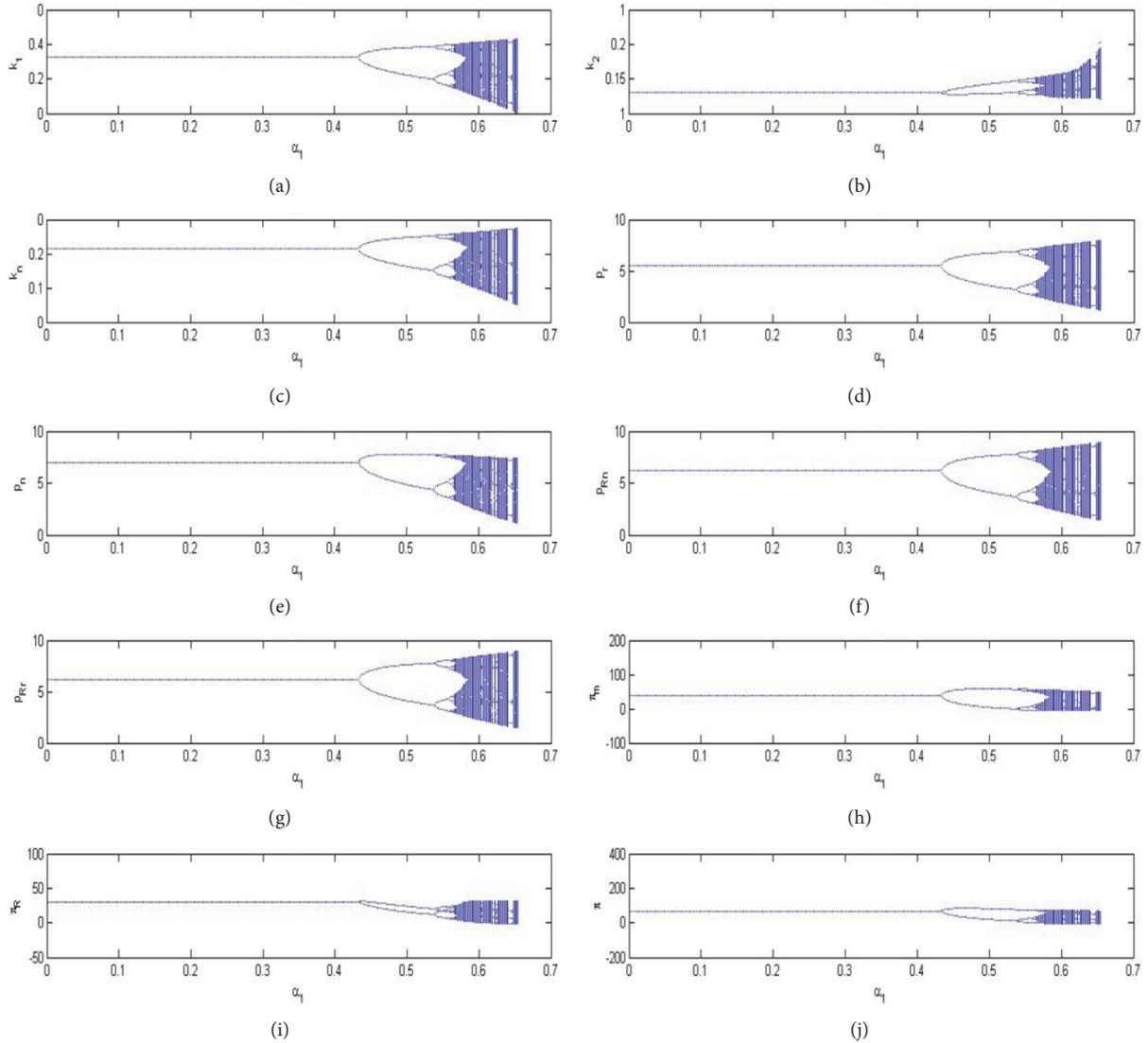


FIGURE 2: Bifurcation diagrams of the supply chain system (9) with the adjustment speed  $\alpha_1$ .

The influence of initial state on the supply chain system is shown in (9).

Figure 5 reveals the sensitive dependence for the dynamic system on initial values. The sensitivity of system (when loosing stability) to the manufacturer's initial remanufactured products price  $p_r$  is taken as 5 and 5.0001, and small changes of the initial conditions can cause the observed large changes of the system, which is sensitive to initial states.

#### 4. Chaos Control

From the numerical simulations, the adjustment rate and the weight coefficient had great influence on the stability of system (9). If the model parameters fail to locate into the stable region required, the behaviors of the dynamics will be much complicated. In a real economic system, chaos is undesirable and unwarranted. There is therefore the need to avoid or control chaos so that the dynamic system would

work favorably. In this section, our study introduced the time-delayed feedback control [27, 28] to control system chaos. We modified the first equation of system (9) by intercalating a controller  $K_1(p_r(t) - p_r(t + 1))$  and  $K_2(k_n(t) - k_n(t + 1))$  as a small perturbation, where  $K_i > 0 (i = 1, 2)$  is a controlling coefficient. The controlled system is computed in (10) as follows:

$$k_1(t) = \frac{a_1 - 2b_1 p_r(t) - 2b_1 k_n(t) p_r(t)}{2b_1(1 + k_n(t)) p_r(t)}$$

$$k_2(t) = \frac{a_2 - b_2 p_r(t)}{2b_2 p_r(t)}$$

$$p_r(t + 1) = p_r(t) + \alpha p_r(t) (a_2 + b_2(1 + k_2(t)) \cdot (c_r - p_r(t)) - b_2(1 + k_2(t)) p_r(t) - b_1(1 + k_1(t))(1 + k_n(t))((1 + k_n(t)) p_r(t) - c_n))$$

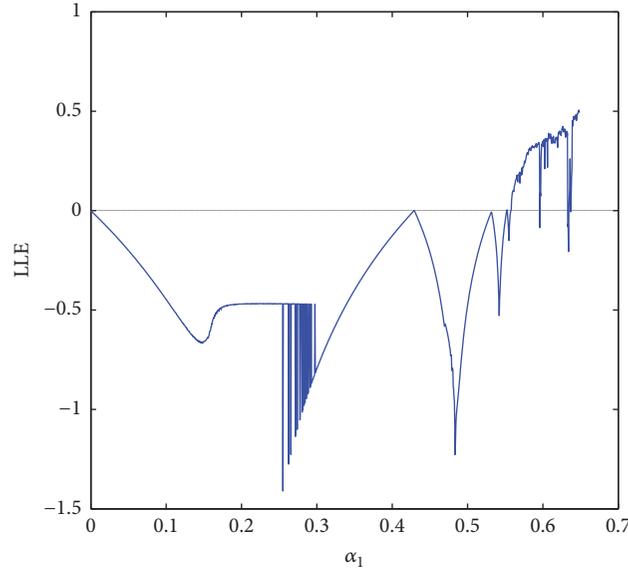


FIGURE 3: The largest Lyapunov exponent (LLE) plot of the supply chain system (10) with the adjustment speed  $\alpha_1$ .

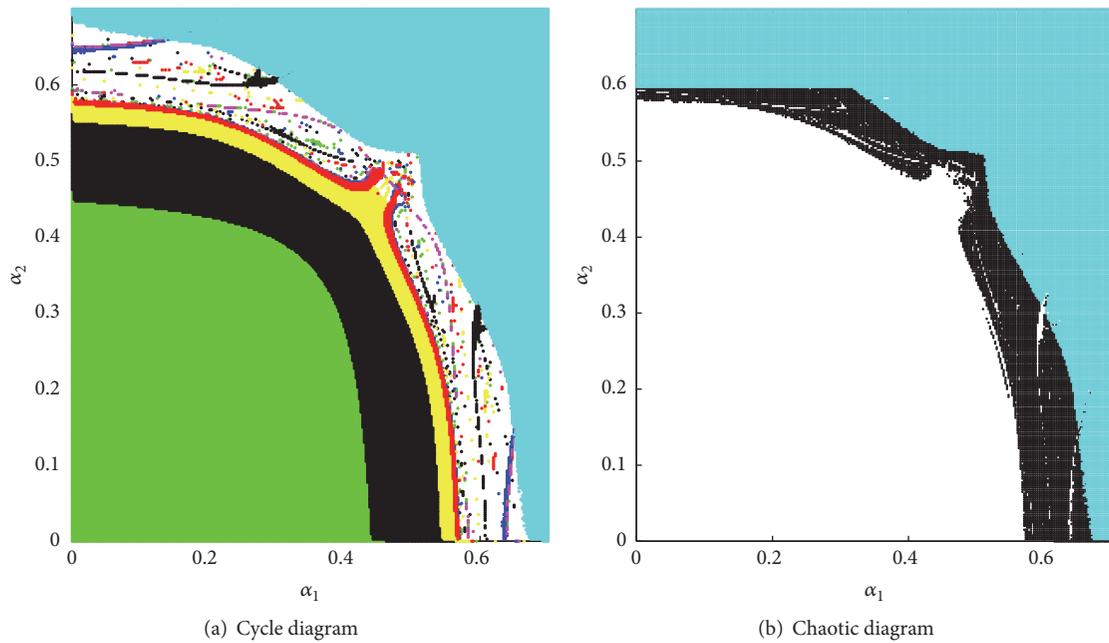


FIGURE 4: Cycle diagram and chaotic diagram of the system with the adjustment speeds  $\alpha_1, \alpha_2$  at the same time.

$$\begin{aligned}
 &+ (1 + k_n(t)) \\
 &\cdot (a_1 - b_1 (1 + k_1(t)) (1 + k_n(t)) p_r(t)) \\
 &+ K_1 (p_r(t) - p_r(t+1)) \\
 k_n(t+1) = &k_n(t) + \beta k_n(t) (p_r(t) (a_1 + b_1 (1 + k_1(t)) \\
 &\cdot (c_n - 2 (1 + k_n(t)) p_r(t))) + K_2 (k_n(t) \\
 &- k_n(t+1)).
 \end{aligned} \tag{10}$$

It is easy to see that the new system (10) has the same equilibriums as system (9) and it takes the following equivalent form:

$$k_1(t) = \frac{a_1 - 2b_1 p_r(t) - 2b_1 k_n(t) p_r(t)}{2b_1 (1 + k_n(t)) p_r(t)}$$

$$k_2(t) = \frac{a_2 - b_2 p_r(t)}{2b_2 p_r(t)}$$

$$p_r(t+1) = p_r(t) + \frac{\alpha_1 p_r(t)}{1 + K_1} (a_2 + b_2 (1 + k_2(t)))$$

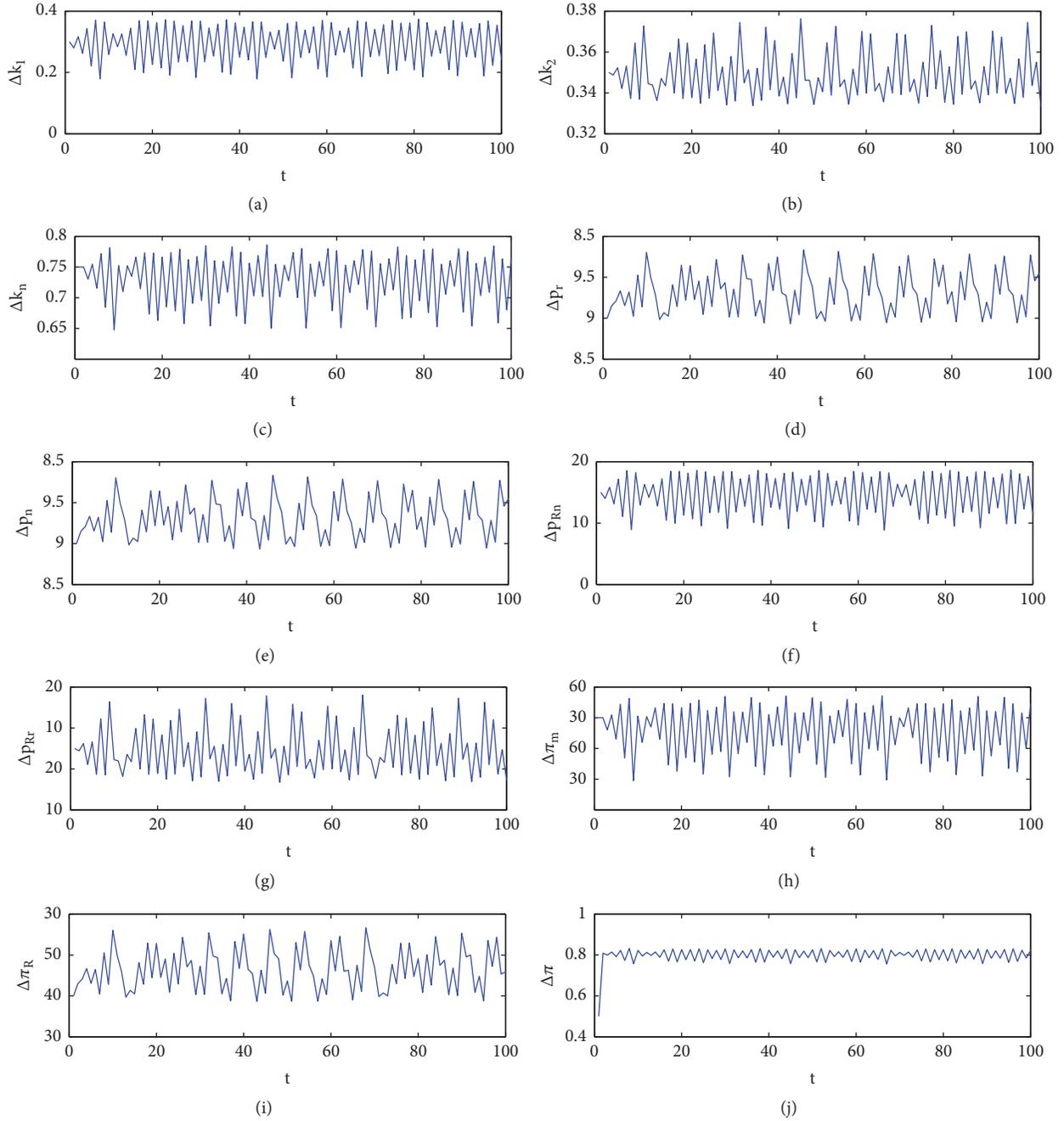


FIGURE 5: Sensitive dependence for the dynamical system (11) on initial conditions.

$$\begin{aligned}
 & \cdot (c_r - p_r(t)) - b_2(1 + k_2(t)) p_r(t) \\
 & - b_1(1 + k_1(t))(1 + k_n(t))((1 + k_n(t)) p_r(t) - c_n) \\
 & + (1 + k_n(t)) \\
 & \cdot (a_1 - b_1(1 + k_1(t))(1 + k_n(t)) p_r(t)) \\
 k_n(t + 1) = & k_n(t) + \frac{\alpha_2 k_n(t)}{1 + K_2} (p_r(t) (a_1 \\
 & + b_1(1 + k_1(t))(c_n - 2(1 + k_n(t)) p_r(t))).
 \end{aligned}$$

(11)

In Figure 6, we set some conditions. For instance when  $K_2 = 0.3$ , it is obviously observed that with the control coefficient  $K_1$  increasing, the supply chain system gradually gets out of chaos and periodic windows and achieves stability when  $K_1 > 0.0817$ . The variation of largest Lyapunov exponent is closely related to the change of the control coefficient ( $K_1$ ), as shown in Figure 7. This figure further confirmed that when the control coefficient changes, the instability of the supply chain system becomes stable, and the characteristics of stable evolution after the double-cycled bifurcation evolve. We can also discuss the system

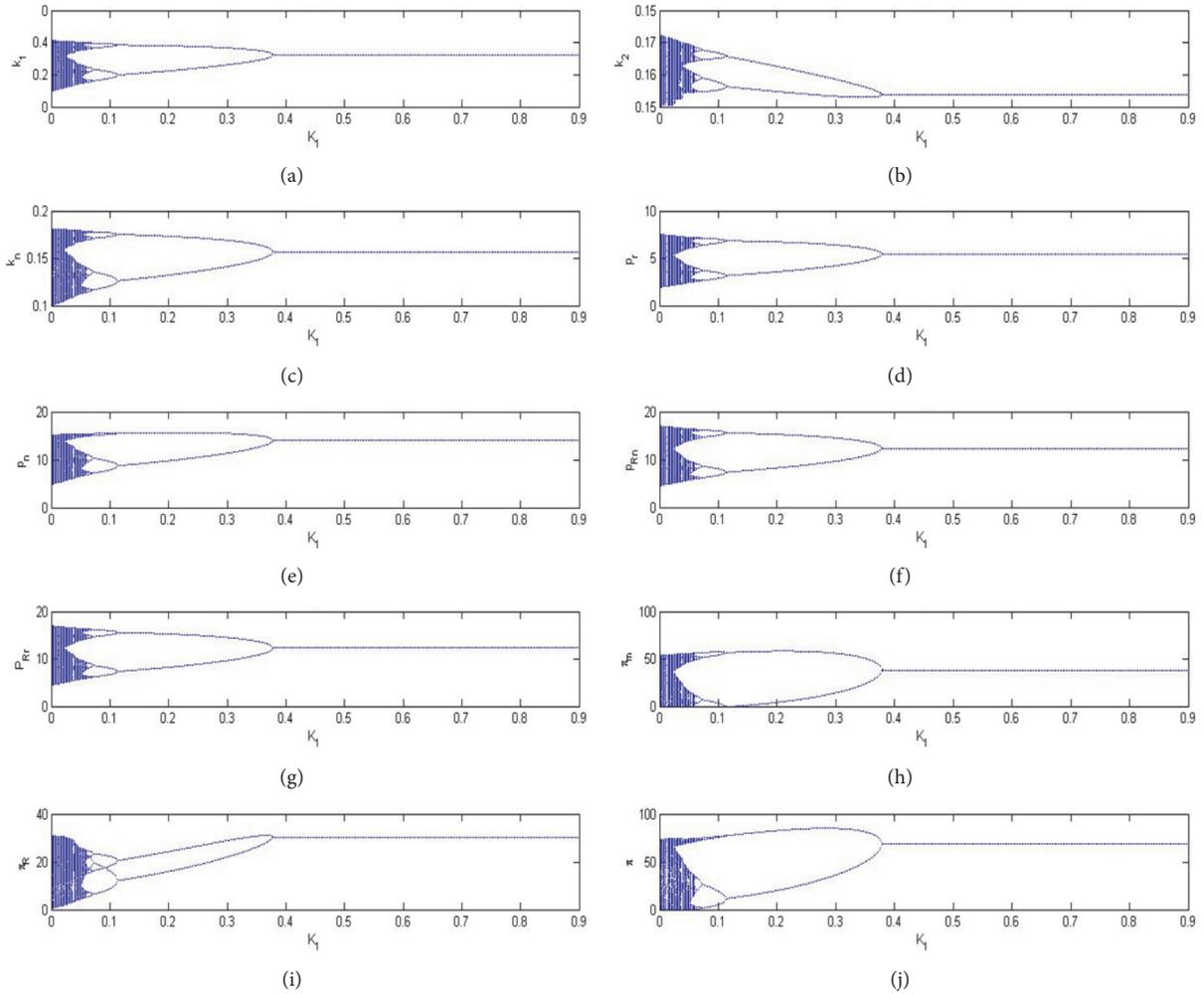


FIGURE 6: Bifurcation diagrams and the maximum Lyapunov exponent of system (11) with the controlling factor  $K_1$ .

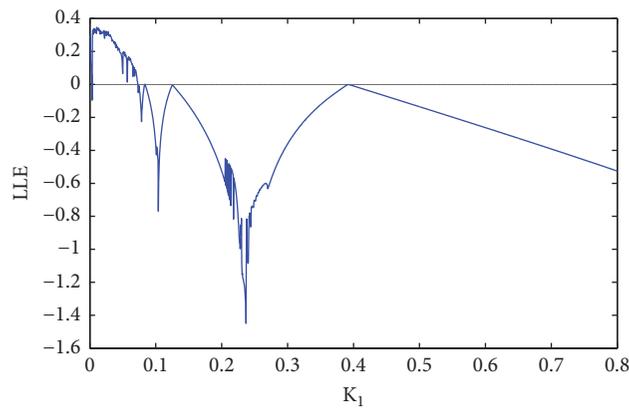


FIGURE 7: The largest Lyapunov exponent (LLE) plot of the supply chain system (10) with the controlling factor  $K_1$ .

stability of the controlled system from the view of the stability region in the two-dimension  $K_1, K_2$ -plane. Thus, the controlled system will be locally asymptotically stable provided that  $K_1$  and  $K_2$  take their values in the plotted region. Figure 8(a) exhibits the cycle diagram whereas

Figure 8(b) depicts the chaotic diagram. By comparing Figures 4(a) and 4(b), we observed that the stability of the supply chain system is significantly enhanced after the implementation of the control system. We can therefore vehemently avow that the supply chain system chaos

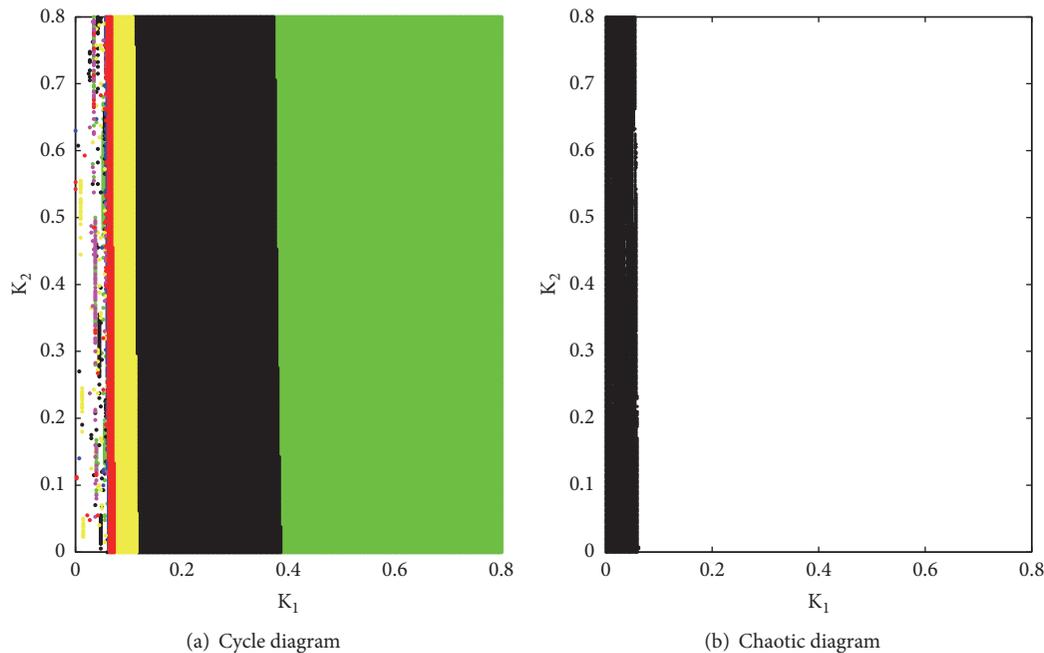


FIGURE 8: Cycle diagram and chaotic diagram of the system with the adjustment speeds  $K_1$ ,  $K_2$  at the same time.

is effectively controlled after employing feedback control approach.

## 5. Conclusion

This study investigated the supply chain in which manufacturers manufacture new products and remanufacture products at the same time using a noncooperative game model approach. The forward supply chain of Stackelberg game model was augmented and developed further for the analysis. The complexity of the models was investigated by simulation. Numerical simulations and the Maximum Lyapunov exponent were used to provide experimental evidence for the complicated evolution behaviors of the supply chain system. Findings of our study revealed that parameters play important role in the stability of the supply chain systems based on analyses of dynamical behaviors of the established game models. Moreover, the time-delayed feedback is a crucial control mechanism suitable for stabilizing the chaotic behaviors of the dynamic systems. The time-delayed feedback demonstrated that the chaotic behavior of the system can be controlled.

The study provided in-depth knowledge on supply chain mechanism within the remanufactured products—composed of one manufacturer and one retailer. This study will therefore serve as a useful guide for policy-makers, supply chain managers, and decision-makers for the implementation of strategic decisions.

## Data Availability

The numerical simulations data used to support the findings of this study were supplied by Liuwei Zhao under license and

so cannot be made freely available. Requests for access to these data should be made to Liuwei Zhao, E-mail address: 136901672@qq.com.

## Conflicts of Interest

The publication of this paper has no conflicts of interest.

## Acknowledgments

This work is supported by the National Nature Science Foundation of China (nos. 71171099, 71471076, 71001028, 71201071, and 71373103) and China Scholarship Council under Grant 20123227110011.

## References

- [1] R. C. Savaskan, S. Bhattacharya, and L. N. Van Wassenhove, "Closed-loop supply chain models with product remanufacturing," *Management Science*, vol. 50, no. 2, pp. 239–252, 2004.
- [2] J. D. Abbey, J. D. Blackburn, and V. D. R. Guide, "Optimal pricing for new and remanufactured products," *Journal of Operations Management*, vol. 36, pp. 130–146, 2015.
- [3] J. D. Abbey, M. G. Meloy, V. D. R. Guide, and S. Atalay, "Remanufactured products in closed-loop supply chains for consumer goods," *Production Engineering Research and Development*, vol. 24, no. 3, pp. 488–503, 2015.
- [4] M. P. De Brito and R. Dekker, "A Framework for Reverse Logistics," in *Reverse Logistics: Quantitative Models for Closed-loop Supply Chains*, R. Dekker, M. Fleischmann, K. Inderfurth, and L. N. Van Wassenhove, Eds., Springer, 2004.
- [5] J. Gaur, M. Amini, P. Banerjee, and R. Gupta, "Drivers of consumer purchase intentions for remanufactured products a study of Indian consumers relocated to the USA," *Qualitative*

- Market Research: An International Journal*, vol. 18, no. 1, pp. 30–47, 2015.
- [6] P. Majumder and H. Groenevelt, “Competition in remanufacturing,” *Production Engineering Research and Development*, vol. 10, no. 2, pp. 125–141, 2001.
- [7] S. C. Bulmus, S. X. Zhu, and R. Teunter, “Competition for cores in remanufacturing,” *European Journal of Operational Research*, vol. 233, no. 1, pp. 105–113, 2014.
- [8] A. Örsdemir, E. Kemahlioğlu-Ziya, and A. K. Parlaktürk, “Competitive quality choice and remanufacturing,” *Production & Operations Management*, vol. 23, no. 1, pp. 48–64, 2014.
- [9] G. Juhong, H. Shuai, H. Liting et al., “Pricing and effort decisions in a closed-loop supply chain under different channel power structures,” *Journal of Cleaner Production*, vol. 112, pp. 2043–2057, 2015.
- [10] C. Xiaogang, Z. Benrong, and W. Hui, “CCustomer preference based decisions on price fixing and coordination in a dual-channel closed-loop supply chain,” *Chinese Journal of Management Science*, vol. 23, no. 6, pp. 107–117, 2015.
- [11] D. Vlachos, P. Georgiadis, and E. Iakovou, “A system dynamics model for dynamic capacity planning of remanufacturing in closed-loop supply chains,” *Computers & Operations Research*, vol. 34, no. 2, pp. 367–394, 2007.
- [12] D. Hammond and P. Beullens, “Closed-loop supply chain network equilibrium under legislation,” *European Journal of Operational Research*, vol. 183, no. 2, pp. 895–908, 2007.
- [13] J.-M. Chen and C.-I. Chang, “The co-opetitive strategy of a closed-loop supply chain with remanufacturing,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 2, pp. 387–400, 2012.
- [14] Y. Guo and J. Ma, “Research on game model and complexity of retailer collecting and selling in closed-loop supply chain,” *Applied Mathematical Modelling: Simulation and Computation for Engineering and Environmental Systems*, vol. 37, no. 7, pp. 5047–5058, 2013.
- [15] W. Wenbin, D. Qingli et al., “Remanufacturing Closed-loop Supply Chain Network Equilibrium Model based on Allowance and Penalty,” *Operations Research and Management Science*, vol. 19, no. 1, pp. 65–72, 2010.
- [16] J. Changfei, C. Erbao, and L. Mingyong, “Analysis on green marketing strategy of duopoly retailing market based on the evolutionary theory,” *Journal of Systems Engineering*, vol. 27, no. 3, pp. 383–389, 2012.
- [17] C. Xiangyun, Z. Yongguang, W. Yixuan et al., “Research of low-carbon policy to promote automotive parts remanufacturing in china: a case study of auto engine remanufacturing,” *Systems Engineering – Theory & Practice*, vol. 33, no. 11, pp. 2811–2821, 2013.
- [18] V. D. R. Guide Jr. and L. N. van Wassenhove, “The evolution of closed-loop supply chain research,” *Operations Research*, vol. 57, no. 1, pp. 10–18, 2009.
- [19] G. D. Hatcher, W. L. Ijomah, and J. F. C. Windmill, “Design for remanufacture: A literature review and future research needs,” *Journal of Cleaner Production*, vol. 19, no. 17–18, pp. 2004–2014, 2011.
- [20] G. C. Souza, “Closed-loop supply chains: a critical review, and future research,” *Decision Sciences*, vol. 44, no. 1, pp. 7–38, 2013.
- [21] A. Tukker, M. J. Cohen, U. De Zoysa et al., “The Oslo declaration on sustainable consumption,” *Journal of Industrial Ecology*, vol. 10, no. 1–2, pp. 9–14, 2006.
- [22] S. Vercraene, J.-P. Gayon, and S. D. Flapper, “Coordination of manufacturing, remanufacturing and returns acceptance in hybrid manufacturing/remanufacturing systems,” *International Journal of Production Economics*, vol. 148, pp. 62–70, 2014.
- [23] Y. Wang, V. Wiegerinck, H. Krikke, and H. Zhang, “Understanding the purchase intention towards remanufactured product in closed-loop supply chains: An empirical study in China,” *International Journal of Physical Distribution and Logistics Management*, vol. 43, no. 10, pp. 866–888, 2013.
- [24] S. Wei, O. Tang, and E. Sundin, “Core (product) Acquisition Management for remanufacturing: a review,” *Journal of Remanufacturing*, vol. 5, no. 4, 2015, Open access.
- [25] X. Xia, K. Govindan, and Q. Zhu, “Analyzing internal barriers for automotive parts remanufacturers in China using grey-DEMATEL approach,” *Journal of Cleaner Production*, vol. 87, no. 1, pp. 811–825, 2015.
- [26] J. Ma and H. Wang, “Complexity analysis of dynamic non-cooperative game models for closed-loop supply chain with product recovery,” *Applied Mathematical Modelling: Simulation and Computation for Engineering and Environmental Systems*, vol. 38, no. 23, pp. 5562–5572, 2014.
- [27] L. Zhao, “Nonlinear complex dynamics of carbon emission reduction Cournot game with bounded rationality,” *Complexity*, vol. 2017, Article ID 8301630, 7 pages, 2017.
- [28] L. Zhao, “Analysis and control of the complex dynamics of a multimarket Cournot investment game with bounded rationality,” *Discrete Dynamics in Nature and Society*, vol. 2016, Article ID 7342405, 10 pages, 2016.