A Hybrid Estimation of Distribution Algorithm for Multi-Objective Multi-Sourcing Intermodal Transportation Network Design Problem Considering Carbon Emissions

Shou-feng Ji and Rong-juan Luo *

School of Business Administration, Northeastern University, Shenyang 110167, China; sfji@mail.neu.edu.cn
* Correspondence: rongjuanluo@163.com; Tel.: +86-024-8368-0613

Abstract: The increasing concern on global warming is prompting transportation sector to take into account more sustainable operation strategies. Among them, intermodal transportation (IT) has already been regarded as one of the most effective measures on carbon reductions. This paper focuses on the model and algorithm for a certain kind of IT, namely multi-objective multi-sourcing intermodal transportation network design problem (MO_MITNDP), in which carbon emission factors are specially considered. The MO_MITNDP is concerned with determining optimal transportation routes and modes for a series of freight provided by multiple sourcing places to find good balance between the total costs and time efficiencies. First, we establish a multi-objective integer programming model to formulate the MO_MITNDP with total cost (TTC) and maximum flow time (MFT) criteria. Specifically, carbon emission costs distinguished by the different transportation mode and route are included in the cost function. Second, to solve the MO_MITNDP, a hybrid estimation of distribution algorithm (HEDA) combined with a heterogeneous marginal distribution and a multi-objective local search is proposed, in which the from the Pareto dominance scenario. Finally, based on randomly generated data and a real-life case study of Jilin Petrochemical Company (JPC), China, simulation experiments and comparisons are carried out to demonstrate the effectiveness and application value of the proposed HEDA.

Keywords: multi-objective optimization; intermodal transportation network design; hybrid estimation of distribution algorithm; carbon reduction; Pareto dominance

1. Introduction

As one of the most important logistics activities, freight transportation which happens throughout the whole production circulation process from material purchase to finished product sale is crucial to the economic growth and market trade. With the rapid growth of global purchasing, transportation has played an increasingly important role in world resource integrations. On the other hand, due to the downsides such as energy consuming and air pollution, transportation is always located at the crossroads of economic and environmental concerns. For example, as pointed by Environmental Protection Agency (USA), transportation is the second largest contributor of carbon emission, causing about 28% of the total emissions in US in the past years [1]. Accumulation of greenhouse gas in the atmosphere, especially CO_2, has been a main source of global warming and attracts increasing concerns by different sectors [2]. For example, the European Commission announced that the major objective corresponds to transportation sector is to reduce the carbon emissions at least 60% of 1990 levels by 2050 [3]. In the past years, carbon emissions from other transportation modes have experienced a relatively slight decreasing while the road transportation is keeping a rapid growth. Road

transportation is still the dominant mode in inland freight transportation system, and meanwhile it is also a carbon intensive mode which contributes two-thirds of total carbon emissions in transportation sector [4]. With the growth of transportation volume and economic, Zheng and Wang 2010 has forecasted that carbon emissions from the road freight transportation will keep a continual increase by 2020 globally [5]. In China, the current energy structure is dominated by coal (about 64%) and oil (about 18%) [2], which also has been seen as the hardest obstacle to the low-carbon transportation. In view of these, different measures and regulations have been established, such as developing new energy technologies, prompting green transportation equipment, issuing relevant laws, etc. According to Benjaafar et al., 2013 [6], mode selection and route design are of great importance for low-carbon transportation systems, which might be more significant than policies, regulations, laws established by national governments at different levels or international organizations. In this regard, intermodal transportation (IT) has been put on the agenda by many countries and regions, as well as been promoted as a highly desirable measure with respect to improving the resource, environment and ecology sustainability.

Unlike other conventional transportation patterns, in which the transportation modes are independently operated without any coordination, IT aims at integrating more than one mode together such that it can provide more flexible "door to door" service according to the personalized requirements of customers [7]. The optimization of IT is essentially to provide a satisfied solution for the intermodal transportation company with reducing transportation cost, transportation time and carbon emissions simultaneously. For carbon emissions reduction, it might be explained that the heavier emission modes are replaced by the lighter ones [8,9]. In general, there are three kinds of modes in IT, i.e., waterway, railway and road, and the former two usually contribute less carbon emissions than the latter. However, intermodal transportation systems have an intractable drawback that it is more likely to result in a larger transportation distance than under a single transportation mode [10]. Since the costs and carbon emissions are mainly from the transportation process, the increased transportation distance (under intermodal transportation situation) may not actually benefit the optimization of the whole transportation systems. Moreover, with the development of the diversified market, the transportation companies are faced with new challenges to meet different preferences as some customers expect a lower cost while some others focus on a fast delivery. In view of the aforementioned observations, the operating decision of IT is in fact a very complex issue. The complex transportation system and resource of freight, the preference of the customers and environmental sustainability (low carbon emissions) ought to be considered to some extent.

This paper focuses on the multi-objective multi-sourcing intermodal transportation network design problem (MO_MITNDP) with the aim to find a good trade-off between the transportation cost (TC), maximum flow time (MFT), and carbon emissions (CE). Among them, TC, MFT and CE are associated with economic, efficiency, and societal factors respectively. Furthermore, we have taken in account two kinds of technique constraints, i.e., constraints of transportation modes and constraints of each node's capacity which are realistic existing to make our model closer to the real-life intermodal transportation situations. Besides, the constraint for each node's capacity is also a strategy to balance the transportation pressure among the stations, leading to a better sustainability of the transportation system.

To formulate the MO_MITNDP, several sourcing places and a single destination are considered. Indeed, this framework is inspired by the well-known complexity of supply chains in the increasing global purchasing and manufacturing environments [11], which has been widely used in many practical transportation systems, such as petrochemical, steel, pharmaceutical industries, etc. The freight transportation company has acted as a freight integrator to transport freights from the sourcing places to the purchasing destination. On the whole, the overall intermodal transportation network is divided into several stages that consist of stations, ports or logistic centers, which are particularly grounded on the geographical characteristics and real intermodal infrastructure construction, such as business types and service radiant scopes. Meanwhile, it is assumed that the freights are irreversibly transported

from the current stage to the following one without dismounting, which is also a typical assumption in the intermodal transportation problem (Steadieseifi et al., 2014) [12]. The crossing logistics are considered as the business handling capacity constraints for each node of a certain stage when different freights from the corresponding sourcing places are assembled in the same node in a period of time. These constraints are quite close to the real-life situations and significant for balancing the logistics pressures among the stations, as well as reducing the idle transportation resources. In other words, such constraints can disperse the transportation pressure in the station of high-density traffic to the others belong to the same stage, which will further alleviate the city's traffic congestion and air pollution. The aim of this paper is to determine both the route and mode selection for freight derived from different sourcing places. The main contributions of this paper are listed as follows.

Firstly, the MO_MSITNDP is formulated as a multi-objective integer programming model. As for the network structure, the entire intermodal transportation network is divided into several stages according to the geography and social factors. Meanwhile, the capacity constrain of each intermediate node is considered to make the solution reliable in practice and balance the logistics pressure.

Secondly, a modified estimation of distribution algorithm (EDA), namely hybrid EDA (HEDA) is proposed to address the mathematical model, which is the first time to apply EDA-based meta-heuristic for MO_MITNDP. Based on the features of MO_MITNDP, we specially design the improvement strategies to enhance the effectiveness and robustness of the proposed method in multi-objective handling capacity.

Thirdly, we perform a series of numerous experiments and comparisons based on randomly generated instances to evaluate the effectiveness of HEDA. To further demonstrate the potential application value of HEDA, a case study of Jilin Petrochemical Company (JPC), China, is carried out such that we can obtain the related management insights from the real-world perspectives.

The remainder of the paper is organized as follows. Section 2 reviews the relevant literature. In Section 3, the mathematical model of MO_MITNDP is described in detail. Following a detailed description of the proposed HEDA in Section 4, we conduct numerous experiments and comparisons with other algorithms in Section 5, and a case study analysis is performed as well. Finally, conclusions together with future works are presented in Section 6.

2. Literature Review

Over the past decades, the IT has attracted many attentions as a cost-effectiveness and environment sustainability alternatives. The existing works can be classified as strategy planning level (Meng and Wang, 2011 [13]; Limbourg and Jourquin, 2009 [14]), tactical planning level (Lam and Gu, 2016 [7]; Verma et al., 2012 [15]), and operational planning level (Francesco et al., 2013 [16]; Bandeira et al., 2009 [17]) according to the decision horizon of the planning problems, and more works can refers to SteadieSeifi et al., 2014 [12]. A majority of the works on ITs consider the cost or time as the single objective and discuss cost-saving or time-efficiency ability compared to the unimodal freight transportation (Bouchery and Fransoo, 2014 [3]). However, the published works regarding cost and time as the bi-objective optimization problem or multiple objectives are limited. Yang et al., 2011 [18] presented a multi-objective optimization model considering the total cost, transportation time and time variability to elevate the competitiveness of 36 candidate routes from China to Indian Ocean. Resat and Turkay, 2015 [19] established a bi-objective mixed-integer programming model with minimizing the total cost and time to determine the transportation modes in the designed intermodal network. Domuta et al., 2012 [20] proposed a modified bi-objective Martines' Algorithm to find the optimal route in the intermodal transportation network with the time window and the objective is to minimize the travel time and cost. Lam and Gu, 2016 [7] formulated the port hinterland intermodal transportation network design as a bi-objective optimization problem with minimizing the transit cost and time.

In view of the green logistics, carbon emissions have been an increasing concern in the transportation sector [7]. Nevertheless, few current studies consider carbon emissions in the field of

intermodal transportation network. In order to estimate the trade-offs between the transportation cost and carbon emission in the intermodal network and show the modification of the mode and route choice with the changing of the trade-offs, Namseok et al., 2009 [21] developed a decision-support method using the multi-objective optimization. Rudi et al., 2016 [22] established a multi-commodity intermodal network flow model considering the in-transit inventory to make route choices. The carbon emission factor is considered in the objective function which is to minimize the total truckload with the transportation cost, time and carbon emission criteria. Demir et al., 2015 [23] used the sample average approximation method to solve the stochastic intermodal network design problem. The method is applied to provide solutions with the different combination of weights corresponding to the direct transportation, carbon emission-related and time-related cost in the objective function. Considering the cost and carbon emission optimization situations, Bouchery and Fransoo, 2014 [3] presented an intermodal network optimization model to provide decisions for both the terminal location and allocation while Lam and Gu, 2016 [7] studied the carbon emissions in the green intermodal network as a constraint with variable levels.

In fact, comparing with single objective, multi-objective optimization can made trade-offs for conflict sub-objectives to satisfy practical fields, which has received wider applications in decision making (Zheng et al., 2015) [2]. In addition, the increasing concerns on the environmental sustainability have promoted it a hotspot to reduce the carbon emission in transportation sector (Lee et al., 2017) [24]. In that regard, The MO_MITNDP considers the cost and time as the objectives, which could help to provide flexible solutions for the decision makers to satisfy the diversified market.

The generation of optimal routes is a crucial part of the intermodal transportation network and efficient routes can achieve great resources saving and improve the service quality (Kılıç and Gök, 2014) [25]. According to Verma and Verter, 2010 [26], the main contribution of rail-truck intermodal transportation is the reliability embodied in on-time delivery. The authors make the first attempt to develop a framework of rail-truck intermodal transportation for hazardous goods, in which the route selection is driven by the lead time from diversified customers. Considering the cooperation of inland terminals simulated by a discrete event model, An et al., 2010 [27] designed a service network for the intermodal barge transportation to select the best route with the goal to saving cost. Similarly, Braekers et al., 2013 [28] presented a decision support model for intermodal barge transportation network which involves a major seaport and several hinterland ports to determine the optimal shipping routes aiming to improving the overall profits. Ishfaq and Sox, 2012 [29] integrated an operation model and a location-allocation model for intermodal hub. In their designed hub logistics network, arcs that represent different transportation modes are alternative to increase cost benefits. In this paper, the intermediate nodes in the network are the first time to be divided into several stages grounded on the geographical characteristic and real intermodal infrastructure construction.

The ITs faced by the freight integrator is generally a NP-hard problem in which the routes selection, modes combination, terminal location or container scheduling are involved (Li et al., 2015) [30] and huge computational efforts are needed. The meta-heuristic algorithms have an outstanding performance in the combinatorial optimization and have been successfully used in the intermodal freight transportation optimization. Kılıç and Gök, 2014 [25] formulated the problem as an initial route programming for urban transportation network, and a route generation algorithm is designed to find solutions with objectives reflecting the interests of stakeholders. To handle the inland container transportation problem considering hard time constraints, Sterzik and Kopfer, 2013 [31] established a comprehensive model which contains vehicle routing and empty repositioning and a Tabu Search heuristic is used to minimizing the total operation time. Nossack and Pesch, 2013 [32] addressed a truck scheduling problem with time window in intermodal container transportation by a two-stage heuristic. The first stage is route construction heuristic which determines a feasible initial solution for the problem; the second stage is route improvement heuristic which improves solution quality by a neighborhood operator. Genetic algorithm is used in Zhang et al., 2015 [33] to effectively save time and cost.

As a novel swarm intelligence algorithm based on probability model, EDA has been a research hotspot in the area of intelligent computing (Wang et al., 2012 [34]; Wang et al., 2016 [35]). Contrast with genetic algorithm, there is no mutation and crossover operators for EDA, whereas a competitive mechanism is adopted in EDA to estimate the probability model according to excellent individuals found in the search process. Thereafter, the new population can be generated through sampling the estimated probability model. Although EDA has been successfully used in various combinatorial optimization problems, there is no research works on EDA for ITs. Essentially, using EDA-based methods to solve MO_MSITNDP is equivalent to confirming the matching probability of the transportation modes and intermediate nodes associated with given freight. Finally, the optimal or approximate optimal relationship will be found. In this regard, this paper engages in raising the EDA-based solution method for dealing with MO_MSITNDP.

3. Problem Description and Model Formulation

3.1. Formulation of Proposed Mathematical Model

MO_MSITNDP considers multiple sourcing places and a single purchasing destination. Denote N the total number of sourcing places and M the total number of stages, then the scale of MO_MSITNDP can be defined as $N \times M$. As shown in Figure 1, there are totally $M + 1$ predefined transportation stages (i.e., stage 0, stage 1, ..., stage M) in the network, where stage 0 is a virtual point without intermediate nodes. The transport request $A_1, ..., A_N$ deriving from the corresponding sourcing places $1, ..., N$ will be responded by transporting the freights from segment 1 to M (i.e., destination B). During the transportation process, the transport requests A_i $(i = 1, ... N)$ are processed independently without dismounting or consolidation. In this case, there exist crossing logistics in the intermediate nodes owing to the share of intermodal terminal, and the capacity constraint of each node must be considered. All the freights must be irreversibly transported from the stage 1 to M in sequence. The changes of transportation modes only occur on the stages rather than transportation process. Next, we will propose the multi-objective integer programming model of MO_MSITNDP.
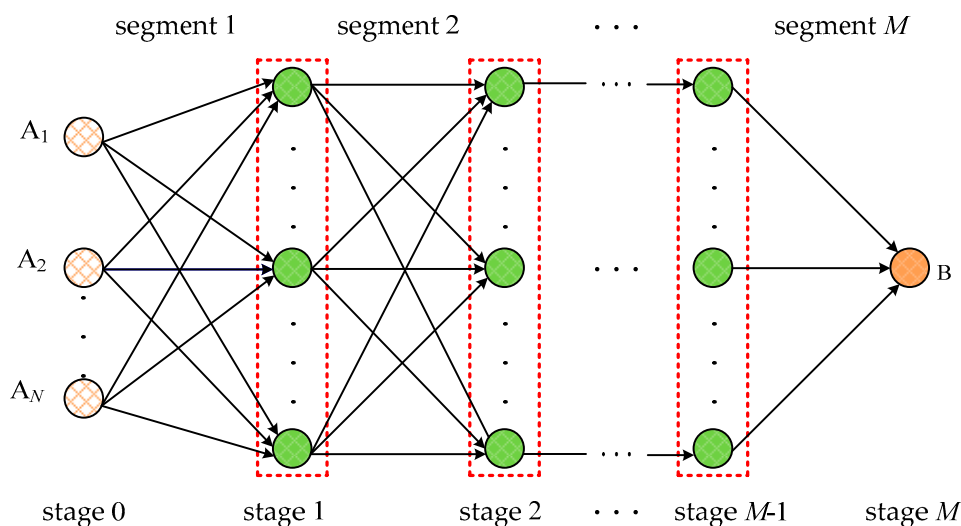
Figure 1. Structure of multi-sourcing intermodal transportation network.

In our model, the objective functions to be minimized consist of the total cost (TTC) and maximum flow time (MFT). Furthermore, the TTC includes three parts, i.e., transportation cost, transfer cost of transportation mode and carbon emissions cost, thus our solutions can reflect the economic benefit, time efficiency and environmental sustainability (society) criteria. The MFT is the maximum completion time of all freights, so a lower MFT indicates a higher efficiency of the whole systems. Accordingly,

our ultimate objective is to find a good trade-off among the economy, efficient and sustainability sub-objectives, and these three criteria are usually in conflict with each other. The optimization of IT is essentially to provide a satisfied solution for the intermodal transportation company with reducing the transportation cost, transportation time and carbon emissions simultaneously.

Parameters:

$\mathcal{M}$:   An infinite number;

Q_i :   Amount of freight from sourcing place i, $i \in \{1, \ldots, N\}$;

R_j :   Set of available nodes of stage j, $j \in \{1, \ldots, M-1\}$;

R_0 :   Available nodes of stage 0 and $|R_0|=N$;

R_M :   Available node of stage M and $|R_M|=1$;

S_j :   Set of available transportation modes of stage j, $j \in \{1, \ldots, M\}$;

W_{kj} :   The handling capacity for the kth node of stage j, $j \in \{1, \ldots, M-1\}$, $k \in \left\{1, \ldots, \left|R_j\right|\right\}$, $W_{10}, W_{20}, \ldots, W_{N0}$, and $W_{1M} = \mathcal{M}$;

$e_i^w(j-1, j, l, h)$:   Carbon emission cost of freight i from the lth node in stage $j-1$ to the hth node in stage j under the transportation mode w, $i \in \{1, \ldots, N\}$, $w \in S_j$, $l \in R_{j-1}, h \in R_j$;

$C_i^w(j-1, j, l, h)$:   Transportation cost of freight i from the lth node in stage $j-1$ to the hth node in stage j under the transportation mode w, $i \in \{1, \ldots, N\}$, $w \in S_j$, $l \in R_{j-1}, h \in R_j$;

$T_i^w(j-1, j, l, h)$:   Transportation time of freight i from the lth node in stage $j-1$ to the hth node in stage j under the transportation mode w, $i \in \{1, \ldots, N\}$, $j \in \{1, \ldots, M\}$, $w \in S_j$, $l \in R_{j-1}, h \in R_j$;

$\xi_i^{w,v}(j, k)$:   Switch cost of freight i from transportation mode w to v for the kth node of stage j, $i \in \{1, \ldots, N\}$, $j \in \{1, \ldots, M-1\}$, $w \in S_j, v \in S_{j+1}$, $k = 1, \ldots, \left|R_j\right|$;

$\tau_i^{w,v}(j, k)$:   Switch time of freight i from transportation mode w to v for the kth node of stage j, $i \in \{1, \ldots, N\}$, $j \in \{1, \ldots, M-1\}$, $w \in S_j, v \in S_{j+1}$, $k = 1, \ldots, \left|R_j\right|$; and

$U^w(j-1, j) = \begin{cases} 1 & \text{if transportation mode } w \text{ exists between stage } j-1 \text{ and stage } j \\ 0 & \text{otherwise} \end{cases}$.

Decision variables:

$x_i^w(j-1, j, l, h) = \begin{cases} 1 & \text{if transportation mode } w \text{ is selected for freight } i \\ & \text{from the } l\text{th tnode in stage } j-1 \text{ to the } h\text{th node in stage } j \\ 0 & \text{otherwise} \end{cases}$;

$y_i^{w,v}(j, k) = \begin{cases} 1 & \text{if transportation mode is transfered from } w \text{ to } v \\ & \text{for freight } i \text{ in the } k\text{th node of stage } j \\ 0 & \text{otherwise} \end{cases}$;

st_{ij} the start time at the stage j for freight i; and

ct_{ij} the leave time at the stage j for freight i.

Based on the aforementioned definitions, we detail the mathematical model of MO_MSITNDP as follows.

$$\min f_1 = \sum_{i \in \{1, \ldots, N\}} \sum \left\{ \begin{array}{l} \sum\limits_{j \in \{1, \ldots, M\}} \left(\sum\limits_{w \in S_j} \sum\limits_{l \in R_{j-1}} \sum\limits_{h \in R_j} \left(\begin{array}{l} C_i^w(j-1, j, l, h) \cdot x_i^w(j-1, j, l, h) \\ +e_i^w(j-1, j, l, h) \cdot x_i^w(j-1, j, l, h) \end{array} \right) \right) \\ + \sum\limits_{j \in \{1, \ldots, M-1\}} \sum\limits_{k \in R_j} \sum\limits_{w \in S_j} \sum\limits_{v \in S_{j+1}} \xi_i^{w,v}(j, k) \cdot y_i^{w,v}(j, k) \end{array} \right\}; \quad (1)$$

$$\min f_2 = \max_{i \in N} \left\{ \begin{array}{l} \sum\limits_{j \in \{1, \ldots, M\}} \sum\limits_{w \in S_j} \sum\limits_{l \in R_{j-1}} \sum\limits_{h \in R_j} T_i^w(j-1, j, l, h) \cdot x_i^w(j-1, j, l, h) \\ + \sum\limits_{j \in \{1, \ldots, M-1\}} \sum\limits_{k \in R_j} \sum\limits_{w \in S_j} \sum\limits_{v \in S_{j+1}} \tau_i^{w,v}(j, k) \cdot y_i^{w,v}(j, k) \end{array} \right\}; \quad (2)$$

s. t.

$$\sum_{j \in \{1,\dots,M\}} \sum_{w \in S_j} \sum_{l \in R_{j-1}} \sum_{h \in R_j} x_i^w(j-1,j,l,h) = 1; \forall\, i \in \{1,\dots,N\}; \tag{3}$$

$$\sum_{k \in R_j} \sum_{w \in S_j} \sum_{v \in S_{j+1}} y_i^{w,v}(j,k) = 1;\ \forall\, i \in \{1,\dots,N\}, j \in \{1,\dots,M-1\}; \tag{4}$$

$$\sum_{i \in \{1,\dots,N\}} \sum_{w \in S_j} \sum_{l \in R_{j-1}} x_i^w(j-1,j,l,h)Q_i \leq W_{hj}; \forall\, j \in \{1,\dots,M-1\}, h \in R_j; \tag{5}$$

$$x_i^w(j,j+1,g,l) + x_i^v(j+1,j+2,l,h) \geq 2y_i^{w,v}(j+1,k); \forall i \in \{1,\dots,N\}, j \in \{0,\dots,M-2\},$$
$$k \in R_{j+1}, w \in S_{j+1}, v \in S_{j+2}; \tag{6}$$

$$x_i^w(j,j+1,l,h) \leq U^w(j,j+1);\ \forall i \in \{1,\dots,N\}; j \in \{0,\dots,M-1\}, w \in S_{j+1}; \tag{7}$$

$$st_{ij} = ct_{i,j-1} + \sum_{w \in S_j} \sum_{l \in R_{j-1}} \sum_{h \in R_j} T_k^w(j-1,j,l,h)x_i^w(j-1,j,l,h); \forall i \in \{1,\dots,N\}, j \in \{1,\dots,M\}, k \in R_j; \tag{8}$$

$$ct_{ij} = st_{ij} + \sum_{k \in R_j} \sum_{w \in S_j} \sum_{v \in S_{j+1}} \tau_i^{w,v}(j,l) \cdot y_i^{w,v}(j,k);\ \forall i \in \{1,\dots,N\}, j \in \{1,\dots,M-1\}; \tag{9}$$

$$st_{ij} \geq ct_{i,j-1};\ \forall i \in \{1,\dots,N\}, j \in \{1,\dots,M\}; \tag{10}$$

$$x_i^w(j-1,j,l,h) \in \{0,1\};\ \forall i \in \{1,\dots,N\}, j \in \{1,\dots,M\}, l \in R_{j-1}, h \in R_j, w \in S_j; \tag{11}$$

$$y_i^{w,v}(j,k) \in \{0,1\}; \forall i \in \{1,\dots,N\}, j \in \{1,\dots,M-1\}, k \in R_j, w \in S_j, v \in S_{j+1}; \tag{12}$$

$$z_{ijk} \in \{0,1\}; \forall i \in \{1,\dots,N\}, j \in \{0,\dots,M-1\}, k \in R_j; \tag{13}$$

$$st_{ij} \geq 0; \forall i \in \{1,\dots,N\}, j \in \{1,\dots,M\}; \tag{14}$$

$$ct_{ij} \geq 0; \forall i \in \{1,\dots,N\}, j \in \{0,\dots,M-1\}. \tag{15}$$

Equations (1) and (2) are the objective functions. Equation (1) is to minimize the TTC which includes transportation cost, transfer cost of transportation mode, and carbon emission cost, whereas Equation (2) minimizes the MFT. Constraint (3) ensures that only one transportation mode can be chosen for two adjacent stages. Constraint (4) ensures the mode transferring only can be executed at the intermodal terminal. Constraint (5) represents the capacity restrictions for each node with respect to given stage. Constraint (6) is the transportation continuity constraints for all routes. Constraint (7) is the transportation mode constraints between two adjacent stages associated with given freight and nodes. Equations (8) and (9) are the definitions of arriving times and departing times with regard to given freight, stages, and selected nodes. Constraint (10) implies that the overall structure of MO_MSITNDP is a directed acyclic graph from the time perspective. Constraints (11)–(15) are the ranges of values for variables.

3.2. Modification of Established Model

In view of the aforementioned model, the MO_MSITNDP is essentially a bi-objective optimization problem with capacity constraints, i.e., Constraint (5). Thus, we should introduce relevant constraints handling techniques into HEDA to realize its global search. Unlike single objective optimization problem, the traditional penalty function method may be unusable for multi-objective optimization scenario. In addition, since the solution representation of MO_MSITNDP is a matrix which combines transportation modes with intermediate nodes together, the repair-based constraints handling techniques are more likely to result in a lower efficiency of computation, especially for large scale or strong constraints conditions. Owing to such intractability, we apply the multi-objective-based constraints handling approach [36,37], in which the degree of constraint violation is regarded as an additional objective function. In this regard, the Constraints (5) can be transformed to the third

objective function, which denotes the total degree of constraint violation with regard to all intermediate nodes as shown by Equation (16):

$$\min f_3 = \sum_{j \in \{1,\dots,M-1\}} \left\{ \max_{h \in R_j} \left\{ \sum_{i \in \{1,\dots,N\}} \sum_{w \in S_j} \sum_{l \in R_{j-1}} x_i^w(j-1,j,l,h)Q_i - W_{hj}, 0 \right\} \right\}; \tag{16}$$

Afterwards, the modification of the established model can be given as follows:

$$\min\{f_1, f_2, f_3\}. \tag{17}$$

where the first objective functions f_1 and f_2 are the same as Equations (1) and (2). Except for removing Constraint (5), other constraints of this modified version keep unchanged as the previous model.

In this work, HEDA deals with the modified version model in scene of the Pareto dominance. The criteria f_1 and f_2 are the ultimate optimization objective associated with the MO_MSITNDP, whereas the additionally introduced objective function f_3 is utilized to guide the search direction from the feasibility perspectives.

4. HEDA for MO_MSITNDP

4.1. Solution Representation

Since the intermediate nodes and transportation modes need to be determined simultaneously, we specially design a matrix $\boldsymbol{\pi} = [\pi_{ij}]$ $(i = 1, \dots, N; j = 1, \dots, 2M-1)$ to represent MO_MSITNDP with the scale $N \times M$, in which the element π_{ij} denotes the index of selected node (if $j \bmod 2 = 0$) or transportation mode (if $j \bmod 2 \neq 0$) of the ith sourcing place. In addition, since freight i $(i = 1, \dots, N)$ need to undergo M nodes and M segments in turns after departing from its sourcing places, the sourcing places (i.e., stage 0) and the common destination (i.e., stage M) are not necessary to be considered. Obviously, if $j \bmod 2 = 0 \cap \pi_{ij} \leq \left| R_{\lceil j/2 \rceil} \right|$, then π_{ij} denotes the selected intermediate node of freight i in stage $\lceil j/2 \rceil$, for $i = 1, \dots, N; j = 1, \dots, 2M-1$. Similarly, π_{ij} denotes the selected transportation mode of freight i in segment $\lceil (j+1)/2 \rceil$, if $j \bmod 2 \neq 0 \cap \pi_{ij} \leq \left| S_{\lceil (j+1)/2 \rceil} \right|$, for $i = 1, \dots, N; j = 1, \dots, 2M-1$. Here, we take a simple example with the size $N \times M = 4 \times 5$ to illustrate this solution representation approach. Considering the matrix

$$\pi_{ij} = \begin{bmatrix} \textcircled{1} & 3 & \textcircled{2} & 4 & \textcircled{1} & 3 & \textcircled{1} & 1 & \textcircled{2} \\ 2 & 4 & 1 & 4 & 2 & 4 & 2 & 2 & 1 \\ 1 & 2 & 2 & 2 & 2 & 2 & 2 & 4 & 1 \\ 2 & 3 & 1 & 3 & 1 & 3 & 2 & 3 & 2 \end{bmatrix}_{N \times (2M-1) = 4 \times 9},$$

which indicates that freight 1 selects the transportation mode 1, 2, 1, 1 and 2 from segment 1 to 5 (marked by circles) and selects the intermediate node 3, 4, 3 and 1 in terms of stage 1 to 4, and so on.

4.2. Multi-Objective Handling Method

To store the non-dominated solutions (i.e., approximate Pareto set), a Pareto archive (PA) is introduced. For a multi-objective optimization problem with K criteria, if we say that π_1 dominates π_2 (i.e., $\pi_1 \prec \pi_2$) then both of the following conditions should be satisfied simultaneously: [38] (1) $f_k(\pi_1) \leq f_k(\pi_2)$, for $k = 1, \dots, K$, and (2) it has at least one $l \in \{1, \dots, K\}$ that $f_l(\pi_1) < f_l(\pi_2)$. Let PS denote the population size of HEDA, $\boldsymbol{N}\pi_i(gen)(i = 1, 2, \dots, PS)$ the ith individual in current population at generation gen, $AC(gen)$ the length of PA at generation gen, $\boldsymbol{P}\pi_i(gen)(i = 1, \dots, AC(gen))$ the ith non-dominated solution of PA at generation gen. The newly generated individual $\boldsymbol{N}\pi_i(gen)(i = 1, 2, \dots, PS)$ will be added to PA if and only if it can dominate at least

one solution in current PA, and, meanwhile, the solution that is dominated by newly generated individual will be removed from PA. Here, we propose the update procedure of PA as shown in Figure 2.

Algorithm 1: Update the Pareto archive.

01	**For** *i*:= 1 to *PS* **do**
02	**For** *ii*:= 1 to *AC*(*gen*) **do**
03	**If** $N\boldsymbol{\pi}_i(gen)) \prec \boldsymbol{P\pi}_{ii}(gen)$ **then**
04	$N\boldsymbol{\pi}_i(gen)) \rightarrow$ PA , and remove $\boldsymbol{P\pi}_{ii}(gen)$ from the current PA;
05	**For** *iii*:= *ii* + 1 to *AC*(*gen*) **do**
06	**If** $N\boldsymbol{\pi}_i(gen) \prec \boldsymbol{P\pi}_{iii}(gen)$ **then**
07	Remove $\boldsymbol{P\pi}_{ii}(gen)$ from the current PA;
08	**End**;//end if
09	**End**; //end for *iii*
10	**End** //end if
11	**Else**
12	**Continue**; //skip to the next loop
13	**End**; //end else
14	**End**; //end for *ii*
15	**End**; //end procedure
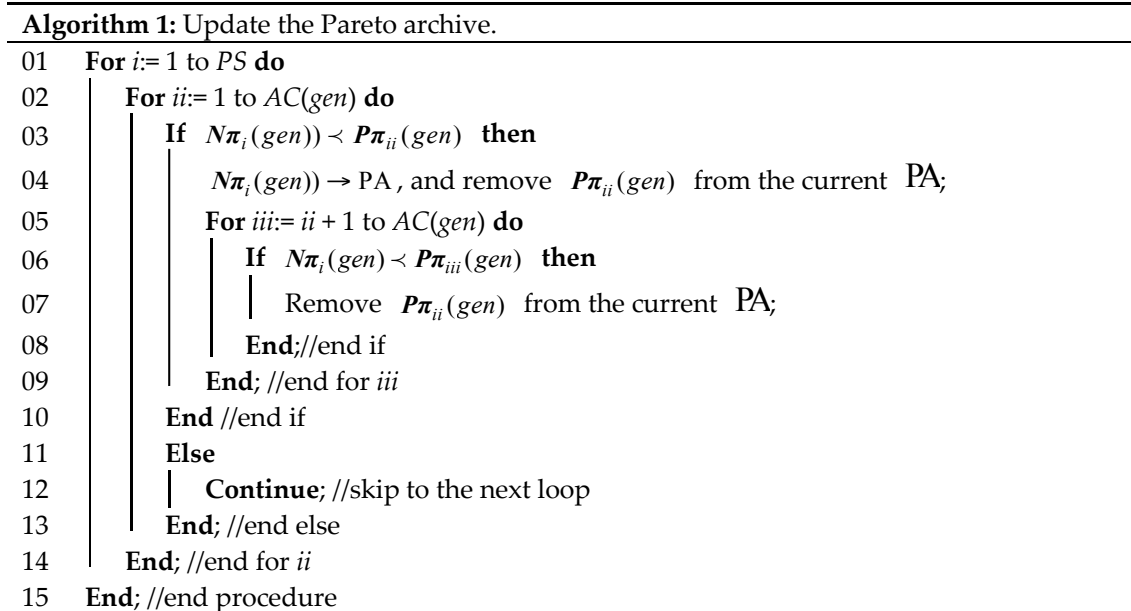
Figure 2. The procedure of updating the Pareto archive.

4.3. The Proposed Heterogeneous Probability Model and Update Mechanism

When using EDA to solve the discrete combinatorial optimization problems, the probability model is generally described as joint probability distribution function with independent random variables. Considering the inherent features of MO_MSITNDP, a heterogeneous marginal distribution law (HMDL) is put forward for serving as HEDA's probability model, which is utilized to represent the distribution of promising solution regions. The probability model $\boldsymbol{pr}(gen)$ can be formulated as follows:

$$
\begin{aligned}
&\boldsymbol{pr}_i(gen) = \boldsymbol{P}^i(gen) \cup \boldsymbol{R}^i(gen) \\[4pt]
&= \left[
\begin{array}{ccccc|ccccc}
p_{11}^i(gen) & \dots & p_{1j}^i(gen) & \dots & p_{1M}^i(gen) & r_{11}^i(gen) & \dots & r_{1j}^i(gen) & \dots & r_{1M-1}^i(gen) \\
\vdots & \dots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \dots & \vdots \\
p_{s1}^i(gen) & \cdots & p_{sj}^i(gen) & \cdots & p_{sM}^i(gen) & r_{r1}^i(gen) & \cdots & r_{rj}^i(gen) & \cdots & r_{rM-1}^i(gen) \\
\vdots & \dots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \dots & \vdots \\
p_{maxS,1}^i(gen) & \cdots & p_{maxS,j}^i(gen) & \cdots & p_{maxS,M}^i(gen) & r_{|R_1|,1}^i(gen) & \cdots & r_{|R_j|,j}^i(gen) & \cdots & r_{|R_M|,M-1}^i(gen)
\end{array}
\right], \\[4pt]
&\underbrace{\hphantom{p_{11}^i(gen) \dots p_{1M}^i(gen)}}_{\boldsymbol{P}^i(gen)} \qquad \underbrace{\hphantom{r_{11}^i(gen) \dots r_{1M-1}^i(gen)}}_{\boldsymbol{R}^i(gen)} \\
&\text{for } i = 1,\dots,N,
\end{aligned}
\tag{18}
$$

where $maxS = \max\limits_{j\in\{1,\dots,M\}}\{|S_j|\}$ is the maximum value of transportation modes, $p_{sj}^i(gen)$ is the probability of transportation mode s to be selected in the jth segment for the ith freight at generation gen, and $r_{rj}^i(gen)$ is the probability of selecting the rth node in stage j for the ith freight at generation gen.

Since there exists transportation mode constraint between stage $j-1$ and j (for $j = 1,\dots,M$), the relate probability should be set to 0 if $u^w(j-1,j)$=0 (for $j = 1,\dots,M$, $w = 1,\dots,maxS$). Therefore, the probability model can be initialized as follows:

$$
\boldsymbol{P}^i(0) = \begin{bmatrix} u^1(0,1) & \cdots & u^{maxS}(0,1) \\ u^1(1,2) & \cdots & u^{maxS}(1,2) \\ \vdots & \cdots & \vdots \\ u^1(M-1,M) & \cdots & u^{maxS}(M-1,M) \end{bmatrix} \cdot \begin{bmatrix} 1/|S_1| & 1/|S_2| & \cdots & 1/|S_M| \\ 1/|S_1| & 1/|S_2| & \cdots & 1/|S_M| \\ \vdots & \vdots & \cdots & \vdots \\ 1/|S_1| & 1/|S_2| & \cdots & 1/|S_M| \end{bmatrix}_{maxS \times M} ,
$$

$$
\boldsymbol{R}^i(0) = \begin{bmatrix} 1/|R_1| & 1/|R_2| & & 1/|R_M| \\ 1/|R_1| & 1/|R_2| & \cdots & 1/|R_M| \\ \vdots & \vdots & \cdots & \vdots \\ 1/|R_1| & 1/|R_2| & \cdots & 1/|R_M| \end{bmatrix} .
$$

(19)

Probability model update is one of the crucial operators for EDA [32], which is used to accumulate the historical information of excellent individuals found during the search process. Thereafter, the new populations are generated by sampling the updated probability model. As mentioned above, PA is dynamically updated by new generated individuals according to the dominance rule (see algorithm in Figure 1), so that PA can be seen as a set of excellent individuals. In view of this, we apply the non-dominated solutions that located in PA to update the probability model. Denote *LR* the learning rate of HEDA and x, y the temporary variables, then the probability model update method is given in Figure 3.

Algorithm 2: Update the probability model.

01　Randomly select an individual $\boldsymbol{\pi}^e = [\pi_{ij}^e]_{N \times (2M-1)}$ from the current PA;

02　Set $x := 1, y := 1$

03　**For** $i := 1$ to N **do**

04　　**For** $j := 1$ to $2M - 1$ **do**

05　　　**If** $j \bmod 2 \neq 0$ **then**

06　　　　$p_{\pi_{ij}^e, x}^i(gen) := p_{\pi_{ij}^e, x}^i(gen) + LR$;// concerned with transportation modes

07　　　　Normalize the probability $p_{sx}^i(gen)$ for each $s = 1, \ldots, maxS$;

08　　　　$x := x + 1$;//the next segment

09　　**End**

10　　**Else**

11　　　　$r_{\pi_{ij}^e, y}^i(gen) := r_{\pi_{ij}^e, y}^i(gen) + LR$;//concerned with intermediate nodes

12　　　　Normalize the probability $r_{ry}^i(gen)$ for each $r = 1, \ldots, |R_y|$;

13　　　　$y := y + 1$;//the next stage

14　　　**End**;//end else

15　　**End**;//end for j

16　**End**;//end procedure

Figure 3. The procedure of updating the probability model.

4.4. New Population Generation Method

HEDA generates new population by sampling the updated probability model (i.e., HMDL). In the sampling procedure, the new generated individuals should be consistent with the HMDL to guide HEDA's global search direction, and, meanwhile, to keep the population diversity, relevant random factors may need to be considered. That is, we tend to select a random variable in HMDL with larger probability for producing new individuals, whereas ones corresponding to relative smaller probability also have a chance to be selected. In this regard, we propose a sampling approach based on roulette wheel selection, for which the selection probability of a random variable is proportional to its probability in HMDL. Denote $\boldsymbol{FP}_j^i(gen) = [FP_{1,j}^i(gen), \ldots, FP_{maxS,j}^i(gen), FP_{maxS+1,j}^i(gen)]^T (j = 1, \ldots, M)$ the *j*th marginal distribution function (MDF) of probability matrix $\boldsymbol{P}^i$ at generation gen, $\boldsymbol{FR}_j^i(gen) = [FR_{1,j}^i(gen), \ldots, FR_{|R_j|,j}^i(gen), FR_{|R_j|+1,j}^i(gen)]^T (j = 1, \ldots, M-1)$ the *j*th MDF of

probability matrix R^i at generation *gen*, and x, y, rnd the temporary variables. Then, the procedure of calculating MDF and generating new population are shown in Figures 4 and 5.

Algorithm 3: Calculate MDF of HMDL.

01	**For** *i*:= 1 to *N* **do**
02	**For** *j*:= 1 to *M* **do**
03	Set $FP^i_{1,j}(gen) := 0$;
04	**For** *k*:= 2 to $maxS+1$ **do**
05	$FP^i_{k,j}(gen) := FP^i_{k-1,j}(gen) + p^i_{k-1,j}(gen)$;//calculate MDF
06	**End**;//end for *k*
07	**End**;//end for *j*
08	**For** *j*:= 1 to M − 1 **do**
09	Set $FR^i_{1,j}(gen) := 0$;
10	**For** *k*:= 2 to $|R_j|+1$ **do**
11	$FR^i_{k,j}(gen) := FR^i_{k-1,j}(gen) + r^i_{k-1,j}(gen)$;//calculate MDF
12	**End**;//end for *k*
13	**End**;//end for *j*
14	**End**;//end procedure

Figure 4. The procedure of calculating MDF.

Algorithm 4: Generate new population.

01	**For** *h*:= 1 to *PS* **do**
02	**For** *i*:= 1 to *N* **do**
03	Set $x := 1$, $y := 1$;
04	**For** *j*:= 1 to 2M − 1 **do**
05	$rnd := random[0,1)$;
06	**If** $j \bmod 2 \neq 0$ **then**
07	**For** *k*:= 1 to $maxS$ **do**
08	**If** $FP^i_{k,x}(gen) \leq rnd < FP^i_{k+1,x}(gen)$ **then**
09	$N\pi_{h,i,j}(gen) := k$;//sampling transportation mode
10	$x := x+1$;//the next segment
11	**End**;//end if
12	**End**;//end for *k*
13	**End**
14	**Else**
15	**For** *k*:= 1 to $|R_j|$ **do**
16	**If** $FR^i_{k,y}(gen) \leq rnd < FR^i_{k+1,y}(gen)$ **then**
17	$N\pi_{h,i,j}(gen) := k$;//sampling intermediate node
18	$y := y+1$;//the next stage
19	**End**;//end if
20	**End**;//end for *k*
21	**End**;//end else
22	**End**;//end for *j*
23	**End**;//end for *i*
24	**End**;//end for *h*

Figure 5. The procedure of generating new population.

4.5. Problem-Dependent Local Search

As mentioned above, HEDA may pay little attention to the local search during its iterative process. Hence, it would be highly advisable to embed problem-dependent local search into HEDA to exploit more promising solutions in relatively narrow regions of solution space. Since the solution representation of MO_MSITNDP is a $N \times M$ matrix, traditional neighborhood structures, such as Insert, Swap, Inverse, etc. [39], may be unsuitable. In this work, we propose a new neighborhood structure named partially interchanged neighborhood (PIN), in which two random positions $c_1, c_2 \in \{1, \ldots, 2M-1\}$, $c_1 < c_2$ are selected first and then we interchange the partial elements between c_1 and c_2 associated with two randomly selected r_1 and r_2, $r_1, r_2 \in \{1, \ldots, N\}$, $r_1 \neq r_2$. The mechanism of PIN can be illustrated by a simple example, as shown in Figure 6.
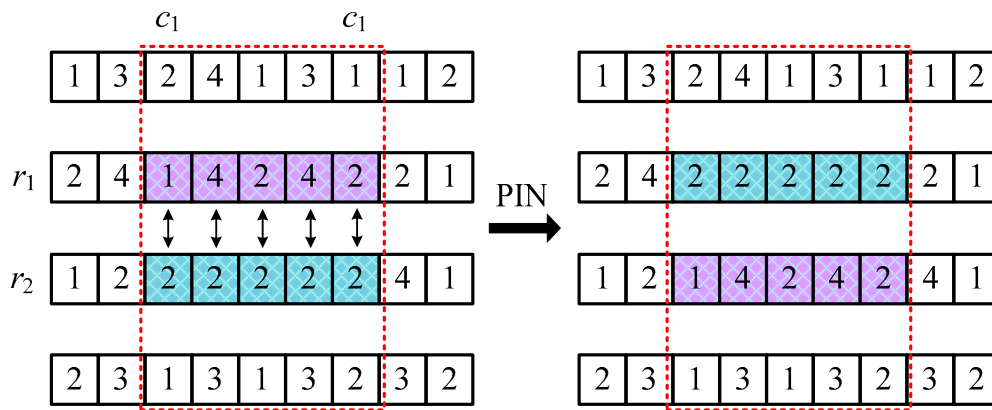

Figure 6. The schematic diagram of partially interchanged neighborhood (PIN).

Thereafter, we perform the PIN-based local search for all non-dominated solutions in PA, and meanwhile the first-move rule is adopted in order to avoid excessive exploitation. That is, once a relatively better solution is found from the neighborhood of current solution, the local search will be terminated. Let $\beta = \ell(\pi)$ denote a neighborhood solution associated with performing PIN operator one time on π, cnt the temporary variables. Then, the proposed local search operator is given in Figure 7.

Algorithm 5: PIN-based local search.

01	**For** $i := 1$ to $AC(gen)$ **do**
02	$\quad cnt := 0;$
03	$\quad$ **Repeat**
04	$\quad\quad \beta := \ell(P\pi_i(gen))$;//construct the neighborhood solution via PIN operator
05	$\quad\quad cnt := cnt + 1;$
06	$\quad$ **Until** ($cnt = 2M - 1$ **or** $\beta \prec P\pi_i(gen)$)
07	$\quad \beta \rightarrow P\pi_i(gen)$;//replace the current solution with its neighborhood solution
08	**End**;//end for i
09	Update the current PA;

Figure 7. The procedure of PIN-based local search.

4.6. Procedure of HEDA

Let *maxGen* denote the maximum number of iteration for HEDA. Based on aforementioned subsections, the procedure of HEDA is shown in Figure 8.

Algorithm 6: HEDA.	
01	**For** *gen*:= 1 to *maxGen* **do**
02	**If** *gen*=1 **then**
03	Randomly initialize the population $N\boldsymbol{\pi}_i(gen=1)$, for $i=1,2,...,PS$;
04	Construct the initial PA $P\boldsymbol{\pi}_i(gen=1)$, for $i=1,...,AC(gen=1)$;
05	Initialize the probability model $pr_i(gen=1)$, for $i=1,...,N$;
06	**End**
07	**Else**
08	Update the probability model $pr_i(gen)$, for $i=1,...,N$;// Algorithm 2
09	Generate new population $N\boldsymbol{\pi}_i(gen)$, for $i=1,2,...,PS$;// Algorithm 4
10	Update the PA $P\boldsymbol{\pi}_i(gen)$, for $i=1,...,AC(gen)$;// Algorithm 1
11	Perform the PIN-dependent local search;// Algorithm 5
12	**End**;//end else
13	**End**;//end for *gen*
14	**Output** $P\boldsymbol{\pi}_i(maxGen)$, for $i=1,...,AC(maxGen)$;

Figure 8. The procedure of HEDA.

5. Computational Results and Comparisons

To evaluate the performance of HEDA, we first perform computational experiments and comparisons based on randomly generated instances. Then, a real-life case study of Jilin Petrochemical Company (JPC), China, is carried out to demonstrate the application value of the proposed HEDA.

5.1. Experimental Setup

Since there is no benchmark data for MO_MSITNDP, we consider fifteen different combinations of $N \times M$ that include 5×3, $10 \times \{3,4,5\}$, $20 \times \{4,5,6,7\}$, $30 \times \{6,7,8,9\}$, $40 \times \{9,11\}$ and 50×13 for generating test instances. For each combination, we randomly generate 10 instances such that there is a total of 150 instances. The generation rules of these instances are listed as follows: (1) the amount of the *i*th freight $Q_i \in [100,500]$, for $i = 1,\ldots,N$; (2) the number of available nodes in the *j*th stage $|R_j| \in [1,6]$, for $j = 2,\ldots,M-1$, obviously $|R_0| = N$ and $|R_M| = 1$; (3) the transportation mode constraint $u^w(j-1,j)$, for $j = 1,\ldots,M$, is randomly established by assigning waterway, railway and road in each segment with 0 or 1; (4) the carbon emission cost $e_i^w(j-1,j,l,h) \in [1,100]$, transportation cost $C_i^w(j-1,j,l,h) \in [1,100]$, transportation time $T_i^w(j-1,j,l,h) \in [1,10]$, for $i = 1,\ldots,N$, $j = 1,\ldots,M$, $w = S_j$, $l \in R_{j-1}$; (5) the switch cost $\xi_i^{w,v}(j,k) \in [1,10]$, switch time $\tau_i^{w,v}(j,k) \in [1,3]$, for $i = 1,\ldots,N$, $j = 1,\ldots,M-1$, $w \in S_j, v \in S_{j+1}$,]; and (6) the handling capacity is calculated as $W_{kj} = \sum_{i=1}^{N} Q_i / r \cdot |R_j|$, $r \in [0.1,0.5]$, for $j = 1,\ldots,M-1$, $k = 1,\ldots,|R_j|$, $W_{10}, W_{20},\ldots,W_{N0}$ and W_{1M} are set to be infinite values.

We independently run all algorithms 20 times, and the results are all based on the average level of 10 instances with respect to the related $N \times M$ combinations. In the results, MIN, MAX, AVG and SD represent the minimum, maximum, average and standard deviation, respectively, corresponding to the related performance metrics. The unique terminated condition for all algorithms is the maximal evaluation times of objective functions, which is set to $1000 \times N \times M$. All algorithms are coded in C++ in Microsoft Visual Studio 2008 environment, and all the simulations are conducted on a PC with Intel Core-i5 3.30 GHz processor with 4 GB memory.

5.2. Performance Metrics

To systematically verify the performance of HEDA, three performance metrics [40] are adopted as follows:

(1)   Overall non-dominated solutions number (*ONSN*)

Denote $\sum \mathbb{PF} = \mathbb{PF}_1 \cup \mathbb{PF}_2, \cup \ldots, \cup \mathbb{PF}_L$ is the union of non-dominated solutions with regard to *Algorithm* 1, . . . , *l*, . . . , *L*, and then $ONSN(\mathbb{PF}_l)$ is the number of non-dominated solutions in $\mathbb{PF}_l$ which are not dominated by solutions in $\sum \mathbb{PF}$ as shown in Equation (20). The larger the value of $ONSN(\mathbb{PF}_l)$ is, the better performance of *Algorithm l*.

$$ONSN(\mathbb{PF}_l) = | \mathbb{PF}_l \setminus \{ x \in \mathbb{PF}_l \mid \exists y \in \sum \mathbb{PF} : y \prec x \} | . \tag{20}$$

(2)   Ratio of non-dominated solutions (*RNDS*)

Denote the ratio of solutions which are not dominated by solutions in $\sum \mathbb{PF}$, shown as following Equation (21). The higher the value of $RNDS(\mathbb{PF}_l)$ is, the better performance of *Algorithm l*.

$$RNDS(\mathbb{PF}_l) = | \mathbb{PF}_l \setminus \{ x \in \mathbb{PF}_l \mid \exists y \in \sum \mathbb{PF} : y \prec x \} | / | \mathbb{PF}_l | . \tag{21}$$

(3)   Average distance (*DIR*)

Denote $d_{yx}(\pi_i)$ is the shortest normalized distance from solution y in $\sum \mathbb{PF}$ to $\mathbb{PF}_l$ [40], then DIR can be defined in Equation (22). Obviously, a smaller $DIR(\mathbb{PF}_l)$ means a better convergence performance to $\sum \mathbb{PF}$, as well as a better distribution of $\mathbb{PF}_l$.

$$DIR(\mathbb{PF}_l) = \sum d_{yx}(\mathbb{PF}_l) / | \sum \mathbb{PF} | , \quad y \in \sum \mathbb{PF} . \tag{22}$$

5.3. Comparisons Results

5.3.1. Comparison with Existing Algorithms

In this subsection, we compare HEDA with three algorithms: (1) HEDA_noLS, a variant of HEDA in which the PIN-based local search operator is removed, and the parameter setting is the same as HEDA; (2) NSGAII [41], a well-known non-dominated sorting multi-objective genetic algorithm, which has been widely used for various multi-objective optimization problems; and (3) PGA [31], a relatively new reported genetic algorithm for a special kind of intermodal transportation problem which is quite similar to MO_MSITNDP. It should be mentioned that all the algorithms are executed based on the modified model in Section 3.2.

For the parameters setting, we set the population size $PS = 50$ for all algorithms, and, meanwhile, HEDA's learning rate $LR = 0.06$, the crossover and mutation probability of NSGAII and PGA are set to 0.8 and 0.2, respectively. It should be noted that we have embedded the PIN-based local search operator in to the NSGAII and PGA to ensure the fairness of comparisons.

We first make an observation on the approximate Pareto fronts with regard to the first instance of 10×5, 20×5, 30×7, 40×9 and 50×13, espectively. Furthermore, the obtained results of HEDA, HEDA_noLS, NSGAII and PGA can be seen in Figure 9.

It is can be obviously seen in Figure 9 that the solutions in Pareto set obtained by HEDA are better (both in terms of the quality and quantity of solutions) than the ones from the three other algorithms in all cases of variable problem sizes, which implies HEDA share a high superiority in solving MO_MSITNDP. Furthermore, the statistic results of these four algorithms are listed in Tables 1–4.

We can see in Tables 1–3 that the performance of HEDA is obviously better than other three compared algorithm associated with *ONSN*, *RNDS* and *DIR* metrics. It is clear that HEDA's performance is more superior than its variant HEDA_noLS, which demonstrate effectiveness of our proposed PIN-based local search. For most instances, HEDA can find more non-dominated solutions

than the other algorithms, and its solutions possess a shorter distance to the union of non-dominated solutions. Thus, not only does HEDA ensure an outstanding convergence to the optimal Pareto font, but it also keeps a reasonable distribution and diversity on its approximated Pareto font. The feasible ratio of HEDA reaches to 100%, as well as other algorithms, which demonstrated the reliability of our proposed HEDA. Accordingly, the proposed HEDA can be seen as an effective solution for addressing MO_MSITNDP.

In terms of the runtime, Figure 10 shows that HEDA has lower runtimes than NSGAII and PGA, which indicates that the running speed of HEDA is faster than NSGAII and PGA. The runtimes of HEDA are a little bit higher than HEDA_noLS, implying that the PIN-based local search of HEDA causes relative lower time consumption. In addition, the feasibility rates for all algorithms are 100%.

(a)

(b)

(c)

(d)

(e)
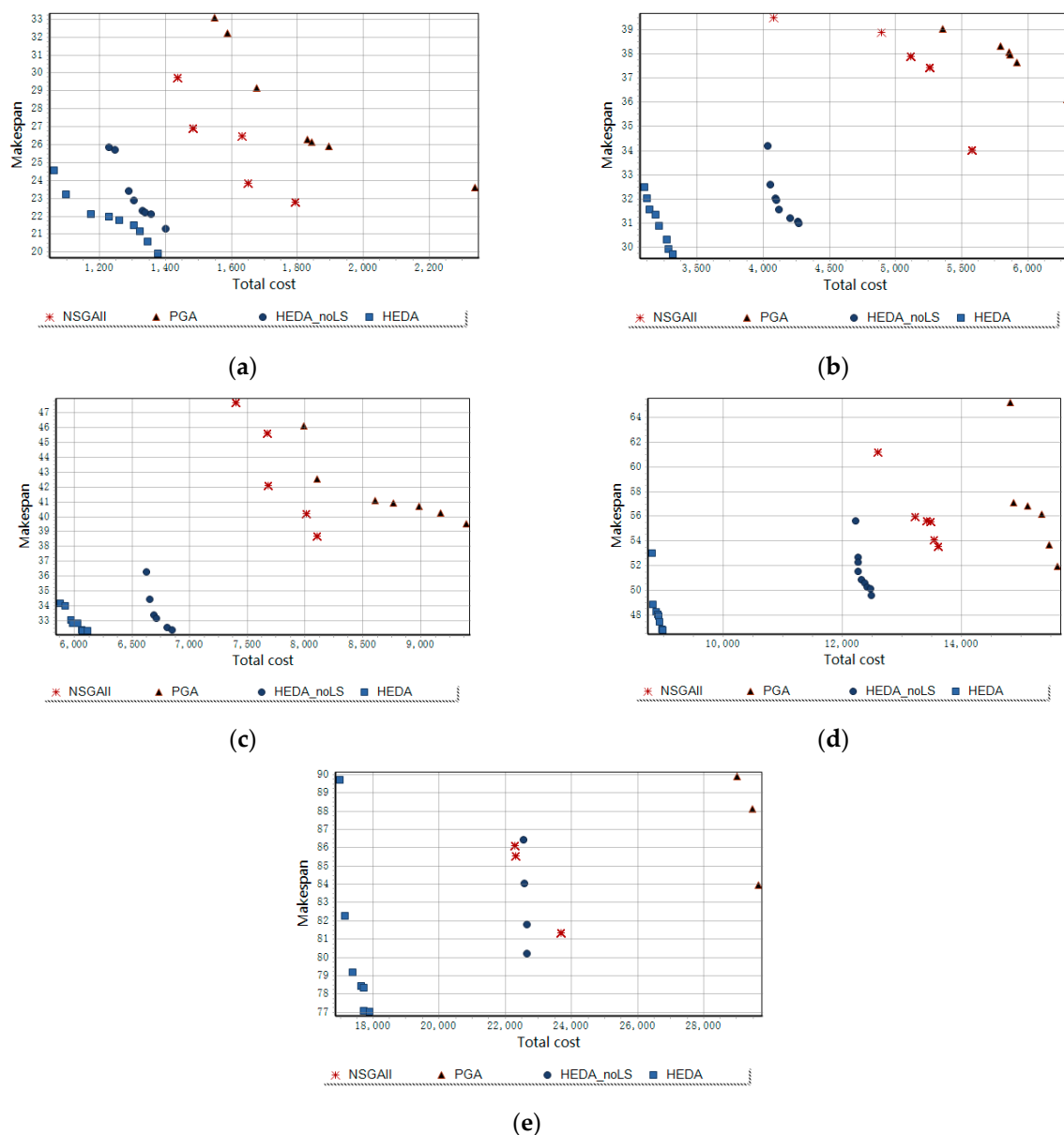
Figure 9. The Obtained approximate Pareto fonts of all algorithm. (**a**) 10×5; (**b**) 20×5; (**c**) 30×7; (**d**) 40×9; (**e**) 50×13.

**Table 1.** Comparisons of HEDA, NSGAII, PGA and HEDA-noLS (*ONSN*).

Instance	HEDA vs. NSGAII						HEDA vs. PGA						HEDA vs. HEDA_noLS					
	NSGAII			HEDA			PGA			HEDA			HEDA_noLS			HEDA		
	MAX	AVG	SD	MAX	AVG	SD	MAX	AVG	SD	MAX	AVG	SD	MAX	AVG	SD	MAX	AVG	SD
5×3	44	20.24	14.45	5	3.71	0.63	3	1.57	0.79	5	3.81	0.59	4	2.52	1.26	4	2.81	0.96
10×3	53	6.29	13.99	8	5.48	1.65	2	0.10	0.43	6	4.57	1.22	8	2.76	2.04	8	3.81	2.38
10×4	52	3.76	12.22	10	5.81	1.87	0	0.00	0.00	12	6.52	2.34	9	2.43	2.57	9	5.67	2.40
10×5	36	6.81	12.55	9	4.86	2.05	1	0.14	0.35	9	4.62	2.08	6	2.00	1.90	6	3.43	1.47
20×4	0	0.00	0.00	9	3.81	1.94	0	0.00	0.00	9	4.24	1.82	4	0.71	1.16	7	4.00	1.57
20×5	0	0.00	0.00	8	3.67	1.98	0	0.00	0.00	9	3.86	1.61	2	0.52	0.73	7	3.95	1.43
20×6	0	0.00	0.00	8	4.38	1.76	0	0.00	0.00	8	3.76	1.69	6	1.19	1.82	8	3.71	1.91
20×7	0	0.00	0.00	10	4.48	1.99	0	0.00	0.00	8	3.48	1.94	3	0.29	0.70	7	3.76	1.74
30×6	0	0.00	0.00	5	2.52	1.18	0	0.00	0.00	7	2.48	1.40	6	0.86	1.83	7	2.48	1.65
30×7	0	0.00	0.00	7	2.38	1.40	0	0.00	0.00	5	2.48	1.33	3	0.67	0.99	4	2.00	1.07
30×8	0	0.00	0.00	4	1.95	1.09	0	0.00	0.00	5	2.33	1.28	3	0.29	0.76	8	2.29	1.69
30×9	0	0.00	0.00	4	2.00	1.07	0	0.00	0.00	6	2.10	1.38	3	0.29	0.76	3	1.67	0.64
40×9	0	0.00	0.00	12	5.33	3.48	0	0.00	0.00	11	4.00	2.98	2	0.10	0.43	10	3.52	3.19
40×11	0	0.00	0.00	6	2.33	1.32	0	0.00	0.00	6	3.14	1.39	1	0.10	0.29	5	1.86	1.04
50×13	0	0.00	0.00	9	2.00	1.69	0	0.00	0.00	6	2.05	1.13	0	0.00	0.00	10	2.48	2.15
Average	12.33	2.47	3.55	7.60	3.65	1.67	0.40	0.12	0.10	7.47	3.56	1.65	4.00	0.98	1.15	6.87	3.16	1.69

Table 2. Comparisons of HEDA, NSGAII, PGA and HEDA-noLS (*RNDS*).

Instance	HEDA vs. NSGAII						HEDA vs. PGA						HEDA vs. HEDA_noLS					
	NSGAII			HEDA			PGA			HEDA			HEDA_noLS			HEDA		
	MAX	AVG	SD	MAX	AVG	SD	MAX	AVG	SD	MAX	AVG	SD	MAX	AVG	SD	MAX	AVG	SD
5 × 3	0.44	0.21	0.15	1.00	1.00	0.00	0.67	0.33	0.18	1.00	1.00	0.00	1.00	0.68	0.35	1.00	0.78	0.24
10 × 3	0.53	0.07	0.14	1.00	1.00	0.00	0.25	0.01	0.05	1.00	1.00	0.00	1.00	0.55	0.38	1.00	0.69	0.34
10 × 4	0.54	0.04	0.13	1.00	0.99	0.02	0.00	0.00	0.00	1.00	1.00	0.00	1.00	0.37	0.33	1.00	0.85	0.30
10 × 5	0.36	0.07	0.13	1.00	1.00	0.00	0.25	0.04	0.09	1.00	1.00	0.00	1.00	0.43	0.38	1.00	0.99	0.05
20 × 4	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.75	0.17	0.26	1.00	1.00	0.00
20 × 5	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.50	0.13	0.19	1.00	0.99	0.03
20 × 6	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	1.00	0.20	0.30	1.00	1.00	0.00
20 × 7	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.50	0.06	0.13	1.00	1.00	0.00
30 × 6	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.86	0.13	0.27	1.00	1.00	0.00
30 × 7	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	1.00	0.17	0.26	1.00	1.00	0.00
30 × 8	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.75	0.07	0.19	1.00	1.00	0.00
30 × 9	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	1.00	0.12	0.30	1.00	1.00	0.00
40 × 9	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	1.00	0.05	0.21	1.00	1.00	0.00
40 × 11	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	1.00	0.07	0.23	1.00	1.00	0.00
50 × 13	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00
Average	0.12	0.03	0.04	1.00	1.00	0.00	0.08	0.03	0.02	1.00	1.00	0.00	0.82	0.21	0.25	1.00	0.95	0.06

Table 3. Table 3. Comparisons of HEDA, NSGAII, PGA and HEDA-noLS (DIR).

Instance	HEDA vs. NSGAII						HEDA vs. PGA						HEDA vs. HEDA_noLS					
	NSGAII			HEDA			PGA			HEDA			HEDA_noLS			HEDA		
	MIN	AVG	SD	MIN	AVG	SD	MIN	AVG	SD	MIN	AVG	SD	MIN	AVG	SD	MIN	AVG	SD
5 × 3	36.64	123.36	67.77	0.00	1107.72	1012.93	61.66	116.48	74.67	0.00	60.73	51.70	0.00	15.66	19.03	0.00	25.87	53.26
10 × 3	181.89	544.88	265.34	0.00	2286.86	5945.60	158.87	793.87	400.31	0.00	55.16	246.67	0.00	183.72	191.64	0.00	128.18	163.28
10 × 4	269.57	1225.25	758.57	0.00	1550.08	6811.30	647.41	1960.32	800.04	0.00	0.00	0.00	0.00	390.81	331.60	0.00	93.72	138.05
10 × 5	1598.90	3824.40	1870.46	0.00	9359.46	17,387.86	2038.44	4826.59	2376.99	0.00	203.33	504.80	32.09	829.86	593.75	0.00	652.71	968.25
20 × 4	1054.05	2843.73	1558.43	0.00	0.00	0.00	1954.42	4134.30	1889.59	0.00	0.00	0.00	157.67	1649.53	1260.86	0.00	233.46	401.39
20 × 5	1216.20	4652.32	3060.08	0.00	0.00	0.00	2946.40	6184.59	2407.24	0.00	0.00	0.00	1096.54	2973.23	1518.30	0.00	404.02	581.21
20 × 6	1650.11	5755.40	2848.66	0.00	0.00	0.00	1967.32	7092.64	3234.82	0.00	0.00	0.00	217.86	2257.73	1475.94	0.00	691.68	1160.49
20 × 7	1343.27	8610.60	3947.62	0.00	0.00	0.00	2449.42	10,036.65	6093.77	0.00	0.00	0.00	833.73	4881.32	2931.44	0.00	371.87	954.23
30 × 6	2254.08	6060.90	3079.91	0.00	0.00	0.00	2639.69	8937.82	5210.97	0.00	0.00	0.00	1347.68	4364.83	2613.23	0.00	1512.26	3231.31
30 × 7	1785.60	6760.44	4068.56	0.00	0.00	0.00	4635.60	11,795.78	6546.49	0.00	0.00	0.00	1730.08	4839.59	3264.24	0.00	1572.35	2319.77
30 × 8	3116.28	8257.92	4925.34	0.00	0.00	0.00	5636.08	15,186.47	8625.25	0.00	0.00	0.00	2451.58	8163.42	6419.59	0.00	1162.14	3194.25
30 × 9	3917.91	9219.88	5022.42	0.00	0.00	0.00	6149.95	15,340.45	10,717.45	0.00	0.00	0.00	2664.78	5933.61	2339.85	0.00	1003.98	2603.61
40 × 9	5795.40	19,982.45	10,818.12	0.00	0.00	0.00	9074.13	34,600.99	21,887.46	0.00	0.00	0.00	4809.12	15,213.67	10,688.43	0.00	525.18	2348.70
40 × 11	2128.72	7800.28	4562.72	0.00	0.00	0.00	6379.80	20,857.98	9415.57	0.00	0.00	0.00	3564.77	7778.40	5030.00	0.00	439.48	1358.51
50 × 13	4697.53	10,692.39	6753.19	0.00	0.00	0.00	10,426.97	22,517.95	12,695.50	0.00	0.00	0.00	5898.19	16,269.53	11,700.77	0.00	0.00	0.00
Average	2069.74	6423.61	3573.81	0.00	953.61	2077.18	3811.08	10,958.86	6158.41	0.00	21.28	53.54	1653.61	5049.66	3358.58	0.00	587.79	1298.42
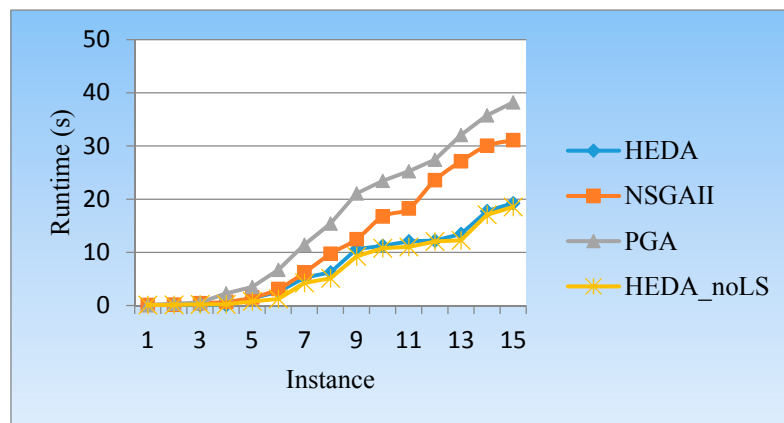
Figure 10. The trends of run time for the algorithms. * The feasibility rates of these four algorithms are 100%.

5.3.2. Comparison with Optimization Solver

In this subsection, the developed HEDA is compared with optimization solver. Owing to the linear characteristic of the mathematical model, we choose CPLEX 12.5, which is a popular solver used in optimization area. The multiple objectives are transformed into a single objective problem by using the epsilon constraint method [42] to satisfy the solving conditions of CPLEX. It should be noted that we do not purse the optimal solutions regarding the branch-and-cut phase addressed by CPLEX to obtain more feasible solutions to construct the approximate Pareto font. That is, once the feasible solution is found, CPLEX will stop and turn to solve the next problem in terms of another upper bound. Upper limit of time for CPLEX is 600 s. The results of *ONSN*, *RNDS*, *DIR* and Time are shown in Table 4.

In Table 4, we can find that, for the small scale instances, HEDA has lower runtimes and the quality of the solutions is also at an acceptable level. For the middle problem size, HEDA has advantages in runtimes and can provide better solutions than CPLEX. Furthermore, CPLEX is unable to solve the large scale instances, such as instance 40×11 and 50×13, whereas HEDA can obtain solutions easily, which indicates HEDA is more capable with respect to all instances.

Table 4. Comparisons of HEDA and CPLEX under a single run.

Instance	CPLEX				HEDA			
	ONSN	*RNDS*	*DIR*	Time	*ONSN*	*RNDS*	*DIR*	Time
5×3	47 *	1.00	0.00	<0.1	7	0.47	107.47	<0.1
10×3	55 *	1.00	0.00	1.23	10	0.58	645.23	<0.1
10×4	61 *	1.00	0.00	2.37	13	0.89	961.30	<0.1
10×5	38 *	1.00	0.00	5.24	8	0.91	809.26	<0.1
20×4	8	0.88	879.01	8.78	10	0.98	2.54	1.38
20×5	11	0.87	669.25	25.33	11	1.00	0.00	2.56
20×6	9	0.74	420.21	38.36	10	1.00	0.00	5.25
20×7	7	0.89	321.52	1 h limit	9	1.00	0.00	6.24
30×6	4	0.77	264.67	2 h limit	7	1.00	0.00	10.62
30×7	7	0.74	74.45	2 h limit	6	1.00	0.00	11.23
30×8	3	0.52	254.91	2 h limit	5	1.00	0.00	12.08
30×9	4	0.69	141.12	2 h limit	6	1.00	0.00	12.21
50×13	5	0.56	957.76	4 h limit	11	1.00	0.00	13.45
40×11	-	-	-	-	7	1.00	0.00	17.77
50×13	-	-	-	-	10	1.00	0.00	19.25

* Optimal solutions.

5.4. Case Study

In this section, we conduct a case study based on the ethylene propylene rubber production from Jilin Petrochemical Company (JPC), China. JPC is the only enterprise that has the production equipment of ethylene propylene rubber in China. The major raw materials for the ethylene propylene rubber are ethylene, propylene and third monomer. The ethylene needed is mainly self-produced, while the other materials rely on outsourcing and importing. In the JPC's sourcing plan, propylene will be purchased from Australia, while B fork borneol third monomer, double loop pentadiene third monomer and B diene third monomer are purchased from Japan, Philippines and America, respectively. In this background, an intermodal transportation company is hired by JPC to transport these freights from the sourcing places. Considering the geographical and economic distribution, 10 cities are chosen as intermediate nodes in the intermodal network (see Figure 11). The intermodal transportation network is divided into six stages and the nodes in each stage are shown in Table 5. The quantity of freights from each sourcing places are in Table 6. The relevant data about the variable transportation cost, carbon emission cost and average time for each transportation mode are in Table 7, in which the distances between the nodes by the three modes are collected from http://www.jctrans.com, www.12306.cn, and www.jdcsww.com in terms of the waterway, railway and road. Furthermore, the carbon emission relevant calculations for the three transportation modes are according to Duan and Heragu (2015) [1].

Figure 11. Intermodal transportation network for JPC.

Table 5. Nodes in each stage.

Stage	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Nodes	Hong Kong (1) Fuzhou (2) Wenzhou (3) Shanghai (4)	Wuhan (1) Nanjing (2)	Zhengzhou (1) Jinan (2)	Beijing (1) Tianjin (2)	Jilin (1)

Table 6. Quantity of freight from each sourcing place.

| Sourcing Place	Tokyo	America	Australia	Manila
Supply/TEU	80	50	50	120

Table 7. Parameters of transportation cost and speed for each mode.

Mode	Waterway		Railway	Road
Transportation cost ($/TEU-km)	0.2 (Ship)	0.18 (Barge)	0.5	2
Carbon emission cost ($/TEU-km)	0.018 (Ship)	0.015 (Barge)	0.03	0.05
Average speed (km/h)	40 (Ship)	30 (Barge)	70	70

According to the data above, we use HEDA to solve this case study and the non-dominated solutions are shown in Table 8 and Figure 12. Especially, to analyze the carbon emissions in our designed intermodal network, the carbon emission cost (CEC) is given.

Table 8. The Pareto set obtained by HEDA.

Solution	Tokyo	America	Australia	Manilas	TTC/$	MFT/h	CEC/$	Feasibility (Yes/No)
1	1-4-2-1-1-1-2-2-2	1-1-3-2-1-1-1-1-3	1-1-2-2-1-2-3-2-2	1-2-2-1-2-2-1-2	110,799	601	2330	Yes
2	1-1-2-2-1-2-2-2-2	1-1-2-2-1-1-3-2-3	1-1-2-2-1-1-2-2-2	1-1-2-2-1-2-2-2-2	99,097	614	2041	Yes
3	1-1-2-1-1-2-2-2-2	1-1-2-2-1-2-3-1-3	1-1-2-2-1-2-2-1-2	1-1-2-1-1-2-2-1-2	92,188	617	1854	Yes
4	1-4-2-2-1-2-2-1-2	1-1-2-2-1-1-2-2-2	1-1-2-1-1-2-3-2-2	1-1-2-2-1-2-2-1-2	92,141	662	1850	Yes
5	1-1-2-2-1-2-3-2-2	1-1-2-2-1-1-3-1-3	1-1-2-1-1-2-2-1-2	1-1-2-2-1-2-2-1-2	103,661	610	2106	Yes
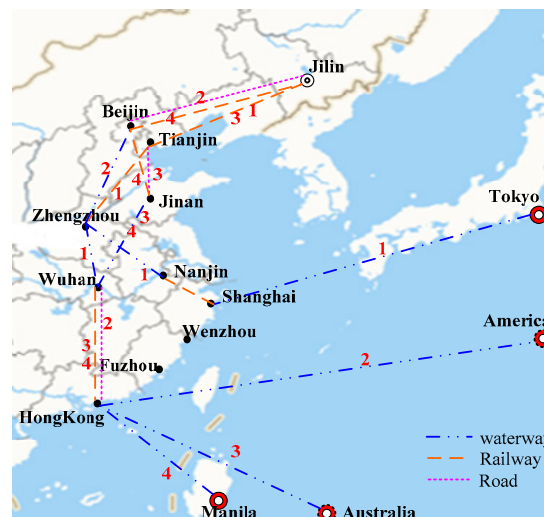
Figure 12. Transportation routes obtained via HEDA.

In Table 8, there are five non-dominated solutions provided by the HEDA for the intermodal transportation company to make transportation decision. By careful analysis, it can be found that the solutions have a good trade-off among cost, time and carbon emissions. For example, total cost and carbon emission cost of Solution 4 is 92,141 and 1850, respectively, which are a little higher than Solution 3 (92,188 and 617). However, the transportation time of Solution 3 is much less than Solution 4. It illustrates that time and carbon emissions can be greatly reduced without large sacrifice of cost.

Obviously, we can find that Solution 1 has a shortest maximum flow time of 601 while Solution 4 can provide a lowest total transportation cost with 92,141. Therefore, Solution 1 is proper if a fast delivery is needed, whereas Solution 4 should be adopted with respect to total cost focus. In addition, the rate of road use is very low; it has been replaced by waterway and railway, which have lighter carbon emission, implying that the carbon emission could be reduced in this network. Moreover, the stage division for an intermodal network that is mainly based on the physical characteristics of the available transportation infrastructures is a key preparation for the model construction. Through the gained solutions, it is obvious that some nodes are at a low frequency to be selected while some ones are invariably being passed. For example, the second node and third node are never selected in stage

1 for all solutions, which implies that we can consider removing the two nodes in stage division to decrease the computing complexity.

5.5. Management Insights

The approach and findings of this study will provide some practical insights for the logistics industry to make a tactical decision. An increasing number of transportation companies operate as freight integrators to provide one-stop service for customers. To enhance the competitiveness, they are supposed to optimize the transportation planning within the existing infrastructure and the expectations of customers. In this paper, the intermodal company can optimize the total cost and maximum flow time to satisfy diversified markets. In addition, carbon emissions are considered to achieve the environment sustainability and response to low-carbon policy and regulation. The model and proposed algorithm will enable the freight integrator to make a trade-off among the total cost, time and carbon emissions while providing highly complicated solutions to improve the competitiveness.

Furthermore, the crossing logistics is involved in the model as the handling capacity in each node. Such constraints can make our model close to reality and ensure the feasibility of solutions. Besides, it can balance the transportation flow among the nodes, which is beneficial to mitigate the traffic congestion and air pollution in busy node and reduce the idle resource in other nodes when the total cost or time change less.

As mentioned above, a main objective of intermodal transportation is to reduce the road transportation rate. In general, unless the fastest delivery is preferred, waterway and railway will cover the entire long-distance transportation and reduce the rate in the last-mile trucking transportation in intermodal transportation network. Therefore, intermodal transportation would be further advocated if the hard carbon emission regulations are set and government should focus on improving intermodal infrastructure construction, especially the port construction where waterway is available and high speed railway construction when waterways do not exist.

6. Conclusions

In this paper, the intermodal transportation problem was modeled as a MO_MITNDP problem concerning both the total cost and the time criteria. What mostly distinguished our model is the stage dividing according to the real intermodal infrastructure construction such as business type, geography characteristic and service radiant scope. We stand on the freight integrator's perspective to enhance the competitiveness and honor loyal customers by offering complicated decisions according to the personalized requirements. Specifically, the handling capacity in each node is programmed as a constraint, which will disperse transportation pressure by reducing the probability of freight assembling. To solve this hard constraint, the multi-objective-based constraints handling approach is adopted. Then, a hybrid estimation of distribution algorithm was designed to solve the problem. Numerical comparison results showed the effectiveness and robustness of the proposed algorithm for MO_MITNDP. Meanwhile, the case study from the petrochemical industry helped the intermodal company provide a highly door-to-door service in the trade-off among transportation cost, emission cost and the maximum flow time. The simulation results confirmed that we can save time and carbon emission without much sacrifice of transportation cost by the optimization of intermodal transportation network. Furthermore, this work is the first to consider the estimation of distribution algorithm when solving an intermodal transportation problem. In future work, it is interesting to propose other evolutionary algorithms to solve such problems and research on the low-carbon transportation network design to achieve the economic effectiveness and the environmental effectiveness.

Author Contributions: Both authors contribute equally to this work. R.L. wrote the initial manuscript draft; and S.J. performed several significant revisions on it.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Duan, X.; Heragu, S. Carbon Emission Tax Policy in an Intermodal Transportation Network. In Proceedings of the IIE Annual Conference, Nashville, TN, USA, 30 May–2 June 2015.
2. Zheng, H.Y.; Wang, L. Reduction of carbon emissions and project makespan by a Pareto-based estimation of distribution algorithm. *Int. J. Prod. Econ.* **2015**, *164*, 421–432. [CrossRef]
3. Bouchery, Y.; Fransoo, J. Cost, carbon emissions and modal shift in intermodal network design decisions. *Int. J. Prod. Econ.* **2014**, *164*, 388–399. [CrossRef]
4. Li, Y.; Bao, L.; Li, W.; Deng, H. Inventory and Policy Reduction Potential of Greenhouse Gas and Pollutant Emissions of Road Transportation Industry in China. *Sustainability* **2016**, *8*, 1218. [CrossRef]
5. Piecyk, M.I.; Mckinnon, A.C. Forecasting the carbon footprint of road freight transport in 2020. *Int. J. Prod. Econ.* **2010**, *128*, 31–42. [CrossRef]
6. Benjaafar, S.; Li, Y.; Daskin, M. Carbon footprint and the management of supply chains: Insights from simple models. IEEE Trans. *Autom. Sci. Eng.* **2013**, *10*, 99–116. [CrossRef]
7. Lam, S.L.; Gu, Y. A market-oriented approach for intermodal network optimisation meeting cost, time and environmental requirements. *Int. J. Prod. Econ.* **2016**, *171*, 266–274. [CrossRef]
8. Kim, N.S.; Wee, V.B. Toward a better methodology for assessing CO_2 emissions for intermodal and truck-only freight systems: A European case study. *Int. J. Sustain. Transp.* **2014**, *8*, 177–201. [CrossRef]
9. Zheng, Y.; Liao, H.; Yang, X. Stochastic Pricing and Order Model with Transportation Mode Selection for Low-Carbon Retailers. *Sustainability* **2016**, *8*, 48. [CrossRef]
10. Craig, A.J.; Blanco, E.E.; Sheffi, Y. Estimating the CO_2 intensity of intermodal freight transportation. *Transp. Res. D Transp. Environ.* **2013**, *22*, 49–53. [CrossRef]
11. Kaur, H.; Singh, S.P. Flexible dynamic sustainable procurement model. *Ann. Oper. Res.* **2017**. [CrossRef]
12. Steadieseifi, M.; Dellaert, N.P.; Nuijten, W.; Woensel, T.V.; Raoufi, R. Multimodal freight transportation planning: A literature review. *Eur. J. Oper. Res.* **2014**, *233*, 1–15. [CrossRef]
13. Meng, Q.; Wang, X. Intermodal hub-and-spoke network design: Incorporating multiple stakeholders and multi-type containers. *Transp. Res. B Meth.* **2011**, *45*, 724–742. [CrossRef]
14. Limbourg, S.; Jourquin, B. Optimal rail-road container terminal locations on the European network. *Transp. Res. E Log.* **2009**, *45*, 551–563. [CrossRef]
15. Vermam, M.; Verter, V.; Zufferey, N. A bi-objective model for planning and managing rail-truck intermodal transportation of hazardous materials. *Transp. Res. E Log.* **2012**, *48*, 132–149. [CrossRef]
16. Francesco, M.D.; Lai, M.; Zuddas, P. Maritime repositioning of empty containers under uncertain port disruptions. *Comput. Ind. Eng.* **2013**, *64*, 827–837. [CrossRef]
17. Bandeira, D.L.; Becker, J.L.; Borenstein, D.A. DSS for integrated distribution of empty and full containers. *Decis. Support Syst.* **2009**, *47*, 383–397. [CrossRef]
18. Yang, X.; Low, J.M.W.; Tang, L.C. Analysis of intermodal freight from China to Indian Ocean: A goal programming approach. *J. Transp. Geogr.* **2011**, *19*, 515–527. [CrossRef]
19. Resat, H.G.; Turkay, M. Design and operation of intermodal transportation network in the Marmara region of Turkey. *Transp. Res. E Log.* **2015**, *83*, 16–33. [CrossRef]
20. Domuţa, C.; Radu, D.; Enache, M.; Aştilean, A. Application for Intermodal Transportation Routing using Bi-objective Time-dependent Modified Algorithm. *Acta Electroteh.* **2012**, *53*, 154–159.
21. Namseok, K.; Janic, M.; Wee, B.V. Trade-off between carbon dioxide emissions and logistics costs based on multi-objective optimization. *Transp. Res. Rec. J. Transp. Res. Board* **2009**, *2139*, 107–116.
22. Rudi, A.; Fröhling, M.; Zimmer, K.; Schultman, F. Freight transportation planning considering carbon emissions and in-transit holding costs: A capacitated multi-commodity network flow model. *Eur. J. Transp. Log.* **2016**, *5*, 123–160. [CrossRef]
23. Demir, E.; Burgholzer, W.; Hrušovský, M.; Arıkan, E.; Jammernegg, W.; Woensel, T.V. A green intermodal service network design problem with travel time uncertainty. *Transp. Res. B Meth.* **2015**, *42*, 1439–1443. [CrossRef]

24. Lee, C.T.; Hashim, H.; Ho, C.S.; Fan, Y.V.; Klemeš, J.J. Sustaining the low-carbon emission development in Asia and beyond: Sustainable energy, water, transportation and low-carbon emission technology. *J. Clean. Prod.* **2017**, *146*, 1–13. [CrossRef]

25. KılıÇ, F.; Gök, M. A demand based route generation algorithm for public transit network design. *Comput. Oper. Res.* **2014**, *51*, 21–29. [CrossRef]

26. Vermam, M.; Verter, V. A lead-time based approach for planning rail–truck intermodal transportation of dangerous goods. *Eur. J. Oper. Res.* **2010**, *202*, 696–706. [CrossRef]

27. An, C.; Macharis, C.; Janssens, G.K. Corridor network design in hinterland transportation systems. *Flex. Serv. Manuf. J.* **2012**, *2*, 294–319.

28. Braekers, K.; An, C.; Janssens, G.K. Optimal shipping routes and vessel size for intermodal barge transport with empty container repositioning. *Comput. Ind.* **2013**, *64*, 155–164. [CrossRef]

29. Ishfaq, R.; Sox, C.R. Design of intermodal logistics networks with hub delays. *Eur. J. Oper. Res.* **2012**, *220*, 629–641. [CrossRef]

30. Li, L.; Negenborn, R.R.; Schutter, B.D. Intermodal freight transport planning—A receding horizon control approach. *Transp. Res. C Emerg.* **2015**, *60*, 77–95. [CrossRef]

31. Sterzik, S.; Kopfer, H.A. Tabu Search Heuristic for the Inland Container Transportation Problem. *Comput. Oper. Res.* **2013**, *40*, 953–962. [CrossRef]

32. Nossack, J.; Pesch, E. A truck scheduling problem arising in intermodal container transportation. *Eur. J. Oper. Res.* **2013**, *230*, 666–680. [CrossRef]

33. Zhang, J.; Zhang, Q.; Zhang, L. *A Study on the Paths Choice of Intermodal Transport Based on Reliability, 2014*; Springer: Berlin, Germany, 2015; pp. 305–315.

34. Wang, S.; Wang, L.; Xu, Y.; Liu, M. An effective estimation of distribution algorithm for the flexible job-shop scheduling problem with fuzzy processing time. *Int. J. Prod. Res.* **2013**, *51*, 3778–3793. [CrossRef]

35. Wang, L.; Wang, S.; Zheng, X. A hybrid estimation of distribution algorithm for unrelated parallel machine scheduling with sequence-dependent setup times. *IEEE. J. Autom. Sin.* **2016**, *3*, 235–246.

36. Salcedo-Sanz, S. A survey of repair methods used as constraint handling techniques in evolutionary algorithms. *Comput. Sci. Rev.* **2009**, *3*, 175–192. [CrossRef]

37. Wang, Y.; Cai, Z.; Guo, G.; Zhou, Y.Q. Multiobjective optimization and hybrid evolutionary algorithm to solve constrained optimization problems. *IEEE Trans. Syst. Man Cybern. B* **2007**, *37*, 560–575. [CrossRef]

38. Li, H.; Zhang, Q. Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II. *IEEE Trans. Evolut. Comput.* **2009**, *13*, 284–302. [CrossRef]

39. Nowicki, E.; Smutnicki, C. A fast tabu search algorithm for the permutation flow-shop problem. *Eur. J. Oper. Res.* **1996**, *91*, 160–175. [CrossRef]

40. Qian, B.; Wang, L.; Huang, D.X.; Wang, X. Multi-objective no-wait flow-shop scheduling with a memetic algorithm based on differential evolution. *Soft Comput.* **2009**, *13*, 847–869. [CrossRef]

41. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evolut. Comput.* **2002**, *6*, 182–197. [CrossRef]

42. Bérubé, J.F.; Gendreau, M.; Potvin, J.Y. An exact-constraint method for bi-objective combinatorial optimization problems: Application to the traveling salesman problem with profits. *Eur. J. Oper. Res.* **2009**, *194*, 39–50. [CrossRef]