IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 10, 2018, accepted April 9, 2018, date of publication April 12, 2018, date of current version May 2, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2826228

A Product-Focused, Cloud-Based Approach to Door-to-Door Railway Freight Design

WEICHUAN YIN[1,2], SHIWEI HE[2], YINGQUN ZHANG[2], AND JI HOU[2]
[1]MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing 100044, China
[2]School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

Corresponding author: Shiwei He (shwhe@bjtu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61374202 and Grant 71701014, in part by the Research Project of China Railway Company under Grant 2017X004-D and Grant 2017X004-E, and in part by the Natural Science Foundation of Beijing Municipality under Grant 9164032.

ABSTRACT This paper proposes a product-based framework that exploits cloud computing resources to improve the logistics of transporting door-to-door railway freight products. First, cluster analysis is used to determine the demand points on branch highway service networks and at stations on main railway service networks, enabling the establishment of a transfer service network between the two. Then, we develop a mathematical optimization model for door-to-door railway freight transportation products, in which the service time of arcs in the network is calculated based on railway freight car trajectory data. Finally, we examine the real-world case of the Wuhan–Dalang section of the Beijing–Guangzhou railway trunk line, demonstrating the practical value of big data and cloud computing technology for calculating the relevant parameters, as well as the adaptability of a framework that highlights product considerations.

INDEX TERMS Big data, cloud environment, door-to-door transport, freight product model, railway Transportation.

I. INTRODUCTION AND RELATED WORK

A. INTRODUCTION

Recent economic development has led to profound changes in the structure of transport demand, primarily reflected in the dramatic decrease in bulk cargo demand (i.e., the importation and exportation of bulk commodities) and rapid growth in the demand for scattered goods (i.e., electrical equipment, pharmaceutical products, stationery, etc.) [1]. Hence, the transformation to modern logistics must be accelerated for railway freight transport operations to adapt to changes in the transportation market, which will also benefit and help develop railway transport enterprises. Any such transformation should focus on customers' needs and given the diversification of customer demand serve all logistics processes, offering safe, efficient, economical, and reliable service. Railway freight transportation products [2] (formerly, scheduled freight train plans, which avoid many stops and focus on time limitations) can significantly enhance the quality of railway freight transportation services and customer satisfaction. Under such circumstances, railway freight transportation should also consider logistics services for additional improvement. Door-to-door railway freight products are an extension of railway freight transportation products, which include pick-up and delivery services. Research has not yet defined door-to-door railway freight products or examined their unified operation between China and other countries.

Cloud computing environments can better meet the logistical needs of railway enterprises; their rich resources can improve the efficiency and service quality of rail companies, strengthen internal management, and convert to new operation patterns, leading to a market-oriented form (i.e., an operation mechanism based on market demand).

The China Railway Corporation has thousands of yard stations and hundreds of thousands of trucks, and maintains a cargo volume on the scale of several billion tons per year [2]. Therefore, big data produced by railway transportation has a very large scale; how best to use big data to better guide transportation production is a very meaningful problem. Cloud computing technology can also help to solve the difficult problem of big data processing.

Traditional operation patterns of freight rail transport, focusing only on production, are gradually converted to the market pattern. Contrary to these patterns, the current freight transportation market requires diversity, individualization, and logistics, thus seriously restricting the coordinated and sustainable development of railway freight transportation.

Building a door-to-door railway service network based on a formal integrated transport network and establishing a complete transport chain directed by a control system is an irresistible trend, and a strategic choice that is not only based in scientific research, but also improves the efficacy and economic benefit of the freight rail system. Crainic divided freight transport plan optimization into three levels (strategic, tactical and operational), and summarized the literature and problem descriptions for each level [3]. This study investigates the tactical level.

In this study, we examine product-focused door-to-door railway freight design (PDRFD)—analyzing door-to-door railway transport service operation in a cloud-based environment—and propose a cloud-based approach framework that includes an optimization model based on a time–space network and large railway freight car trajectory datasets. The remainder of this paper is structured as follows. PDRFD is described in Section 2. A mathematical model is formulated in Section 3 and our solution approach is explained in Section 4. To validate the proposed model, we provide a case study based on China's Beijing–Guangzhou railway corridor in Section 5. Finally, Section 6 outlines our conclusions.

B. RELATED WORK

PDRFD is still a new and relatively unexplored problem. Few policy papers address door-to-door railway freight transport organization; in one such paper, Islam examined and identified barriers to and enablers for the European rail freight transport services as a transport chain and pointed out that the modern supply chain requires total transport chain [4], [5]. Relevant studies are mainly about the service network design (SND) problem; considering their similarities to door-to-door services using multimodal transport, we also analyze research on multimodal transport network design in this study.

SND is well studied, mostly through the lens of multicommodity network theory. Magnanti built a model of network design based on integer programming, presented a variety of algorithms, summarized the unified framework of the algorithm, and summarized network design problems in detail [6]. Crainic studied the design of freight service networks for multicommodity flow and multiple modes of transport and provided mathematical models and solution algorithms [7]. Minoux studied the optimization network design problem [8]. Nozick and Morlok [9] studied the medium-term planning of rail transit service networks and used integer linear programming to design an intelligent algorithm by relaxing the model. Ziliaskopoulos and Wardell [10] studied time–space path search algorithms for intermodal freight networks that account for transit time and linkage-arc transfer delay; Jeong *et al.* [11] considered Europe's radial freight network and established a linear integer programming model with a hub location constraint. Ceselli et al. proposed three models to describe problems such as the routing, dispatch, and locomotive operation of clipper freight transport. The model strictly meets time window constraints to ensure timely cargo delivery. A case study was used to verify and optimize cargo transport organization [12]. Cacchiani *et al.* [13] studied cargo transport organization in railway networks operating under the mixed logistics of passenger and freight transport. Yaghini and Akhavan [14] reviewed representative models and solved methods of multicommodity flow network design in railway operation planning, such as plan marshalling, route design, and empty truck dispatch. Yaghini *et al.* [15], [16] designed a mixed integer programming model based on hybrid heuristics to solve multicommodity flow network problems and applied it to the dynamic dispatching of rolling stock. Lin and Chen [17] conducted a hierarchical discussion of network design and time criteria in the courier industry, establishing a 0–1 integer programming model.

SND can be divided into two areas based on considerations of service time: static and dynamic service network design (DSND) problems. Representative achievements in DSND are illustrated by Crainic, who classified freight SND into frequency SND and DSND and reviewed the literature [18]. Dall'Orto explored the issue of DSND and proposed a time-dependent, stochastic formulation that aims to optimize the DSND problem over a given planning horizon [19]. Zhu considered traffic allocation and railway grouping strategy with decision-making variables to build a three-layer time–space network [20]. Additional applied research has investigated SND-problem-based evolutionary algorithms. Alanis *et al.* [21] proposed optimized solutions for a social-aware network with multi-objective routing and load balancing. Xu *et al.* [22] proposed a matching algorithm to optimize the social network content delivery problem. Duan *et al.* [23] proposed a new variable partitioning strategy in Benders decomposition to deal with a wide class of mixed-integer nonlinear programming problems, including fixed-charge multicommodity network design. Fakhri and Ghatee [24] demonstrated the effect of recognizing heterogeneity in time values on the design of a hub network for freight transportation. Lin *et al.* [25] considered the environmental benefit of building railways to design railway networks and developed a bi-level 0-1 programming model to balance investment, transportation, and environmental costs.

In the new era of cloud computing, an increasing number of enterprises and organizations migrate their business to cloud environments, developing services using big data. For example, researchers studying Facebook described how the social media site migrated its internal applications to hybrid cloud implementations [26]. Shen *et al.* [27] discussed how to migrate social-network applications to the cloud, reducing costs by taking advantage of rich cloud resources and on-demand calls and billing. From the perspective of big data processing, Zhang *et al.* [28] built an optimization model for efficiently migrating geographically distributed data to the cloud. Companies such as Facebook and Twitter processed Petabyte-level data using big data technologies and tools such as MapReduce and Hadoop by employing cloud resources.

Kemp *et al.* [29] proposed a cloud service-oriented approach for managing and analyzing big data required by transport applications. Lv *et al.* [30] proposed an OpenCL-accelerated point feature histogram method for railway track point cloud processing. Azizian *et al.* [31] proposed solutions for efficient data delivery based on transmission scheduling methods where vehicles gather data from mounted sensors. Luo *et al.* [32] designed an effective approach to improving prediction performance for cloud services by incorporating some intelligent techniques into quality-of-service prediction methods. Tawalbeh *et al.* [33] discussed networked health-care and the role of mobile cloud computing and big data analytics in its enablement.

The most representative achievements in SND were produced by Crainic, whose ideas have been foundational. Although Crainic has described his classification methods and general model analysis in detail, and proposed a research frame for such problems, his papers have not offered concrete manifestations of each classification nor designed solving algorithms. Furthermore, there are differences between PDRFD and SND. The decision variable of SND is whether to serve an arc, whereas for PDRFD it is to decide which product should be operated and at what frequency, synthesizing the consideration of stations, road sections, door-to-door transport situations, political and economic factors, and resource constraints stemming from Origin–Destination (OD) demand. Full transport planning SND for PDRFD must also account for the design of the access service network at both ends, which has not been done in existing literature. Furthermore, traditional railway optimization does not introduce cloud computing technology as an auxiliary method, even though optimization model parameters can be improved by the application of railway big data and cloud computing technology.

Railway SND has garnered increasing attention, and service network models and algorithms have been proposed based on the analyses described here. Although the differences between PDRFD and SND are significant (and PDRFD has excellent application prospects), few researchers have addressed a PDRFD optimization model or algorithm with big data technology applied. SND is a fundamental problem in theoretical research [7] that is generally applicable to rail networks. Our research on PDRFD is based on the SND problem, so it is also applicable to other rail networks.

II. PROBLEM DESCRIPTION

A. PRODUCT-FOCUSED DOOR-TO-DOOR RAILWAY FREIGHT DESIGN (PDRFD)

PDRFD differs from existing concepts of railway freight design problems. In existing railway freight design, railway transport enterprises design a transport service on the trunk railway network to meet the consignor's most basic cargo displacement demands. PDRFD considers not only the demand for trunk railway transportation, but also considers the demand for branch highway transport (see Fig. 1), resulting in a transportation service that meets the consignor's demand for transportation from the point of delivery to the point of arrival.

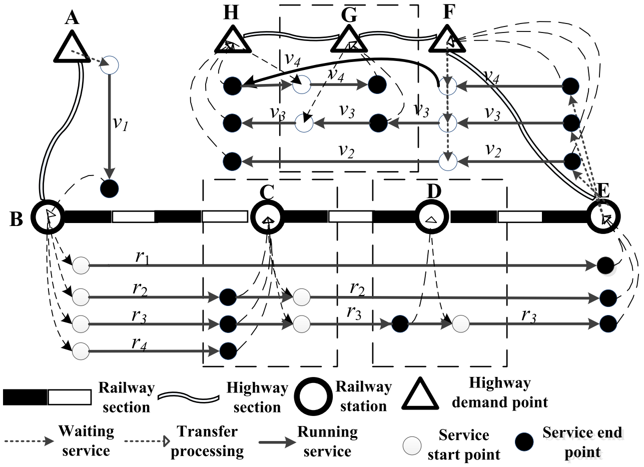


FIGURE 1. Illustration of a service network for door-to-door freight product.

As shown in Fig. 1, a shipment of goods from branch highway demand point A needs to be sent to branch highway demand points F and H, and highway transfer service to points F and H are known to be provided by trunk railway freight station E. Possible door-to-door freight products designed for this shipment include trunk railway freight product plans r_1–r_4 and branch highway freight product plans v_1–v_4. The transport services between two railways, a railway and highway, and two highways are illustrated via waiting, transfer, and running service arcs.

The construction of door-to-door railway freight service networks is based on transfer service decisions regarding which trunk railway freight station or technical station to connect to a highway demand point. Hub-and-spoke network theory can be used to explain door-to-door transfer transport services [34], [35], with the trunk railway freight station abstracted as a hub and the highway demand point as a spoke. Door-to-door railway freight product transport can be separated into three processes: gathering describes the process of traveling from spoke to hub, transport occurs between two hubs, and evacuation moves from hub to spoke. A hub-and-spoke network can be divided into single- and multi-assignment network structures based on the relationship between hub and non-hub nodes. This study examines a single-assignment network structure for the trunk railway station and branch highway demand point; however, non-hub nodes can be connected, i.e., a highway demand point is connected to other demand points. The door-to-door freight product transfer service network based on a single-assignment hybrid hub-and-spoke network is shown in Fig. 2.

For example, highway demand point 2 can connect to trunk railway freight stations A and B and highway demand point 4 can connect to trunk railway freight stations B and C, as shown in the upper half of Fig. 2. The objective of

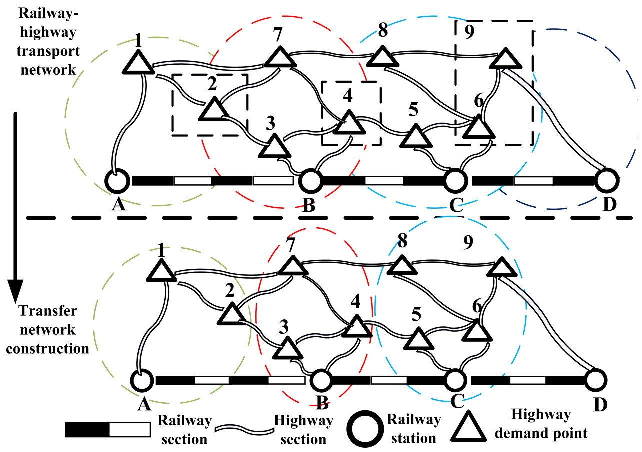


FIGURE 2. Illustration of a transfer service network construction for a door-to-door freight product.

constructing a door-to-door railway freight service network is deciding which highway demand point to connect with which trunk railway freight station. After the construction of door-to-door railway freight service network, highway demand point 2 can only connect to trunk railway freight station A and highway demand point 4 can only connect to trunk railway freight station B, as shown in the lower half of Fig. 2.

B. ANALYSIS OF DOOR-TO-DOOR RAILWAY TRANSPORT SERVICE OPERATION

At present, door-to-door railway freight transport organization can be described as a self-support half-way transport mode (SHTM). First, freight cargo is transported to the rail terminal by the highway transportation party and is transferred to trains belonging to the operation department of a door-to-door railway information cloud platform (DRICP). Finally, door-to-door service is arranged by the highway transportation party, fulfilling "last-mile" logistics. During this process, the highway transportation party could be individual drivers or third-party logistics companies responsible for the cargo from origination to rail terminal and from the rail terminal to its destination. Operation in this mode is shown in Fig. 3.

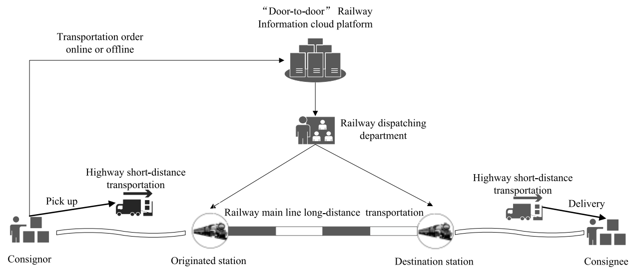


FIGURE 3. Self-support half-way carriage mode.

The main target clients of self-support half-way carriage mode are consignors who have sources of goods and carriers, individual drivers, or third-party logistics companies. However, DRICP operation departments' scheduling with highway transportation parties does not maximize their efficiency.

This study proposes a new freight traffic organization mode, called self-support whole-way transport mode (SWTM), which offers complete service throughout the transportation process by forming railway enterprises' own logistics team, i.e., replacing the highway transportation party with the railway enterprises' own logistics team to enhance operational efficiency. Freight cargo is transported from its origin to its destination using only the DRICP operation department. This process forms the basis of the proposed approach and the mathematical model in the next section.

III. PROPOSED MODEL FORMULATION

A. ASSUMPTIONS

We assume that freight demand is relatively stable during the decision-making period, and that the amount, characteristics, service level, and delivery time limits of the freight demand have been determined.

We further assume that the branch highway transport service capacity of a railway station is relatively stable during the decision-making period and that the travel speed of a branch highway transport service is also relatively stable.

B. FORMULATION

Based on the notations in Table 1, we propose the following mixed integer programming model for PDRFD (MIPM-PDRFD).

Objective function (4) maximizes the total door-to-door railway freight income, which includes transport revenue from Eq. (1), transport cost from Eq. (2), and penalty costs for unsatisfied demand from Eq. (3). The model constraints are as follows.

$$Rev = \sum_{d\in D} Q_d \psi_d \tag{1}$$

$$\begin{aligned} Cost_1 = & \sum_{s\in S}\sum_{k\in K} \eta_k f_{k,s} + \sum_{a\in A}\sum_{p\in P}\sum_{d\in D} Q_d x_{d,p}\delta_p^a \eta_a \\ & + \sum_{n\in N}\sum_{p_r\in P_r}\sum_{d\in D} Q_d x_{d,p_r}\lambda_{p_r}^n \gamma_n \end{aligned} \tag{2}$$

$$Cost_2 = \sum_{u\in U}\sum_{d\in D} Q_d x_{d,u}\varpi_u \tag{3}$$

$$Maxmize\ F = Rev - \mathrm{Cost}_1 - \mathrm{Cost}_2 \tag{4}$$

Constraint sets (5) and (6) ensure that each cargo demand can choose only one service arc by trunk railway transport and branch highway transport, respectively.

$$\boldsymbol{s.t.} \quad \sum_{p_r\in P_r} x_{d,p_r} + x_{d,u_r} = 1 \quad \forall d \in D \tag{5}$$

$$\sum_{p_v\in P_v} x_{d,p_v} + x_{d,u_v} = 1 \quad \forall d \in D \tag{6}$$

TABLE 1. Notation.

Symbol	Quantity
1)	*Sets*
$G = (N, E)$	Physical network of railways and highways, composed of railway freight yard N_r, highway demand point N_v, and section E;
N	Set of stations, indexed by n ($n \in N$);
E	Set of sections, indexed by e ($e \in E$);
P	Set of service arcs, indexed by p ($p \in P$), $P = P_r \cup P_v$, including trunk railway services $p_r \in P_r$, and branch highway services $p_v \in P_v$;
A	Set of arcs, indexed by a ($a \in A$), $A = A_r \cup A_v$, including trunk railway arcs $a_r \in A_r$ and branch highway arcs $a_v \in A_v$;
K	Set of product types $k \in K$;
S	Set of services $s \in S$;
D	Set of demand $d \in D$;
U	Set of superpaths for the service path of unsatisfied demand, $U = U_r \cup U_v$, $u \in U$.
2)	*Parameters*
Q_d	Amount of demand d;
δ_p^a	Service arc p is served by arc a, $\delta_p^a \in \{0,1\}$;
δ_p^s	Service arc p is included by service s or not, $\delta_p^s \in \{0,1\}$;
δ_s^e	Service s consists of section e, $\delta_s^e \in \{0,1\}$;
T_d	Time requirement of demand d;
$Lb_{k,s}$	Lower-bound service capacity of product k service s;
$Ub_{k,s}$	Upper-bound service capacity of product k service s;
ψ_d	Unit transport price calculated by railway department for demand d;
ϖ_u	Unit penalty price for demand d which is not satisfied;
η_k^s	Fixed transport price for product k service s;
η_a	Cost of arc a;
$Capv_n$	Capacity of transfer transport service to the highway point of railway station n;
Cap_n	Capacity of pass-through services at railway station n;
γ_n	Cost of transfer transport service to the highway point of railway station n;
Q_k^s	Maximum number of service s in product p;
N_e	Capacity of pass through service of section e;
t_a	Processing time of arc a;
$t_{p_r}^n$	Transfer processing time of trunk railway service arc p_r in railway station n;
$\lambda_{p_r}^n$	Trunk railway service arc p_r is included by railway station n, $t_{p_r}^n \in \{0,1\}$;
Θ	Large positive number.
3)	*Decision Variables*
$f_{k,s}$	Amount of service s in product p, $f_{k,s} \in \mathbb{Z}^+$;
$x_{d,p}$	Demand d is transported by service arc p or not, $x_{d,p} \in \{0,1\}$;
$y_{d,a}$	Amount of demand d assigned in arc a, $y_{d,a} \in \mathbb{Z}^+$.

Constraint sets (7) and (8) restrict the amount of cargo demand assigned to a trunk railway arc and branch highway arc, respectively.

$$y_{d,a_r} = Q_d x_{d,p_r} \delta_{p_r}^{a_r} \quad \forall d \in D,\ a_r \in A_r,\ p_r \in P_r \tag{7}$$

$$y_{d,a_v} = Q_d x_{d,p_v} \delta_{p_v}^{a_v} \quad \forall d \in D,\ a_v \in A_v,\ p_v \in P_v \tag{8}$$

Constraint sets (9) to (12) ensure that the trunk railway and branch highway product service must satisfy the operation requirements for cargo demand.

$$\sum_{p \in P_r} \sum_{a_r \in A_r} \sum_{d \in D} y_{d,a_r} \delta_{p_r}^{a_r} \delta_{p_v}^{s_v} \geq Lb_{k,s_r} f_{k,s_r} \quad \forall k \in K,\ s_r \in S_r \tag{9}$$

$$\sum_{p\in P_r}\sum_{a_r\in A_r}\sum_{d\in D} y_{d,a_r}\delta_{p_r}^{a_r}\delta_{p_r}^{s_r} \le Ub_{k,s_r}f_{k,s_r} \quad \forall k\in K,\ s_r\in S_r \tag{10}$$

$$\sum_{p\in P_v}\sum_{a_v\in A_v}\sum_{d\in D} y_{d,a_v}\delta_{p_v}^{a_v}\delta_{p_v}^{s_v} \ge Lb_{k,s_v}f_{k,s_v} \quad \forall k\in K,\ s_v\in S_v \tag{11}$$

$$\sum_{p_v\in P_v}\sum_{a_v\in A_v}\sum_{d\in D} y_{d,a_v}\delta_{p_v}^{a_v}\delta_{p_v}^{s_v} \le Ub_{k,s_v}f_{k,s_v} \quad \forall k\in K,\ s_v\in S_v \tag{12}$$

Constraint set (13) ensures that a railway station transferring cargo to branch highway demand cannot exceed its service capacity.

$$\sum_{k\in K}\sum_{s_v\in S_v} f_{k,s_v} \le Capv_n \quad \forall n\in N \tag{13}$$

The amount of service provided by a trunk railway product is formulated in Constraint (14).

$$\sum_{s_r\in S_r} f_{k,s_r} \le Q_k^{s_r} \quad \forall k\in K \tag{14}$$

The capacity of the physical section of a trunk railway for cargo to pass through is guaranteed by Constraint (15).

$$\sum_{k\in K}\sum_{s_r\in S_r} f_{k,s_r}\delta_{s_r}^{e}\varepsilon_k \le N_e \quad \forall e\in E \tag{15}$$

Constraint set (16) ensures the pass-through capacity of a railway station.

$$\sum_{d\in D}\sum_{p_r\in P_r} Q_d x_{d,p_r}\delta_{p_r}^{n} \le Cap_n \quad \forall n\in N \tag{16}$$

Constraint set (17) ensures that cargo transported by a door-to-door railway product can meet its time requirement.

$$\begin{aligned}\sum_{n\in N}\sum_{a_r\in A_r}\sum_{p_r\in P_r} & x_{d,p_r}(t_{a_r}\delta_{p_r}^{a_r}+t_{p_r}^{n}\lambda_{p_r}^{n})\\ &+\sum_{a_v\in A_v}\sum_{p_v\in P_v} x_{d,p_v}t_{a_v}\delta_{p_v}^{a_v} \le T_d \quad \forall d\in D\end{aligned} \tag{17}$$

Relations (18) to (21) are decision-variable-related constraints.

$$f_{k,s_r} \le x_{d,p_r}\delta_{p_r}^{s_r}\cdot\Theta \quad \forall d\in D,\ s_r\in S_r,\ p_r\in P_r \tag{18}$$
$$f_{k,s_v} \le x_{d,p_v}\delta_{p_v}^{s_v}\cdot\Theta \quad \forall d\in D,\ s_v\in S_v,\ p_v\in P_v \tag{19}$$
$$f_{k,s_r} \ge x_{d,p_r}\delta_{p_r}^{s_r} \quad \forall d\in D,\ s_r\in S_r,\ p_r\in P_r \tag{20}$$
$$f_{k,s_v} \ge x_{d,p_v}\delta_{p_v}^{s_v} \quad \forall d\in D,\ s_v\in S_v,\ p_v\in P_v \tag{21}$$

Relations (22) to (24) are the decision variable constraints.

$$x_{d,p} \in \{0,1\} \quad \forall d\in D,\ p\in P \tag{22}$$
$$y_{d,a} \in \mathbb{Z}^{+} \quad \forall d\in D,\ a\in A \tag{23}$$
$$f_{k,s} \in \mathbb{Z}^{+} \quad \forall k\in K,\ s\in S \tag{24}$$

IV. SOLUTION APPROACH

A. CLOUD PLATFORM TOOL

The cloud platform utilized in this study is a computer cluster with integrated data mining algorithms, which has performed well in processing big data and has good development prospects. This work calculates each process service arc time using railway freight car trajectory data and regards this platform as a suitable tool. We use the Ali cloud platform for processing big data. MapReduce is the core of open data processing service (ODPS) in Ali and is suitable for parallel computing of big data (sets larger than 1 Terabyte) [36]. In MapReduce, operations on the large-scale dataset are distributed to each network node to achieve reliability. Each node periodically returns completed work and the new status to distribute the computing task to the cluster, which reduces the customer's costs and completes the task quickly. SQL ODPS uses standard SQL syntax and is a much more efficient computing framework for supporting the SQL computing model, whose efficiency is much higher than the MapReduce model. ODPS can support more than one million offline scheduling tasks and provide stable offline scheduling capability based on the multi dimension of offline task scheduling, online operation, maintenance and monitoring alarm functions (see Fig. 4).

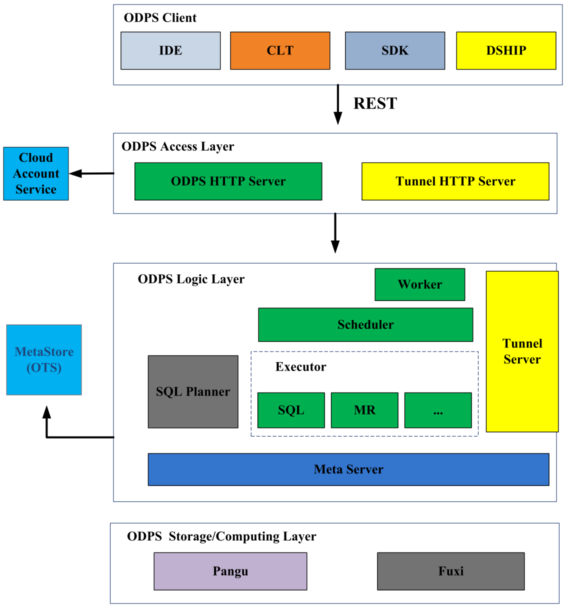


FIGURE 4. ODPS system framework.

B. CLUSTERING METHOD FOR PRODUCT-FOCUSED DOOR-TO-DOOR RAILWAY FREIGHT SERVICE NETWORKS

Section 2A confirmed that the key to constructing a door-to-door railway freight service network is deciding which trunk railway freight station or technical station should provide service to highway demand point (i.e., constructing the hub-and-spoke network). Therefore, we use cloud computing

TABLE 2. Data structure of railway freight car loading–unloading records.

Column Name	Data type	Explanation	Example	Remark
RPT_ID	String	report type	ZCBG（XCBG）	loading report （unloading report）
CAR_NO	String	number of cars	3423754	—
CDY_NAME	String	cargo category	Steel	—
RPT_ADM	String	loading bureau	P	Beijing bureau
ORG_STN	String	loading station	HDP	telegraph code, Handan
DEST_STN	String	unloading station	BDP	telegraph code, Baoding
IN_LINE_DATE	String	loading/unloading start date	2017/6/6	—
IN_LINE_TIME	String	loading/unloading start time	8:00	—
OUT_LINE_DATE	String	loading/unloading end date	2017/6/6	—
OUT_LINE_TIME	String	loading/unloading end time	9:20	—

TABLE 3. Data structure of railway freight car trajectory records.

Column Name	Data type	Explanation	Example	Remark
CAR_NO	String	number of car	3423754	—
WB_CDY_NAME	String	cargo category	Steel	—
ORG_ADM	String	dispatch bureau	N	Wuhan bureau
ORG_STN_NAME	String	dispatch station	Dalang	—
TRAIN_ARR_DEP_ADM	String	current bureau	Q	Guangzhou Railway Group
TRAIN_ARR_DEP_CHN	String	current station	Jiangcun	—
DEP_ARR_FLAG	String	status	A/D	arrive/dispatch
TRAIN_ARR_DEP_DATE	String	date	2017/6/4	—
TRAIN_ARR_DEP_TIME	String	time	17:33:00	—

resources to solve the transfer service network construction problem of door-to-door freight product with K-means clustering.

A fast, simple, nearly linear K-means clustering algorithm can be obtained from the Ali cloud platform. In this paper, we use an integrated K-means clustering algorithm module provided by the Ali cloud platform for programming with the following steps:

(1) Get the OD demand, latitude, and longitude of candidate points as a dataset. (2) Select k objects (values) from the dataset as the initial clustering center; in this case, we use trunk railway stations. (3) Calculate the distance between each object (value) in the dataset and the object (value) of the cluster center (herein, Euclidean distance), then divide the clusters so that each can only have one railway station. (4) Find a new cluster center object (value) within the clusters and meet the Euclidean distance between other objects in the clusters and the center objects, which is the best in the current cluster. (5) Repeat Steps (2)–(3) until the clusters no longer change or they meet the set of convergence conditions. (6) Output each cluster with the constraint that only one railway station is inside it. The railway station of the cluster need not be its center. Next, each cluster represents a hub with a connected spoke.

C. BIG DATA OF RAILWAY FREIGHT CAR TRAJECTORY

We propose a big data method to improve the accuracy of some parameters of our mathematical model, which gives our research a wide range of adaptability. Big data of railway freight car trajectory is collected from railway enterprises and stored in the data structure outlined in Tables 2–3. Railway freight car loading and unloading records provide information about the current load of each station, which affects its service processing arc time. Railway freight car trajectory records provide information about the running service arc time in each section when cargo passes through.

D. K SHORTEST PATH ALGORITHM IN A TIME–SPACE SERVICE NETWORK

At OD demand start point Φ_-, cargo is loaded from branch highway demand A to start station B of the trunk railway transport service, then passes through station C to end station D of the trunk railway transport service, unloading cargo from branch highway demand E to OD demand end point Φ_+ to complete the door-to-door transportation. Time $t = 1, 2, 3, \ldots, 14$ h, and the time–space service network is defined by section running, station processing, station running, station waiting, and transfer transport arcs. Each arc represents the transport service time and space displacement (see Fig. 5).

Door-to-door freight demand must be timely, making it necessary to select a feasible set of time–space service paths for decision optimization within time constraints. However, the number of service arcs in the spatiotemporal network increases with the number of nodes, leading to inefficiency in solving for a feasible service path.

Here, we design a k shortest path algorithm in a time–space service network to solve the feasible path set of door-to-door transport service. The algorithm's underlying principle is that in a time–space service network, the feasible path set of

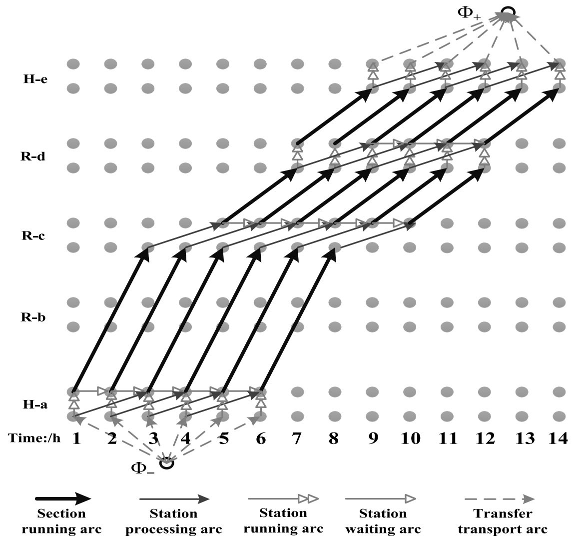


FIGURE 5. A time–space network for transporting goods door-to-door.

door-to-door transport is restricted to a limited, prism-shaped area. As shown in Fig. 6, the service paths in this area are feasible. The k shortest path algorithm finds the first k shortest service paths within the confined area of the time–space prism.

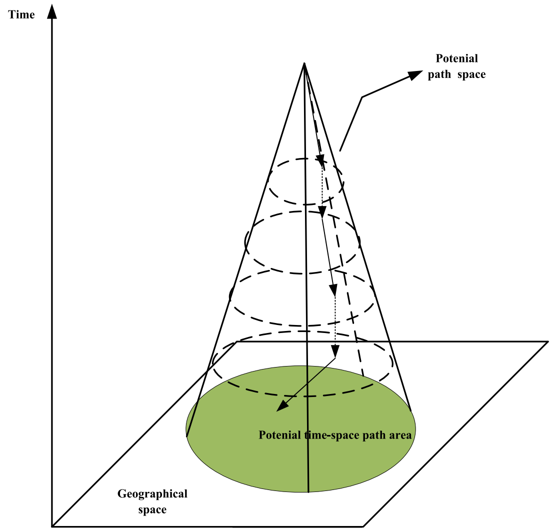


FIGURE 6. A prism time–space shortest path for door-to-door transport service.

Step 1: Dijkstra's algorithm is used to search for the shortest path of a freight demand. If the shortest path time does not meet the time limit requirements of delivery, then the alternative freight demand path is added to a superpath (a virtual path is set to avoid the effect of the algorithm when there is no solution).

Step 2: All nodes of the shortest path are traversed from the start to end points, and the tabu search information is added to the arc segment of each adjacent node in turn. After each tabu has been traversed, Dijkstra's algorithm is used to find the current shortest path as the alternative shortest path to be added to the alternative set. After traversal, the shortest path in alternative set is calculated. If the shortest path time meets the freight demand's delivery time limit, it is a secondary shortest path; otherwise, the number of feasible paths is not equal to k.

Step 3: Clear alternate set using the same method as in Step 2 to traverse the secondary shortest paths; the current shortest path obtained after each tabu search is added to alternative set as the alternative shortest path. The third-shortest path is selected from alternative set under cargo delivery time limits when the traverse is completed. If it is not satisfied, the number of feasible paths is not equal to k; finish the current k shortest paths search. Repeat Step 3 until all k shortest path sets of freight demand are found.

E. SOLUTION APPROACH FRAMEWORK

We propose a PDRFD solution approach framework (PDRFD-SAF) based on cloud resource and big data. The PDRFD-SAF combines with the MIPM-PDRFD mathematical model and the methods described earlier in this section. The core contribution of this study is also based on this PDRFD-SAF.

The PDRFD-SAF includes data reading, data storage, data cleaning, data analysis, and operations research (OR) models. The OR model can be solved exactly using ILOG Cplex software [37]. Because MIPM-PDRFD is the core of cloud-based PDRFD-SAF, the PDRFD-SAF result is exact. A flow chart of the PDRFD solution approach framework is shown in Fig. 7.

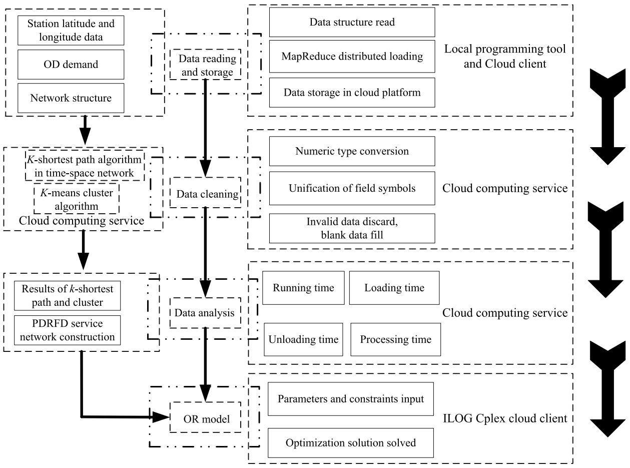


FIGURE 7. Flow chart of the PDRFD solution approach framework.

Step 1 (Data Reading): The structure of the original dataset is read into the local programming software, e.g., Java software development tools, and retrieves row and column delimiter information. For the OR model proposed herein, input data including stations (including highway demands), latitude and longitude, and their OD demand to acquire a physical network structure.

TABLE 4. Latitude and longitude of railway freight station.

Identifier	Station	Longitude and latitude	Distance to previous station (km)
A	Wuhan North Station	114.323981,30.782058	0
B	Xianning Station	114.295448,29.885403	100
C	Puqi Station	113.883741,29.705143	45
D	Linxiang Station	113.490515,29.480848	45
E	Yueyang North Station	113.165584,29.404829	33
F	Xianing Station	113.018618,28.179477	120
G	Zhuzhou North Station	113.137636,27.880593	56
H	Hengyang North Station	112.705173,26.936628	115
I	Leiyang Station	112.829235,26.410085	57
J	Lechang Station	113.363767,25.135103	152
K	Maba Station	113.619791,24.688856	56
L	Jiangcun Station	113.229125,23.309613	165
M	Dalang Station	113.240972,23.243021	8

TABLE 5. Latitude and longitude of partial highway demand points.

Highway demand number	Freight demand source	Longitude and latitude	Transfer railway station	Distance to transfer railway station (km)
1	Wuhan Light Industry Machinery Factory	114.220745,30.600838	Wuhan North Station	4.5
2	Peace Auto Industry Park	114.396911,30.605368		14.0
3	Jiahong International Transport Agency Company of Wuhan	114.281703,30.594435		6.4
4	Hubei Xianning WinTop Car	114.320959,29.877857	Xianning Station	2.6
5	Lianchuang Fangyuan Logistics	114.326719,29.894049		3.2
6	Deppon Logistics	114.270467,29.883663		2.5
7	Darunfa Centre Logistics Centre	113.85544,29.732459	Puqi Station	4.1
8	Hubei Jiuda Machinery Manufacture Limited Company	113.890225,29.694538		1.3
9	Hubei Cibi Desulfurizer of Steel Limited Company	113.759056,29.548745		21.0
10	Chengpeng Agriculture Machinery Sale Limited Company	113.480059,29.478384	Linxiang Station	1.0
11	Linxiang International Car Trade City	113.502401,29.478499		1.5
12	Dongxiang Grease Chemical Limited Company	113.521703,29.48595		3.5
13	Yueyang Paper Mill	113.165818,29.448936	Yueyang North Station	4.9
14	Shenkai Logistics Park	113.258625,29.478738		12.3
15	Yueyang Xinshen Precision Machinery Equipment Limited Company	113.180168,29.379669		3.0
16	Hunan Jianlang Medicine Limited Company	113.148598,29.383893		3.0
17	Yueyang Jingming International Freight Transport Agency Company	113.136444,29.378495		4.0
18	Hengguang International Logistics Park	113.088001,28.31618	Xianing Station	14.0
19	Changsha Minghua Machinery Limited Company	113.017061,28.301323		8.2
20	Wanwei Changsha Wangcheng Logistics Park	112.855558,28.315907		10.0
21	Huann Changhao International Freight Transport Agency Company	112.981819,28.201381		17.5

Step 2 (Data Storage): A new data table structure is built in the cloud database, which corresponds to the data structure in Step 1; big data read by the cloud database is completed using a MapReduce distributed cloud client upload function. Thus, big data is stored in the cloud database.

TABLE 6. Model parameters for big data calculation.

Identifier	Station	Average loading time (min)	Average processing time	Average unloading time	Distance to previous station	Section running time at 80km/h (min)	Section running time at 120 km/h (min)	Section running time at 160 km/h (min)
A	Wuhan North Station	918.25	537.77	988.70	0.00	0.00	0.00	0.00
B	Xianning Station	778.63	130.00	609.13	100.00	75.00	50.00	37.50
C	Puqi Station	766.53	298.00	711.39	45.00	33.75	22.5	16.88
D	Linxiang Station	305.58	160.00	1588.00	45.30	33.975	22.65	16.99
E	Yueyang North Station	984.24	407.00	376.82	32.70	24.525	16.35	12.26
F	Xianing Station	880.33	314.00	455.95	120.00	90.00	60.00	45.00
G	Zhuzhou North Station	953.34	413.00	403.52	56.00	42.00	28.00	21.00
H	Hengyang North Station	1342.98	338.00	572.76	115.00	86.25	57.50	43.13
I	Leiyang Station	1169.46	139.00	863.57	57.40	43.05	28.70	21.53
J	Lechang Station	663.69	118.00	844.92	152.00	114.00	76.00	57.00
K	Maba Station	1275.28	145.00	229.87	56.10	42.075	28.05	21.04
L	Jiangcun Station	1603.29	569.57	1329.19	165.00	123.75	82.50	61.88
M	Dalang Station	687.31	260.00	250.77	7.50	5.63	3.75	2.81

TABLE 7. Partial highway demand volume.

Start highway demand point	End highway demand point	Transfer railway station identifier	Number of cars	Time requirement (/h)
1	15	E	30	24
1	24	G	9	48
1	32	J	6	24
1	40	L	20	48
1	41	L	10	48
2	14	E	18	48
2	20	F	43	48
2	29	I	36	48
2	32	J	24	48
2	41	L	60	48

Step 3 (Data Cleaning): Data cleaning can be completed on the cloud platform, including values of types String and Integer, the number of columns in the numerical calculation, and irregular character unification formats. The k-shortest path algorithm in the time–space network and k-means clustering can be programmed with cloud computing resources to speed up the algorithm; their results are also key OR model parameters.

Step 4 (Data Analysis): A cleaned dataset was analyzed via SQL operations in the cloud platform, which can be calculated from railway freight car trajectory and loading–unloading records, section running time, station processing time, station loading and unloading time, etc.

Step 5: The results of Steps 3 and 4 are input to the OR model and the ILOG CPLEX Cloud client solves the OR model to get the final door-to-door railway freight product.

TABLE 8. Result of trunk railway product.

Product type	Serial number	Trunk service plan	Train	Product type	Serial number	Trunk service plan	Train
Fixed Product	1	M-A	3				
	2	L-H-G-E-A	8				
Express product	1	A-H-L	1	Express product	14	L-H-A	1
	3	A-J-L	4		15	L-J-A	5
	4	A-G-L	2		18	L-G-A	4
	6	A-M	2		21	L-E-A	6
	7	A-E	3		22	M-A	2
	9	A-H	1		24	M-G-A	5
Ordinary Product	1	M-K-H	8	Ordinary Product	29	A-B-C-D	12
	3	M-H-B-A	6		31	A-C-H-L	11
	4	M-G-D-C	10		33	A-D-E-M	6
	5	M-H-G-E-A	5		34	A-G-H-M	7
	8	K-H-E-A	15		37	B-D-F-M	3
	9	K-H-G	3		38	B-E-H-L-M	7
	11	K-E-B	7		42	E-G-H-L	10
	13	G-F-E-D	2		45	E-F-I-J	2
	14	H-G	5		48	F-H-I-K-L	15
	18	I-H	1		50	G-H-K-M	14
	21	L-F-D-A	15		52	I-J-K-M	9
	22	L-I-D-B-A	11		54	J-K-L-M	6
	23	L-K-I	8				

V. CASE STUDY AND NUMERICAL EXPERIMENT

A. CASE STUDY

To validate the benefits derived from the proposed model, we examined the Beijing–Guangzhou railway corridor. One of the most important transport corridors in China, it starts at Fengtai Station in the Beijing Railway Bureau and ends in Guangzhou West Station via 6 provinces and 218 freight stations. It was built for freight transportation operations such as freight arrival, delivery, and transfer, promoting the economy and physical distribution in economic development zones along the corridor, including Beijing–Tianjin–Hebei, the Central Plains Economic Zone, the Yangtze River Economic Belt, the Wuhan metropolitan area, and the Greater Changsha Metropolitan and Zhujiang Delta regions (see Fig. 8).

This study mainly analyzes a 952-meter-long segment from Wuhan North Station to Dalang Station, which has a large demand for bulk cargo and value-added commodities, which is a representative example for our search. Along the segment, there are four special marshalling stations, Wuhan North, Zhuzhou North, Hengyang North, and Jiangcun. In this case study, we used big data techniques to analyze 3,280,803 pieces of arrival and departure information, as well as 773,659 pieces of loading and unloading information; we cover the period from May 18th to June 15th, 2017. Thirteen freight stations with higher demand and 44 highway demand points are analyzed in detail. The result is carried out by the Ali cloud platform, ILOG Cplex cloud platform, and ILOG Cplex on a Windows 7 PC with a 2.20 GHz CPU and 4 GB RAM.

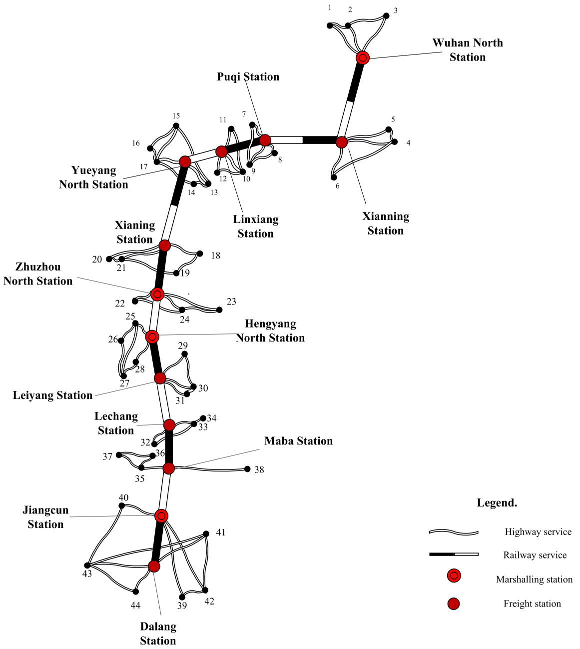


FIGURE 8. Network for Wuhan North Station to Dalang Station.

TABLE 9. Result of branch highway product.

Product type	Serial number	Branch service plan	Car	Product type	Serial number	Branch service plan	Car
Fixed Product	1	A-1-2-3-A	30	Fixed Product	8	H-25-26-27-28-H	25
	2	B-6-4-5-B	20		9	I-29-30-31-I	15
	3	C-7-9-8-C	18		10	J-32-33-35-J	16
	4	D-11-10-12	15		11	K-36-37-35-38-K	18
	5	E-13-14-17-16-15-E	28		12	L-39-42-41-40-L	35
	6	F-18-19-21-20-F	30		13	M-44-43-M	30
	7	G-23-24-22-G	30				
Express product	1	A-1-A	10	Express product	23	G-23-G	6
	2	A-2-A	7		24	G-24-G	8
	3	A-3-A	3		25	H-25-H	5
	4	B-4-B	2		26	H-26-H	2
	5	B-5-B	1		27	H-27-H	4
	6	B-6-B	2		28	H-28-H	1
	7	C-7-C	3		29	I-29-I	1
	8	C-8-C	2		30	I-30-I	1
	9	C-9-C	1		31	I-31-I	1
	10	D-10-D	1		32	J-32-J	2
	11	D-11-D	1		33	J-33-J	2
	12	D-12-D	2		34	J-34-J	1
	13	E-13-E	7		35	K-35-K	2
	14	E-14-E	6		36	K-36-K	1
	15	E-15-E	9		37	K-37-K	3
	16	E-16-E	7		38	K-38-K	2
	17	E-17-E	4		39	L-39-L	8
	18	F-18-F	10		40	L-40-L	10
	19	F-19-F	8		41	L-41-L	7
	20	F-20-F	6		42	L-42-L	5
	21	F-21-F	3		43	M-43-M	12
	22	G-22-G	3		44	M-44-M	8
Ordinary product	1	A-1-2-A	72	Ordinary product	110	G-23-24-G	37
	2	A-2-3-A	65		112	G-22-24-G	41
	8	B-4-5-B	28		119	H-25-26-H	35
	9	B-4-6-B	27		123	H-28-27-H	40
	14	C-7-8-C	27		139	I-29-30-I	20
	16	C-8-9-C	29		140	I-30-31-I	25
	20	D-10-12-D	22		145	J-33-34-J	21
	21	D-11-12-D	25		152	K-36-37-35-K	22
	30	E-13-14-E	36		158	K-35-38-K	26
	32	E-15-16-E	42		175	L-39-40-L	68
	48	E-16-17-E	41		176	L-41-42-L	65
	88	F-18-19-F	43		197	M-43-44-M	46
	92	F-20-21-F	48				

The decision-making period was 48 h; other parameters in the proposed model are shown in Tables 4 to 7. The case study was solved with the approach proposed, as shown in Tables 8 to 9. The transfer service network construction of the door-to-door freight product with K-means clustering algorithm is also shown in Fig. 8.

Ordinary products were the most common in both railways and highways. Fixed products needed 11 trains, express products needed 36, and ordinary products needs 198. Fixed products needed 310 cars, express products needed 190 cars, and ordinary products needed 951 cars in highway service. Wuhan North Station, Jiangcun Station, and Dalang Station

TABLE 10. Comparison for different network scales in two-way computational time.

Number of highway demand points	Number of railway stations	ILOG Cplex Cloud Client		Local ILOG Cplex	
		Objective value (Yuan)	Running time (/s)	Objective value (Yuan)	Running time (/s)
44	13	1435786	20	1435786	69
80	20	3704235	36	3704235	128
120	25	7240830	71	7240830	233
200	35	16530259	108	16530259	587
320	45	37837870	143	37837870	1012

TABLE 11. Efficiency comparison of the cloud computing platform and traditional database.

SQL sentence	Record	SQL Server (/s)	ODPS (/s)
select * from zmzd_0606_trim	7903	1	4
select * from df0611_0615_jgx where fj ='P'	44599	99	8
select * from df0611_0615_jgx	458572	270	10
select * from df0518_0610_trim_time	38996011	503	10

were the top three stations in processing railway cargo and transferring it to highway demand points. Therefore, it was necessary to strengthen the processing capacity and dispatching control of these stations.

B. NUMERICAL EXPERIMENT

A numerical experiment of a different service network scale was applied to test the proposed model. The model was implemented in the ILOG Cplex Cloud Client and ILOG Cplex on a Windows 7 PC with a 2.20 GHz CPU, and 4 GB RAM. As shown in Table 9, when the number of highway demand points was 320, the number of railway stations was 45 and the runtime was 143 s. Fewer than 50 freight stations in China have a larger volume, so ILOG Cplex Cloud Client is acceptable for solving the PDRFD problem. Cloud resources have an enormous advantage over local resources in solving an OR model with complex variables and constraints (see Table 10).

We carried out a computing test between the Ali cloud ODPS and a traditional SQL Server 2008 database to compare their efficiency, processing big data of railway freight train trajectories and loading–unloading records on a Windows 7 PC with a 2.20 GHz CPU, and 4 GB RAM. Overall, computing rates of the Ali cloud ODPS were better than those of the traditional Server SQL database. Its computing advantages were more obvious with increasing SQL sentence complexity or dataset size.

In the first computing test, Ali cloud ODPS took 9 s, but SQL Server only needed 1 s because the local computer submits an SQL request to the Ali cloud computing platform, which requires transmission and initialization times; additionally, Ali splits the ODPS into a distributed task. MapReduce initialization is relatively slow. even if the complexity of the SQL sentence is low. However, advantages of the cloud computing technology compared with the traditional database are presented in Table 11.

A cloud computing mode that does not occupy local resources also allows users to better allocate resources, reduce costs, and improve efficiency. Compared to traditional database technology, cloud computing is very adaptable and has good application prospects.

VI. CONCLUSION

We established an MIPM-PDRFD using a time–space service network and large railway freight car trajectory datasets. We proposed an innovative approach framework for PDRFD and presented a cloud-resource-based solution processing method. Our analysis led to three main conclusions.

(1) It is necessary to design and develop door-to-door railway freight products, which can improve consignor satisfaction and the efficiency of logistics transport.

(2) The cloud-based approach proposed in this paper is a new attempt to solve the PDRFD problem; the case study shows that cloud resource and big data technology change optimization methods and thinking modes.

(3) Ordinary products of both trunk railway services and branch highway services are suitable for cargo with less time sensitivity and large volumes, express products are suitable for high value-added cargo, because they need less delivery time and high-quality logistics service. Overall, ordinary products comprise the largest proportion of door-to-door product.

A limitation of this study is that the volume of railway freight car trajectory data may not be big enough to calculate the more precise real parameters of the MIPM-PDRFD. We therefore aim to obtain more data via cooperation with railway enterprises. In addition, we plan to explore several future research directions:

(1) An integrated PDRFD system that has been solved in the proposed approach framework should be developed.

(2) The complexity of the time–space service network should be analyzed, as simplification methods are expected to be challenging when the number of nodes and arcs increases.

(3) PDRFD problems can also be considered in economic factor optimization.

(4) With the continuous reform of railway operation, building a suitable door-to-door freight product transfer service network should be researched.

(5) More efficient and suitable railway door-to-door freight operation modes should be developed.

REFERENCES

[1] *Ministry of Transport of the People's Republic of China*. Accessed: Nov. 2017. [Online]. Available: http://www.mot.gov.cn/shuju/

[2] *China Railway Company*. Accessed: Nov. 2017. [Online]. Available: http://www.china-railway.com.cn/cpyfw/ysdt/

[3] T. G. Crainic and G. Laporte, "Planning models for freight transportation," *Eur. J. Oper. Res.*, vol. 97, no. 3, pp. 409–438, Mar. 1997.

[4] D. M. Z. Islam, "Barriers to and enablers for European rail freight transport for integrated door-to-door logistics service. Part 1: Barriers to multimodal rail freight transport," *Transp. Problems*, vol. 9, no. 3, pp. 43–56, Mar. 2014.

[5] D. M. Z. Islam, "Barriers to and enablers for European rail freight transport for integrated door-to-door logistics service. Part 2: Enablers for multimodal rail freight transport," *Transp. Problems*, vol. 9, no. 4, pp. 5–13, Apr. 2014.

[6] T. L. Magnanti and R. T. Wong, "Network design and transportation planning: Models and algorithms," *Transp. Sci.*, vol. 18, no. 1, pp. 1–55, 1984.

[7] T. G. Crainic and J.-M. Rousseau, "Multicommodity, multimode freight transportation: A general modeling and algorithmic framework for the service network design problem," *Transp. Res. B, Methodol.*, vol. 20, no. 3, pp. 225–242, Jun. 1986.

[8] M. Minoux, "Networks synthesis and optimum network design problems: Models, solution methods and applications," *Network*, vol. 19, no. 3, pp. 313–360, May 1989.

[9] L. K. Nozick and E. K. Morlok, "A model for medium-term operations planning in an intermodal rail-truck service," *Transp. Res. A, Policy Pract.*, vol. 31, no. 2, pp. 91–107, Mar. 1997.

[10] A. Ziliaskopoulos and W. Wardell, "An intermodal optimum path algorithm for multimodal networks with dynamic arc travel times and switching delays," *Eur. J. Oper. Res.*, vol. 125, no. 3, pp. 486–502, Sep. 2000.

[11] S.-J. Jeong, C.-G. Lee, and J. H. Bookbinder, "The European freight railway system as a hub-and-spoke network," *Transp. Res. A, Policy Pract.*, vol. 41, no. 6, pp. 523–536, Jun. 2007.

[12] A. Ceselli, M. Gatto, M. E. Lübbecke, M. Nunkesser, and H. Schilling, "Optimizing the cargo express service of Swiss federal railways," *Transp. Sci.*, vol. 42, no. 4, pp. 450–465, Nov. 2008.

[13] V. Cacchiani, A. Caprara, and P. Toth, "Scheduling extra freight trains on railway networks," *Transp. Res. B, Methodol.*, vol. 44, no. 2, pp. 215–231, Feb. 2010.

[14] M. Yaghinia and R. Akhavan, "Multicommodity network design problem in rail freight transportation planning," *Procedia-Social Behav. Sci.*, vol. 43, pp. 728–739, Jan. 2012.

[15] M. Yaghini, M. Rahbar, and M. Karimi, "A hybrid simulated annealing and column generation approach for capacitated multicommodity network design," *J. Oper. Res. Soc.*, vol. 64, no. 7, pp. 1010–1020, Jul. 2013.

[16] M. Yaghini and Z. Khandaghabadi, "A hybrid metaheuristic algorithm for dynamic rail car fleet sizing problem," *Appl. Math. Model.*, vol. 37, no. 6, pp. 4127–4138, Jun. 2013.

[17] C.-C. Lin and S.-H. Chen, "The hierarchical network design problem for time-definite express common carriers," *Transp. Res. B, Methodol.*, vol. 38, no. 3, pp. 271–283, Mar. 2004.

[18] T. G. Crainic, "Service network design in freight transportation," *Eur. J. Oper. Res.*, vol. 122, no. 2, pp. 272–288, Feb. 2000.

[19] L. C. Dall'Orto, T. G. Crainic, J. E. Leal, and W. B. Powell, "The single-node dynamic service scheduling and dispatching problem," *Eur. J. Oper. Res.*, vol. 170, no. 1, pp. 1–23, Jan. 2006.

[20] E. Zhu, T. G. Crainic, and M. Gendreau, "Scheduled service network design for freight rail transportation," *Oper. Res.*, vol. 62, no. 2, pp. 383–400, Feb. 2014.

[21] D. Alanis, J. Hu, P. Botsinis, Z. Babar, S. X. Ng, and L. Hanzo, "Quantum-assisted joint multi-objective routing and load balancing for socially-aware networks," *IEEE Access*, vol. 4, pp. 9993–10028, Jan. 2017.

[22] C. Xu, C. Gao, Z. Zhou, Z. Chang, and Y. Jia, "Social network-based content delivery in device-to-device underlay cellular networks using matching theory," *IEEE Access*, vol. 5, pp. 924–937, 2017.

[23] L. Duan, L. Tavasszy, and Q. Peng, "Freight network design with heterogeneous values of time," *Transp. Res. Procedia*, vol. 25, pp. 1144–1150, Jun. 2017.

[24] A. Fakhri and M. Ghatee, "Application of benders decomposition method in solution of a fixed-charge multicommodity network design problem avoiding congestion," *Appl. Math. Model.*, vol. 40, nos. 13–14, pp. 6468–6476, Jul. 2016.

[25] B. Lin, C. Liu, H. Wang, and R. Lin, "Modeling the railway network design problem: A novel approach to considering carbon emissions reduction," *Transp. Res. D, Transp. Environ.*, vol. 56, pp. 95–109, Oct. 2017.

[26] *Moving an Elephant: Large Scale Hadoop Data Migration at Facebook*. Accessed: Nov. 2015. [Online]. Available: http://nosql.mypopescu.com/post/8651712469

[27] Q. Shen, X. Liang, X. Shen, X. Lin, and H. Y. Luo, "Exploiting geo-distributed clouds for a E-health monitoring system with minimum service delay and privacy preservation," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 2, pp. 430–439, Mar. 2014.

[28] L. Zhang *et al.*, "Moving big data to the cloud: An online cost-minimizing approach," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 2710–2721, Dec. 2013.

[29] G. Kemp, G. Vargas-Solar, C. F. Da Silva, P. Ghodous, C. Collet, and P. P. L. Amalya, "Cloud big data application for transport," *Int. J. Agile Syst. Manage.*, vol. 9, no. 3, pp. 232–250, Nov. 2016.

[30] D. Lv, P. Wang, W. Li, and P. Chen, "OpenCL-accelerated point feature histogram and its application in railway track point cloud data processing," in *Proc. Int. Conf. Inform. Control, Autom. Robot.*, Lisbon, Portugal, 2016, pp. 433–438.

[31] M. Azizian, S. Cherkaoui, and A. Hafid, "An optimized flow allocation in vehicular cloud," *IEEE Access*, vol. 4, pp. 6766–6779, Nov. 2017.

[32] X. Luo, Y. Lv, R. Li, and Y. Chen, "Web service QoS prediction based on adaptive dynamic programming using fuzzy neural networks for cloud services," *IEEE Access*, vol. 3, pp. 2260–2269, Nov. 2015.

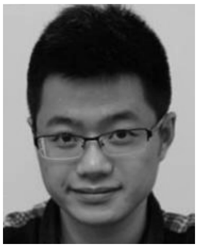
[33] L. A. Tawalbeh, R. Mehmood, B. Elhadj, and H. Song, "Mobile cloud computing model and big data analysis for healthcare applications," *IEEE Access*, vol. 4, pp. 6171–6180, Sep. 2016.

[34] J. Y. Yen, "Finding the K shortest loopless paths in a network," *Manage. Sci.*, vol. 17, pp. 712–716, Jul. 1971.

[35] Y.-L. Chen and H.-H. Yang, "Finding the first K shortest paths in a time-window network," *Comput. Oper. Res.*, vol. 31, no. 4, pp. 499–513, Apr. 2004.

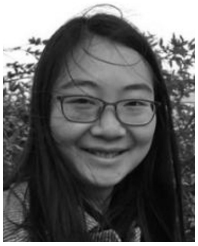
[36] T. M. White, *Hadoop: The Definitive Guide*, vol. 215, no. 11. Sebastopol, CA, USA: O'Reilly Media, Nov. 2009, pp. 1–4.

[37] B. Elbenani and J. A. Ferland, "An exact method for solving the manufacturing cell formation problem," *Int. J. Prod. Res.*, vol. 50, no. 15, pp. 4038–4045, Aug. 2012.



WEICHUAN YIN is currently pursuing the Ph.D. degree with the School of Traffic and Transportation, Beijing Jiaotong University. His research interests include the optimization and modeling of railway freight transportation, and the big data application of railway transportation.

SHIWEI HE is currently a Professor with the School of Traffic and Transportation, Beijing Jiaotong University. His research interests include railway transport organization theory and technology.



YINGQUN ZHANG is currently pursuing the Ph.D. degree with the School of Traffic and Transportation, Beijing Jiaotong University. Her research interests include the optimization and modeling of urban transportation.

JI HOU is currently pursuing the master's degree with the School of Traffic and Transportation, Beijing Jiaotong University. His research interests include the optimization and modeling of railway freight transportation.

• • •