Proactive Two-Level Dynamic Distribution Routing Optimization Based on Historical Data

Xianlong Ge 1,2, Guiqin Xue 1, and Pengzhe Wen 1

1 School of Economics and Management, Chongqing Jiaotong University, Chongqing 400074, China
2 Key Laboratory of Intelligent Logistics Network, Chongqing 400074, China

Correspondence should be addressed to Xianlong Ge; gexianlong@cqjtu.edu.cn and Guiqin Xue; xuegqmail@126.com

Received 27 April 2018; Revised 24 September 2018; Accepted 30 October 2018; Published 21 November 2018

Abstract

In view of the dynamic dispersion of e-commerce logistics demand, this paper uses the historical distribution data of logistics companies to study data-driven proactive vehicle routing optimization. First, based on the classic 2E-VRP problem, a single-node/multistage 2E-VRP mathematical model is constructed. Then, a framework for solving the proactive vehicle routing optimization problem is proposed in combination with the characteristics of the proposed model, including four modules: data-driven demand forecasting methods, customer clustering methods, proactive demand quotas and replenishment strategies, and vehicle routing optimization procedure. The significant feature of the proposed solution framework is that the response to dynamic customers is proactive rather than passive. The solution is applied to the distribution practice of a large logistics company in Chongqing. The results show that the proposed method has better dynamic scene adaptability and customer response capabilities in traffic limit.

1. Introduction

With the rise of cloud computing and big data technologies, the role of data-driven decision-making in commercial operations has become increasingly prominent and provides technical guidance to traditional distribution enterprises from a practical perspective. At the same time, the continuous improvement of the Internet infrastructure also provides a reliable source of data for business decision-making, but massive data has also led to data disasters to a certain extent. On the one hand, diversification of data acquisition means that distribution enterprises can obtain customer data very conveniently and tap customer potential based on the data obtained to change or reshape customer’s demand behavior; on the other hand, even in a short period of time, the amount of data gathered by large enterprises is staggering, which poses a serious challenge for data cleaning and data mining. For example, the gradual opening of some cloud computing platforms and tools such as Amazon Cloud and Alibaba Cloud enables SMEs to take advantage of the powerful computing power of these cloud computing platforms to rapidly respond to changing market demands and shorten the lead time of ordering so that the inconvenience of their customers’ demand for goods is minimized. However, in the distribution process, enterprises can achieve proactive demand response if we can tap out the potential law of dynamic demand, optimizing the allocation of existing logistics resources.

Vehicle routing problem is one of the most extensive and in-depth combinatorial optimization problems. In recent years, the development of e-commerce has led to the trend of decentralization of demand dynamics, which makes the solution of vehicle routing problems more complicated (see [1–4]). In addition, the vehicle restriction strategy is prevalent in major cities around the world. Therefore, in the current urban distribution process, vehicles with heavy loads are often used to transport goods to the urban edge areas and transport to smaller cities by vehicles with smaller loads to meet the needs of people’s daily lives, which led to the research on the two-level dynamic vehicle routing problem [5, 6]. However, this strategy often neglects some potential dynamic customers and lacks practical consideration.

In this paper, we propose a two-echelon proactive vehicle routing problem based on the proactive scheduling method, which is one of the expanding problems of dynamic vehicle
routing problem [7–9]. 2E-PVRP generates a priori knowledge of dynamic customer changes based on the customer data obtained by the delivery companies and uses the prior knowledge to process the dynamic demand information during the distribution process to form a robust vehicle scheduling solution. Due to the complexity of the questions raised, this paper divides the solution into four modules: data driven demand forecasting methods, customer clustering methods, proactive demand quotas and replenishment strategies, and vehicle routing optimization programs. The data driven demand forecasting methods mainly excludes the gross error points in the original data, selects the appropriate measure dimension to evaluate the change of customer attributes, and then predicts the dynamic demand trend. The customer clustering methods will determine the customers who meet a certain distribution into the same cluster mainly based on the dynamic customer forecasting results and static customer data. The proactive demand quotas and replenishment strategies mainly evaluate the statistical characteristics of customer needs in the cluster and allocate the distribution quotas of each stage reasonably. The vehicle routing optimization program mainly adopts the exact or approximate path optimization algorithm to generate the optimal delivery path plan for each cluster mainly based on the dynamic demand information.

The test results show that, in the case of small number of customers, the demand extreme (suddenly a large or small value) is unpredictable and the demand dynamics affect the demand forecasting effect. In this paper, we propose a proactive demand quota method to solve the failure of demand forecasting by using the strategy of multibatch and small-batch delivery. That is to say, a complete service cycle is divided into initial delivery stage and multiple replenishment stages. If there is an extreme demand and a failure forecasting in the initial stage, the shortfall of demand will be supplemented through the late replenishment phase to avoid the serious consequences of the forecast failure.

After acquiring the prior knowledge of dynamic clients by using data-driven forecasting methods, they need to be merged with the original delivery customers to determine the customer base to be served during the full-service period. The customer clustering method is to classify customers with similar attributes into the same type of clusters according to a certain rule and classify the clients with large differences in attributes into different clusters so as to realize the classification of customer behaviors. Due to the effective customer clustering scheme in reducing logistics costs and improving the quality of service, many scholars have adopted the strategy of “planning route after clustering” to optimize the vehicle routing. The authors in [19–25] analyzed the problem of service area allocation in the dynamic vehicle problem with stochastic demand. Ferrucci [26] assumed that the dynamic customer demand obeys Poisson distribution and divided one cycle of distribution activities into several smaller grids areas and calculated the Poisson distribution parameters of each grid area in different distribution stages. Since the Poisson distribution parameters in the single grid area are small and cannot meet the delivery requirements, the author uses the maximum roaming radius and roaming time criteria to cluster the service subareas and optimize the distribution network path. The results show that the proactive scheduling method can effectively improve customer satisfaction.

Ferrucci [27] expanded the space-time Poisson distribution model. Wang [28] discussed the customer clustering problem in 2E-VRP, assumed that the location of Tier 1 and Tier 2 facilities is determined, and the customers are clustered into a Tier 2 facility according to the set evaluation criteria. On this basis, this paper determines the waiting customer group based on the data-driven forecasting method and clusters the customer group to determine the service customer group of

The structure of the article is as follows: In the second part, the paper introduces the development of the vehicle routing problem and its extension problem. The third part is the mathematical model of the problem. The fourth part is the realization process of the solution framework. The fifth part is the use of the enterprise empirical research data, the sixth part is obtaining the management inspiration, and the seventh part is the summary and future work.

2. Literature Review

At present, many scholars at home and abroad have made a great deal of research on the 2E-VRP problem. Among them, demand forecasting, customer clustering method, and vehicle routing solving algorithm in dynamic vehicle scheduling are closely related to this paper. Dynamic customer processing in dynamic vehicle routing has always been the focus of research. Wood [10] uses the three indicators of order quantity, order variation rate, and order frequency as the measure and uses the piecewise linear model to evaluate the reliability of the demand forecast based on the data. The test results show that, in the case of small number of orders, small differences in order quantity, and lower order frequency, the use of point-of-sale real-time shared data can achieve better forecasting results, which is provided in this paper by using multidimensional customer attribute data to predict dynamic customer demand. Data-driven demand forecasting methods, as an effective forecasting strategy, are also widely used by scholars. Thomas [11] designed a real-time heuristic algorithm to predict future customer locations and the probability of occurrence of demand, resulting in a significant reduction in customer service latency. Lima [12] takes customer needs and businesses on their own to leverage historical data and online data monitoring to anticipate customer needs and improve customer responsiveness. Cabello [13] designed low cost forecasting algorithm for uncertain demand to manage bank cash flow. Jan [14] and Ma [15] anticipate changes in customer demand in advance by considering demand forecasting and inventory control. Sha [16] anticipates clients’ expected waiting times significantly using historical demand data to proactively anticipate spare part demand. Willemai [17] and Porras [18] argue that, in the case of limited demand data sets, the demand extreme (suddenly a large or small value) is unpredictable and the demand dynamics affect the demand forecasting effect. In this paper, we propose a proactive demand quota method to solve the failure of demand forecasting by using the strategy of multibatch and small-batch delivery. That is to say, a complete service cycle is divided into initial delivery stage and multiple replenishment stages. If there is an extreme demand and a failure forecasting in the initial stage, the shortfall of demand will be supplemented through the late replenishment phase to avoid the serious consequences of the forecast failure.

After acquiring the prior knowledge of dynamic clients by using data-driven forecasting methods, they need to be merged with the original delivery customers to determine the customer base to be served during the full-service period. The customer clustering method is to classify customers with similar attributes into the same type of clusters according to a certain rule and classify the clients with large differences in attributes into different clusters so as to realize the classification of customer behaviors. Due to the effective customer clustering scheme in reducing logistics costs and improving the quality of service, many scholars have adopted the strategy of “planning route after clustering” to optimize the vehicle routing. The authors in [19–25] analyzed the problem of service area allocation in the dynamic vehicle problem with stochastic demand. Ferrucci [26] assumed that the dynamic customer demand obeys Poisson distribution and divided one cycle of distribution activities into several smaller grids areas and calculated the Poisson distribution parameters of each grid area in different distribution stages. Since the Poisson distribution parameters in the single grid area are small and cannot meet the delivery requirements, the author uses the maximum roaming radius and roaming time criteria to cluster the service subareas and optimize the distribution network path. The results show that the proactive scheduling method can effectively improve customer satisfaction.

Ferrucci [27] expanded the space-time Poisson distribution model. Wang [28] discussed the customer clustering problem in 2E-VRP, assumed that the location of Tier 1 and Tier 2 facilities is determined, and the customers are clustered into a Tier 2 facility according to the set evaluation criteria. On this basis, this paper determines the waiting customer group based on the data-driven forecasting method and clusters the customer group to determine the service customer group of
each secondary facility, so as to optimize the delivery route of the entire distribution system.

The proactive two-level dynamic routing optimization problem studied in this paper belongs to the crossover problem of dynamic vehicle routing problem and secondary network vehicle routing problem. The above three problems are unable to get a reasonable solution in polynomial time, which is a typical NP-hard problem. However, the problem in this paper is more difficult to solve than the above problem. For the vehicle routing optimization problem, the main problem is whether to use the exact algorithm or the approximate algorithm. For an introduction of VRP’s exact solution algorithm, see [29–31]. Santos [32] proposed branch and bound algorithm and its solution to the weighted average method proposed by Santos [33] and Sahraeian, an extended version of 2E-VRP that includes two optimization goals, environmental costs and service satisfaction. However, the exact solution to the algorithm has poor adaptability to the customer scale, so the approximate solution algorithm is more favored by scholars when the problem scale is larger. Commonly used approximate solution algorithms include large neighborhood search algorithm [19, 20, 34], genetic algorithm [35], tabu search algorithm [36], simulated annealing algorithm, and ant colony algorithm. Alonso [37] developed a tabu search algorithm to solve the multivehicle routing problem. Maischberger [38] added perturbation mechanism to tabu search algorithm and used multicore parallel technology, both ensuring the diversity and optimization ability of algorithms and improving the efficiency of solving. Silvestrin [39] embedded tabu search algorithms in iterative local search algorithms for better computational results. Inspired by this, considering that genetic algorithm has a strong global search capability, premature algorithm may exist. Tabu search algorithm has better local search ability but its solution performance is greatly affected by initial solution. Therefore, the genetic algorithm is embedded into the tabu search algorithm to get the optimal solution of the proposed 2E-PVRP.

The above research provides meaningful solutions for vehicle optimal scheduling in dynamic scenarios from different perspectives, but there is still room for further improvement. First, the traditional dynamic VRP considers the first-level distribution network system from the vehicle yard to the customer. In the dynamic distribution scenario, there is a long-term customer response distance, which is difficult to adapt to the demanding requirements of customers; second, the research on urban distribution scheduling problem is mostly reactive scheduling and less use of prior knowledge contained in historical data. By refining the service area to develop a differentiated service strategy that takes into account the differentiated characteristics of regional demand, the distribution process is less systematically optimized from the perspective of proactive/reactive integration optimization.

In order to solve the problem of long response distance in the traditional reactive dispatching mode of urban logistics distribution, this paper uses the basic model of two-level vehicle routing problem to construct a proactive two-level dispatching mathematical model and its solution framework based on historical data. The aim of this step is to realize the efficiency of urban logistics distribution service and enhance the ability of customer demand response in dynamic scenarios. The contribution is as follows: taking the two-level vehicle routing problem as the basic model, the distribution cycle is divided into initial phase and multiple dynamic replenishment phase; the customers with the same attributes are spatially clustered by using the customer spatial allocation feature, and the dynamic demand forecasting is performed according to the historical performance of the demand within each cluster. At the same time, differentiated services and demand quota strategies are designed to optimize distribution costs under different regional differentiation. Due to the two-level network design, when a dynamic customer is generated, it can be served by the transfer station of the clustering area, which can reduce the spatial distance of the dynamic demand response and improve the flexibility of distribution scheduling. Since the required data comes from real distribution enterprise, the proposed method has certain practical significance.

3. Mathematical Model

3.1. Terms and Definitions. (1) The two-level network refers to a distribution network consisting of a hub-type distribution center to a distribution center (first-level network) and a distribution-type logistics center to a customer (second-level distribution network). Single-stage delivery divides the delivery cycle into multiple time intervals, each time interval corresponding to the delivery operation process. Multistage delivery includes the delivery process for many single-stage delivery operations.

(2) Assume that the undirected graph \( G = (N, A) \) represents the distribution network, \( N = N_0 \cup N_F \cup N_S \) represents the node set, and is composed of three types of nodes: hub-type distribution center \( N_0 \), distribution-type logistics center \( N_F \), and customer \( N_S \). \( A = A_F \cup A_S \) represents the set of arcs, where \( A_F \) represents the hub type of a set of arcs between the logistics center \( N_0 \) and the distribution-type logistics center \( N_F \), \( A_S \) represents a set of arcs between the distribution-type logistics center \( N_F \) and its service customer group \( N_S \). The length of the connection arc between any two nodes \( i \) and \( j \) in \( A_F \) and \( A_S \) is denoted as \( d_{ij} \), and the decision variable \( b_i \) is used to identify whether node \( i \) belongs to the primary distribution network node.

(3) Considering the problem of urban traffic limit, the same level distribution network should be used to deliver the same type of vehicles, different levels of networks should be distributed using heterogeneous vehicles, and the vehicles should not be used across levels; that is, the vehicles in the primary distribution network should be prohibited from being secondary distribution networks in the delivery of customer service. Suppose the demand of any node \( i \) is \( q_i \), and the model used in the first-class distribution network is \( K_F \), the capacity constraints of the vehicle \( kf \) is \( F_{G_F} \) and the fixed cost is \( c_G \); the second-class delivery network uses model \( K_S \), the capacity constraints of the vehicle \( ks \) are \( S_G \), and the fixed cost is \( c_G \).
(4) Given a level distribution network and nodes i, j, if vehicle $K_F$ traverses the distribution type logistics centers i and j, $x_{kfi}$ is assumed to be 1; otherwise it is 0; and the corresponding secondary distribution network is denoted by the symbol $x_{kaj}$.

(5) Let $y_{kfi}$ and $y_{kai}$ denote the ownership relationship between nodes and vehicles in the two-level distribution network, respectively. If the ownership relationship is established, the value is 1; otherwise it is 0.

Since we divide the whole distribution cycle into multiple distribution phases and each distribution phase carries out cargo allocation through the proactive demand quota strategy, we first establish a single-phase 2E-PVRP mathematical model to describe the path optimization of each distribution phase. On this basis, a multistage 2E-VRP mathematical model was established to optimize the path of the entire distribution cycle.

In the design of single-stage distribution model, following the traditional VRP research model, we do not consider the impact of distribution center inventory costs on scheduling operations, but only consider the Path-Related costs. The optimal solution of the model is to minimize the distribution operation costs in the whole distribution phase; in the process of multistage distribution, because of the introduction of distribution logistics center for proactive inventory, the optimal solution of this model is that, besides considering the minimum cost of each stage of distribution operation, the objective of minimum operation of distribution logistics center should be added.

### 3.2. Single-Stage 2E-PVRP Mathematical Model

To cope with the changing dynamic customer demand, this article divides the distribution cycle into the initial distribution phase and the multiple replenishment phases. A single distribution phase can be regarded as a static problem. Therefore, for any single distribution stage, the mathematical model of single-stage and two-level dynamic distribution problem is established with the minimum total cost of operation as objective function. The model is as follows:

#### The objective function:

\[
\min Z = \sum_{k_f=1}^{K} \sum_{i=1}^{N_f} \sum_{j=1}^{N_f} x_{kfi} c_{kfi} \\
+ \sum_{k_f=1}^{K} \sum_{i=1}^{N_f} \sum_{j=1}^{N_f} x_{kaj} c_{kaj} + \left[ \sum_{i=1}^{N} q_i \right] c_{KF} \\
+ \left[ \sum_{i=1}^{N} q_i \right] c_{QF}
\]  

#### Restrictions:

\[
\sum_{i=1}^{N} b_i y_{kfi} \leq F_Q 
\]  

\[
\sum_{i=1}^{N} b_i y_{kai} \leq S_Q 
\]  

\[
\sum_{i=1}^{N} (1-b_i) x_{kai} = \sum_{i=1}^{N} (1-b_i) x_{kaj} \quad \forall j \in N_F 
\]  

\[
\sum_{i=1}^{N} \sum_{j=1}^{N} (1-b_i) x_{kai} = \sum_{i=1}^{N} \sum_{j=1}^{N} (1-b_i) x_{kaj} 
\]  

\[
\sum_{i=1}^{N} b_i x_{kfi} = \sum_{i=1}^{N} b_i x_{kfi} 
\]  

\[
\sum_{i=1}^{N} b_i x_{kii} = \sum_{i=1}^{N} b_i x_{kii} 
\]  

\[
\sum_{i=1}^{N} (1-b_i) x_{kai} = \sum_{i=1}^{N} (1-b_i) x_{kaj} 
\]  

\[
\sum_{i=1}^{N} b_i q_i \geq \sum_{i=1}^{N} (1-b_i) q_i 
\]  

\[
x_{kaj} = 1 \Rightarrow S_Q - \sum_{i=1}^{N} y_{kai} q_i \geq q_j 
\]  

\[
d_{ij} + d_{jk} - d_{ik} > 0, \quad \forall i \neq j \neq k \in N 
\]  

\[
b_i x_{kfi} \left(1 - x_{kfi}\right) = \left(1-b_i\right) x_{kai} x_{kaj} = 0, 
\]  

\[
b_i y_{kfi} \left(1 - y_{kfi}\right) = \left(1-b_i\right) y_{kai} y_{kaj} = 0, 
\]  

\[
b_i \left(1 - b_i\right) = 0, \quad \forall i \in N 
\]  

Formulas (1) is the minimum objective function of total operating cost, which consists of three items: the first is the total distribution cost of the two-level distribution network, the second is the fixed vehicle cost of the first-level distribution network, and the third is the fixed vehicle cost of the second-level distribution network. Formulas (2) and (3) are the vehicle load constraints in the two-level distribution network; formula (4) is the access constraint between the hub-type logistics center and the distribution-type logistics center; formula (5) is the access constraint between the distribution-type logistics centers and their services; formulas (6) and (7) represent the access constraints between nodes in a two-level network; formula (8) shows that each distribution-type logistics center in a level-I distribution network can only be
3.3. Multistage 2E-PVRP Mathematical Model. T: The number of stages of a complete distribution cycle

\[
\begin{align*}
H_{QT} & : \text{Service ability of hub logistics center} \\
D_{QT}^j & : \text{Service capability of distribution logistics center} \\
Z_t & : \text{Operation cost of T in a distribution stage} \\
y_{jst} & : \text{The customer J in the distribution stage T whether it is from the vehicle KS service, then take 1, otherwise take 0}
\end{align*}
\]

### Mathematical Formulation

The multistage 2E-PVRP mathematical model with the objective of minimizing the total vehicle dispatching cost of each distribution stage and the operation cost of the distribution logistics center is established as follows:

\[
\min Z_T = cN_F + \sum_{t=1}^{T} \left( Z_t + \sum_{i=1}^{N_i} \left( H_i^1 c^h_i + U_i^1 c^u_i \right) \right)
\]

In addition to the constraints specified by the single stage 2E-PVRP issue, the following constraints need to be fulfilled:

\[
\begin{align*}
\sum_{i=1}^{N_i} D_{QT}^j & \leq H_{QT} \quad (18) \\
\sum_{t=1}^{T} \sum_{j=1}^{N_i} y_{jst} & \leq D_{QT}^j \quad \forall i \in N_F, \forall j \in N_S \quad (19) \\
1 & \leq \sum_{i=1}^{T} y_{jst} \leq T, \quad i \in N_S \cup N_F \quad (20) \\
\sum_{i=1}^{T} y_{jst} & = 1, \quad i \in N_S \quad (21) \\
\forall t & \in T, i \in N_F, \quad H_i^1 = 0 \iff U_i^1 \neq 0 \quad (22) \\
c & > 0, \quad c^h_i \geq 0, \quad c^u_i \geq 0 \quad (23)
\end{align*}
\]

Equation (17) is the minimum total cost of operation objective function in each stage; the first one is the Leasethold cost of distribution logistics center; the second one is the vehicle dispatching cost of each stage and the shortage and surplus cost of distribution logistics centers. Equation (18) is the constraint that secondary network service capability is not greater than the primary network service capability. Equation (19) indicates that each distribution type logistics center needs less than its service. Equation (20) represents the number of times each distribution type logistics center is visited throughout the entire distribution cycle (the initial delivery phase needs to access all the distribution type logistics centers). Equation (21) indicates that any customer can only be served once during the entire service period and no customer needs replenishment. Equivalents (22) denote that, at any distribution stage, the distribution logistics center I does not have both shortage and surplus. Equivalents (23) denote non-negative variable constraints.

### 4. Solution Framework

The proposed method consists of four modules: (1) Data-driven demand forecasting method: the historical performance of dynamic demand is evaluated by using deterministic linguistic value and triangular fuzzy number, and the potential dynamic customer demand is predicted based on the assessment results. (2) Customer clustering method: through the method of clustering algorithm, proactive partitioning is divided and then the distribution logistics center is determined. (3) Proactive demand quotas and replenishment strategies: analyzing the historical performance of customer demand data in each proactive subregion and determining the initial and replenishment supply quotas to timely meet customer needs. (4) Vehicle path optimization procedure: using scan operator to get the initial solution of the delivery path and embedding it into the designed tabu search algorithm to obtain the optimal scheduling scheme of distribution network at all levels.

#### 4.1. Data-Driven Demand Forecasting Methods

First if the firm evaluates the customer's historical performance from x dimensions, then the historical performance data for n clients can be represented by matrix A = (a_{ij}^1, a_{ij}^2, ..., a_{ij}^n; a_{ij}^1, a_{ij}^2, ..., a_{ij}^n; a_{ij}^1, a_{ij}^2, ..., a_{ij}^n). To ensure that the description of historical customer needs is closer to the actual application scenario, different dynamic client attributes are described by determining language values and triangular fuzzy numbers, respectively. The symbol \(x^D\) is used to identify the part of the customer demand attribute that can be accurately described, and the symbol \(x^F\) is used to identify the part of the customer demand attribute that can only be measured by vague language values. The prediction of demand attributes is determined by expert scoring method. Assuming that the predicted value of the attribute dimension \(x\) of the custom \(i\) is \(p_i^x\) and the measured mean value of the attribute dimension in customers within multiple delivery cycles is \(m_i^x\), the logical distance \(d_{i,x}^x\) between the current forecast value and the multiple measured actual mean values can be expressed as follows:

\[
\begin{align*}
\text{if } x & \in x^D \\
\text{then } d_{i,x}^x & = \sqrt{\frac{|p_i^x - m_i^x|^2 + |p_{i2}^x - m_{i2}^x|^2 + |p_{i3}^x - m_{i3}^x|^2}{3}} \\
\end{align*}
\]
\[(p_1^i, p_2^i, p_3^i)\] is the predicted value of the customer attribute using the triangular fuzzy number and \((m_1^i, m_2^i, m_3^i)\) is the average of the customer attributes using the triangular fuzzy number. If a dynamic customer attribute value ranges from 1 to \(n\), its triangular fuzzy number is calculated as

\[
p_i^l = \left( \max \left( \frac{l-1}{n}, 0 \right), \frac{l}{n}, \min \left( \frac{l+1}{n}, 1 \right) \right),
\]

\(l \in [1, n]\)

According to the theory of foreground, the decision-maker has the risk preference when the expected evaluation value is greater than the actual measurement value \((p_i^l > m_i^l)\). When the expected evaluation value is smaller than the actual measurement value \((p_i^l < m_i^l)\), the decision-maker will take the initiative to avoid the risk. Because of the difference of decision-makers' cognition of risk aversion in different situations, the risk factors \(\mu_1\) and \(\mu_2\) are set in the process of calculating the foreground value, reflecting the decision preferences of different decision-makers. The dynamic demand foreground value can be calculated by formula (25).

\[
FV = \begin{cases} 
\mu_1 l d_i^r & p_i^l > m_i^l \\
\mu_2 l d_i^r & p_i^l < m_i^l 
\end{cases}
\]

\[(26)\]

Obviously, when the dynamic demand foreground is positive, the larger the value of \(\mu_1\) is, with the same foreground valuation, the more optimistic the decision-maker customer selection decision is. When the dynamic demand foreground is negative, the larger the value of \(\mu_2\) is, with the same foreground valuation, the more pessimistic the decision-maker clients choose to make their decisions.

4.2. Customer Clustering Methods. Considering that dynamic clients are obtained by proactive risk assessment and cannot ensure certainty in the process of delivery, clustering constrained only with actual demand will lead to the ineffective clustering expansion from the regional boundaries. Therefore, adding to the clustering algorithm the service radius expansion factor \(u\) and the load expansion factor \(v\) allows the demand in the cluster area (including the exact known static customer demand and possible dynamic customer demand) to be greater than the vehicle load. However, the vehicle is not allowed to be overloaded when leaving the distribution center at every distribution stage. On this basis, the proactive divisional scheduling strategy under the service area is divided as follows.

Step 1. Calculate the adjacency matrix between the distribution center and all the customers based on the known customer coordinate data.

Step 2. Search each node \(i\) as a center within the service radius \(R\), record the client nodes falling within the range, and select the circle to which the node with the largest number of client nodes belongs as the proactive service subregion.

Step 3. Generate proactive service subarea.

Step 3.1. Center of gravity to determine the subregion proactive scheduling center.

Step 3.2. Determine whether the total customer demand \(QD\) (iter) within the initial service subarea is greater than the vehicle load \(Q\); if yes, execute Step 3.3; if less, go Step 3.4.

Step 3.3. If the result of Step 3.2 is true, calculate the distance between the customer and the proactive scheduling center falling within the subregion to obtain the client sequence \(M\) sorted in descending order and the first element of the sequence \(M\) removed; get \(QD\) (iter + 1) to implement Step 3.5 judgment conditions.

Step 3.4. Expand the search radius to \(uR\) to determine whether it is true or not, and Step 3.2 is executed; if false, Step 4 is executed.

Step 3.5. Determine whether \(QD\) (iter) < \(Q\) < \(Q\) (iter + 1) is true, and if true, Step 4 is performed; if false, Step 3.2 is executed.

Step 4. Remove the clients contained in the service subregion generated in Step 3 and go to Step 5.

Step 5. Judge whether all the customers have been included in all clusters; if the judgment result is false, execute Step 2; if the judgment result is true, the algorithm is terminated.

As vehicle overloading is not allowed in practical applications, the estimated load in the cluster partition generated by the clustering algorithm is likely to exceed the actual load. Therefore, considering the practical application scenario, the strategy of replenishing the replenishment vehicle between the distribution center and each proactive dispatch center is adopted to solve the distribution difference problem that the actual delivery demand is greater than the subregion of the vehicle load.

4.3. Proactive Demand Quotas and Replenishment Strategies. Since each distribution logistics center covers a relatively stable service area, the key to optimizing the primary distribution path is to evaluate the likely value of demand in that area. The long-term distribution practice of urban distribution companies accumulated a wealth of historical distribution data and it provides a basis for the assessment of the demand for each distribution subregion.

It is assumed that the average historical demand of a distribution-type logistics center is \(\mu_i\), the variance is \(\sigma_i\), and the customer dynamic degree is introduced to evaluate the customer changes in each cluster area. The calculation formula of regional dynamic degree is

\[
DD_i = \frac{n_{dyn}}{n_{stat} + n_{dyn}}
\]

\[(27)\]

\(n_{dyn}\) contains predicted dynamic customers and unpredicted dynamic customers, and \(n_{stat}\) is a static customer in the delivery system.
During the delivery process, due to the new demand in the subarea, the supply of the initial delivery plan is smaller than the actual demand in the subarea, and the design of the replenishment path must be supplemented with the failure of dynamic customer forecasting.

Therefore, the set replenishment probability threshold $p_{accept}$ determines whether the distribution type logistics center $i$ is a replenishment subarea, and the formula is as follows:

$$a_i = \begin{cases} 1, & p_i \geq p_{accept} \\ 0, & p_i < p_{accept} \end{cases}$$  \hspace{1cm} (28)

After determining whether a distribution logistics center's covering subarea is a replenishment subarea, considering the differences in demand levels in the respective areas and their distance from the distribution center, the design is configured to allocate demand quotas for each subarea of the initial path planning based on the historical demand method:

$$q_i^* = \begin{cases} \text{rand} \left( \mu_i, \mu_i + 3\sigma \right), & a_i = 0 \\ \mu_i, & a_i = 1 \end{cases}$$  \hspace{1cm} (29)

where $\text{rand} \left( a, b \right)$ denotes the random number in the interval $[a, b]$ and $\min \left[ a, b \right]$ denotes the smaller of the two.

4.4. Vehicle Routing Optimization Program

(1) **Scan Operator.** Due to the large dependence of tabu search algorithm on initial solution, the initial solution of scanning method is chosen, and the initial solution is optimized by using tabu search algorithm to obtain higher performance solution. The specific steps for constructing the initial solution of the scan operator are as follows: first of all, taking the distribution center as the origin and any customer as the starting point to build the polar coordinates of the starting vector; taking the distribution center as a starting point and the other client as the destination vector to calculate the angle between the vector and the starting vector; the vector angles are sorted first, and then the initial customer sequence is generated based on the path constraints.

(2) **Construct Temporary Solution.** In the process of taboo search, the neighborhood of the current solution is transformed, and the scope of the solution space that can be searched out is expanded to increase the optimization ability of the algorithm. This paper defines the following four kinds of neighborhood optimization operator:

- **T1:** Randomly selected customers are removed from the vehicle and reinserted randomly;
- **T2:** Randomly exchange two randomly selected customers;
- **T3:** Randomly select two subroute segments to exchange with each other;
- **T4:** Choose two customers at random and reverse all the customers located between the two customers.

For each neighborhood operation, the following two acceptance strategies are used: Strategy A, the first improvement, stopping the optimization after the first improvement obtained after the neighborhood transformation; Strategy B, the best improvement, repeating the same operator run $n$ times and choosing the best improvement that appears during the experiment. If running $n$ operations, the current solution has not been improved. Then terminate this neighborhood operation.

For each customer on each path and path, five kinds of local optimization operators and one of the two acceptance strategies are randomly and independently selected with a uniform probability (1/4 and 1/2, respectively). Search depth is $N$ times.

(3) **Construct Contraindications and Taboos.** To avoid the algorithm getting into the local optimum, we need to judge whether the neighborhood solution is better than the historical optimal solution. If the neighborhood solution is superior to the historical optimal solution, the historical optimal solution is updated, and the neighborhood is treated as a taboo object. And then determine whether the taboo table is full; if the taboo table is full, remove the first element of taboo table, move the other elements to the left by one, and insert the taboo object into the tail of taboo table; if taboo table is not full, the taboo object is directly inserted into the first nonzero position. If the neighborhood solution is worse than the historical optimal solution, the depth of search will be increased by 1 to continue to determine whether the solution of the next neighborhood satisfies the taboo condition.

(4) **Quality Evaluation and Processing Method of Solutions.** The algorithm does not accept the transformation of the infeasible solution and discards it directly; that is, if the total amount of the distribution path exceeds the vehicle's nuclear load requirement, it is regarded as an infeasible solution and is directly removed from the solution space.

To sum up, the tabu search routing optimization procedure designed in this paper is shown in Figure 1.

5. Numerical Test

5.1. **Data Description.** To illustrate the feasibility of the proposed method, we use the distribution data of large logistics company in Chongqing, China. Chongqing is an important city in southwestern China and is in the middle and upper reaches of the Yangtze River. Its port trade is well developed. Our selected logistics and distribution company is a larger auto parts distribution company in Chongqing with up to 35 service customers, as shown in Figure 2. The distribution company has two different types of ZA and HA parts systems; the customer may need one of two major categories of parts or at the same time need two major categories of parts. The two types of vehicles owned by the enterprise are marked as B207 and CD101, and their authorized weights are 300 and 150. Due to the fragmented distribution of customers and the dynamic changes in demand, the hub located at the headquarters of the enterprise is located at (520, 280) in Figure 2. Establish distribution logistics center in a few more concentrated areas of customers. The parts produced by the company are first transported from the
B207 vehicle to the distribution logistics center, and then the CD101 vehicle is used from the distribution logistics center to deliver the customer's required parts to the corresponding customer.

Customers that have appeared in the delivery system at the start of delivery are called static customers, and nonappearing customers use data-driven demand forecasting methods to decide whether to include them in the delivery system. The static customer for each cycle is identified by the customer's production plan and is known at the beginning of the cycle. Table 1 is a distribution cycle customer information form.

5.2. Data-Driven Dynamic Customer Identification

(1) Customer Demand Forecast. According to the historical demand of each customer collected by the enterprise for a total of 30 days, the mean and standard deviation are shown in the following table. Using KS nonparametric test to test the historical demand of each customer in historical data, we found that the normal distribution can better reflect the change of the demand of historical data. The fitting results are shown in Table 2.

(2) Dynamic Customer Evaluation. To effectively describe the possible situation of dynamic customers in the service area, the dynamic customer attributes are described by three evaluation indexes: customer dependence $a_1$, payment speed $a_2$, and demand $a_3$. The corresponding weights of the three indicators are 0.4, 0.3, and 0.3. The above three assessment measures belong to the fuzzy evaluation, in which the customer dependency is divided into four categories: dependent, slightly dependent, medium dependent, and heavily dependent. The payment speed is divided into three levels: procrastination, advancement, and immediateness. The expert gives the prediction reference level from his experience, that is, the $l$ value in formula (25), and then calculates the predicted value of the attribute. For the customer dependency and payment speed attributes selected in this paper, formula (25) can be embodied as formulas (30) and (31), where the $l$ value is consistent with the definition order of the customer attribute.

$$a_1 = \begin{cases} \max \left( \frac{l - 1}{4}, 0 \right), & l \in [1, 4] \\ \frac{l}{4}, & l \in [4, 1] \end{cases}$$

$$a_2 = \min \left( \frac{l + 1}{4}, 1 \right)$$
Table 1: Customer information form for complete delivery cycle.

<table>
<thead>
<tr>
<th>Customer type</th>
<th>Customer Index</th>
<th>The number of customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static customers</td>
<td>1,2,5,6,8,9,11,12,14,15,16,17,18,20,22,24,25,26,27,30,32,34,35</td>
<td>23</td>
</tr>
<tr>
<td>Dynamic customer</td>
<td>3,4,7,10,13,19,21,23,28,29,31,33</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 2: 30 Day historical needs distribution, all 35 clients.

<table>
<thead>
<tr>
<th>Customer Number</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Customer Number</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>3.35</td>
<td>19</td>
<td>13.6</td>
<td>3.04</td>
</tr>
<tr>
<td>2</td>
<td>29.98</td>
<td>6.71</td>
<td>20</td>
<td>14.1</td>
<td>3.15</td>
</tr>
<tr>
<td>3</td>
<td>15.7</td>
<td>3.51</td>
<td>21</td>
<td>10.2</td>
<td>2.28</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>3.35</td>
<td>22</td>
<td>27</td>
<td>6.04</td>
</tr>
<tr>
<td>5</td>
<td>11.71</td>
<td>2.62</td>
<td>23</td>
<td>13.6</td>
<td>3.04</td>
</tr>
<tr>
<td>6</td>
<td>10.5</td>
<td>2.35</td>
<td>24</td>
<td>22.94</td>
<td>5.13</td>
</tr>
<tr>
<td>7</td>
<td>17</td>
<td>3.8</td>
<td>25</td>
<td>20</td>
<td>4.47</td>
</tr>
<tr>
<td>8</td>
<td>14.6</td>
<td>3.27</td>
<td>26</td>
<td>15.1</td>
<td>3.38</td>
</tr>
<tr>
<td>9</td>
<td>24.13</td>
<td>5.4</td>
<td>27</td>
<td>11.65</td>
<td>2.6</td>
</tr>
<tr>
<td>10</td>
<td>25.6</td>
<td>5.73</td>
<td>28</td>
<td>18.11</td>
<td>4.05</td>
</tr>
<tr>
<td>11</td>
<td>15.06</td>
<td>3.37</td>
<td>29</td>
<td>13</td>
<td>2.91</td>
</tr>
<tr>
<td>12</td>
<td>16.33</td>
<td>3.65</td>
<td>30</td>
<td>15.08</td>
<td>3.37</td>
</tr>
<tr>
<td>13</td>
<td>16.78</td>
<td>3.75</td>
<td>31</td>
<td>12.97</td>
<td>2.9</td>
</tr>
<tr>
<td>14</td>
<td>13.31</td>
<td>2.98</td>
<td>32</td>
<td>18.66</td>
<td>4.17</td>
</tr>
<tr>
<td>15</td>
<td>13</td>
<td>2.91</td>
<td>33</td>
<td>17.59</td>
<td>3.93</td>
</tr>
<tr>
<td>16</td>
<td>16.7</td>
<td>3.74</td>
<td>34</td>
<td>19.9</td>
<td>4.45</td>
</tr>
<tr>
<td>17</td>
<td>15.82</td>
<td>3.54</td>
<td>35</td>
<td>18.6</td>
<td>4.16</td>
</tr>
<tr>
<td>18</td>
<td>13.9</td>
<td>3.11</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ a_2 = (\max \left( \frac{l - 1}{3}, 0 \right), \frac{l}{3} \min \left( \frac{l + 1}{3}, 1 \right)), \quad l \in [1, 3] \]  

For customer demand, the expert generates the forecast value of customer demand in the distribution cycle according to the statistical information of customer demand provided in Table 2. After getting the predicted value of the three attributes, the decision-maker uses formula (26) to evaluate the dynamic customer and confirm whether to accept the dynamic customer according to the relationship between the predicted value of the expert and the expected value and the risk preference. When calculating the foreground value, \( \mu_1 = 1, \mu_2 = -2.25 \) is taken; since the customer's attribute dimensions are different, the data is normalized. The processed dynamic customer foreground values are shown in Table 3 (note: due to space limitations, only the evaluation results of the first stage dynamic customer prospect value are provided). The dynamic customer evaluation formula is as follows:

\[ f = 0.4 \ast f_1 + 0.3 \ast f_2 + 0.3 \ast f_3 \]  

A negative forecasted value indicates that the predicted value of firm property is less than the previous average value, indicating that the decision-maker’s performance on the client’s property is more pessimistic; the positive result of the foreground value assessment indicates that the firm’s forecast value of the client’s property is higher than that of the previous period, indicating that the decision-maker is more optimistic about the customer service outlook. Therefore, in this paper, dynamic customers with a foreground value greater than zero are included in the distribution network, corresponding to 6 clients with indexes 3, 7, 28, 29, 31, and 33 in Table 3 and determining the service attribution of each customer through the customer clustering method.

5.3. Customer Clustering and Demand Quotas. As shown in Figure 2 and Table 3, the decision-makers are pessimistic about the historical performance of customers 4, 10, 13, 19, 21, and 23. The original distribution plan of this distribution cycle will not consider these customers, only clustering the customers entering the distribution system. The maximum service radius bounded by the enterprise is 100, the service radius is 80, and the service radius maximum expansion scale is 0.25; the expansion factor \( u \) has a step size of 0.05, the load expansion factor \( v \) has a step size of 0.1, and the maximum expansion scale is 0.5. The customer clustering results are shown in Figure 3.

After each cluster area and its customer base are served, it is necessary to determine the distribution quotas for the area during the initial and replenishment phases, based on the historical performance of customer needs in the area. Since the demand of each customer point obeys a normal distribution, the customer's historical demand in
Table 3: Evaluation table of the dynamic client foreground value.

<table>
<thead>
<tr>
<th>Dynamic Customer Index</th>
<th>Customer Dependency (f1)</th>
<th>Payment speed(f2)</th>
<th>Demand (f3)</th>
<th>Foreground value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Predictive value</td>
<td>Past average</td>
<td>Predictive value</td>
<td>Past average</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>21</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>28</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>29</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>33</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 3: Delivery system customer clustering results.

Each clustering subarea will be summed up for 30 days, and the mean and variance will be calculated. According to formula (28) for the distribution acceptance standards, where \( p_{\text{accept}} = 0.35 \), then clusters 1, 2, and 3 are divided into nonreplenishment subregions, and region 4 is divided into replenishment subregions, according to formula (29) to determine the initial allocation of each region as shown in Table 4.

5.4. Vehicle Scheduling Scheme. After determining the customer base to be served in each cluster subarea, the routing optimization program described in Section 4.4 is used to optimize the delivery path based on the initial allocation quota. In the primary distribution network, the initial distribution phase transports the goods to the distribution type logistics center (that is, the distribution type logistics center) starting from the hub-type center according to the distribution quota specified in Table 4. Secondary distribution network traffic will start from the clustering center and directly provide delivery service for the customer according to each customer’s demand—static customers with its actual demand distribution and distribution of dynamic customer to enterprise for its demand forecast.

The whole distribution process is a complete distribution workday consisting of eight hours, which is divided into the initial stage and replenishment stage, each with four hours; the initial stage is only static customer service, and the expert experience is used to conduct dynamic customer evaluation. The dynamic replenishment phase is enabled at half the time (that is, the fourth hour) for static customers that have not yet been served and the dynamic customer service obtained by the proposed method. A complete cycle of distribution job scheduling process is shown in Figure 4. Static customers with their actual demand distribution, dynamic customers with expert predicted value of stock, the actual demand is clear at the customer.

The tabu search algorithm designed by 4.4 part is programmed on the platform of Win 7 32 by using MATLAB 2015B, and the optimal distribution routing of 2E-VRP is generated with the design example data. The algorithm parameters are as follows: the taboo table length is 20, \( N = 10 \), the maximum iterations number is 1000, and the initial solution scale is 100. Table 5 shows the optimal initial and replenishment routings for distribution centers to four cluster centers, and Table 6 shows the initial distribution subpaths for four cluster subregions.

Under the accurate forecast result, the distribution logistics centers in the nonreplenishment clustering area have different levels of surplus stock from Table 6, while the replenishment clustering area has a small amount of demand difference. At the same time, due to the small difference in demand of cluster 4, dispatched vehicles may be worth more than replenishing their replenishment. However, in practical application scenarios, prediction errors are inevitable. In addition, the analysis of the periodic distribution data found
that the forecast number, 33 customers, did not actually generate demand; and the number of 21 and 23 customers which were given up in the beginning of the distribution due to decision-making risk aversion issued new distribution needs in the distribution process. Then, the path of cluster 3 and cluster 4 in the initial path is adjusted locally. The results are shown in Table 7. The actual demand of 29 customers involved in the distribution operation is shown in Table 8.

Because the original intention of PVRP design is to put the commodity closer to the customer, the precondition of its realization is that the distribution logistics center has enough stock to serve the new dynamic customer in the replenishment stage. Therefore, it is stipulated that unit shortage cost $c_i^h = 2$ and unit excess cost $c_i^u = 1$, from which we can get a complete cycle of changes in the stock of goods and the cost of each center, as shown in Table 9.

Given that the leasehold cost of distribution logistics center is 50, the total leasehold cost is 200. It is assumed that the fixed cost of the model B207 is 1000, the unit cost of use is 5, the fixed model cost of the model C101 is 600, and the unit cost is 3. B207 and C101 were used for delivery, respectively. The results of comparison with the two-level vehicle routing scheme designed in this paper are shown in Table 10.

As shown in Table 10, the total cost of PVRP solution delivery is 17903.35, in terms of delivery costs, slightly higher than the delivery of B207 bicycle and C101. However, due to the partition clustering algorithm, the PVRP scheme designed by the PVRP method has shorter running time than

### Table 4: Clustering subarea parameter description table.

<table>
<thead>
<tr>
<th>Number</th>
<th>Clustering 1</th>
<th>Clustering 2</th>
<th>Clustering 3</th>
<th>Clustering 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster center coordinates</td>
<td>(161,318)</td>
<td>(145,125)</td>
<td>(350,110)</td>
<td>(368,290)</td>
</tr>
<tr>
<td>The number of accepted customers</td>
<td>8</td>
<td>6</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>The number of dynamic customers</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Dynamic value</td>
<td>0.2</td>
<td>0.25</td>
<td>0.3</td>
<td>0.44</td>
</tr>
<tr>
<td>Daily average</td>
<td>130.58</td>
<td>98.52</td>
<td>162.79</td>
<td>80.59</td>
</tr>
<tr>
<td>Daily variance</td>
<td>11.11</td>
<td>7.63</td>
<td>10.89</td>
<td>8.12</td>
</tr>
<tr>
<td>Initial rationing</td>
<td>143.07</td>
<td>113.59</td>
<td>180.31</td>
<td>80.59</td>
</tr>
<tr>
<td>Radius expansion scale u</td>
<td>0.1875</td>
<td>0.1875</td>
<td>0.1875</td>
<td>0</td>
</tr>
<tr>
<td>Load expansion factor v</td>
<td>-0.0462</td>
<td>-0.2427</td>
<td>0.2021</td>
<td>-0.4627</td>
</tr>
</tbody>
</table>

### Table 5: A distribution network path table.

<table>
<thead>
<tr>
<th>Delivery phase</th>
<th>Routing index</th>
<th>Routing connection</th>
<th>Routing mileage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial phase</td>
<td>Routing 1</td>
<td>DC-1-2-DC</td>
<td>760.44</td>
</tr>
<tr>
<td></td>
<td>Routing 2</td>
<td>DC-3-4-DC</td>
<td>373.64</td>
</tr>
<tr>
<td>Replenishment phase</td>
<td>Routing 1</td>
<td>DC-4-DC</td>
<td>180.83</td>
</tr>
</tbody>
</table>
Table 6: Distribution initial routing table for each cluster area.

<table>
<thead>
<tr>
<th>Cluster area number</th>
<th>Routing connection</th>
<th>Routing mileage</th>
<th>Routing load</th>
<th>Surplus load</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>c1-17-8-15-7-5-2-3-8-c1</td>
<td>328.5</td>
<td>132.87</td>
<td>10.2</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>c2-14-16-12-6-1-9-c2</td>
<td>402.11</td>
<td>95.97</td>
<td>17.62</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>c3-31-35-33-28-24-c3</td>
<td>237.75</td>
<td>176.86</td>
<td>3.45</td>
</tr>
<tr>
<td></td>
<td>c3-22-18-20-25-27-c3</td>
<td>225.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster 4</td>
<td>c4-26-30-34-32-29-c4</td>
<td>178.26</td>
<td>81.74</td>
<td>-1.15</td>
</tr>
</tbody>
</table>

Table 7: Updated delivery path.

<table>
<thead>
<tr>
<th>Cluster area number</th>
<th>Updated routing connection</th>
<th>Updated routing mileage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 4</td>
<td>c4-21-23-c4</td>
<td>96.6</td>
</tr>
</tbody>
</table>

Table 8: All customer demand information sheets in the whole cycle.

<table>
<thead>
<tr>
<th>code</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>real demand</td>
<td>16.47</td>
<td>33.23</td>
<td>18.84</td>
<td>0</td>
<td>14.2</td>
<td>12.7</td>
<td>20.33</td>
<td>16.33</td>
<td>19.6</td>
</tr>
<tr>
<td>forecast demand</td>
<td>-</td>
<td>-</td>
<td>16.1</td>
<td>16.1</td>
<td>-</td>
<td>-</td>
<td>16.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>code</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>real demand</td>
<td>0</td>
<td>12.6</td>
<td>16.7</td>
<td>0</td>
<td>16.3</td>
<td>15.63</td>
<td>14.2</td>
<td>14.33</td>
<td>14.5</td>
</tr>
<tr>
<td>forecast demand</td>
<td>26.7</td>
<td>-</td>
<td>-</td>
<td>179</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>code</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td>real demand</td>
<td>0</td>
<td>16.8</td>
<td>14.4</td>
<td>14.1</td>
<td>16.9</td>
<td>21.6</td>
<td>16.46</td>
<td>12.9</td>
<td>19.4</td>
</tr>
<tr>
<td>forecast demand</td>
<td>13.7</td>
<td>-</td>
<td>9.63</td>
<td>-</td>
<td>12.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>code</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>real demand</td>
<td>20.8</td>
<td>16.7</td>
<td>16.7</td>
<td>13.7</td>
<td>16.3</td>
<td>19.2</td>
<td>14.8</td>
<td>20.3</td>
<td></td>
</tr>
<tr>
<td>forecast demand</td>
<td>18.8</td>
<td>11.1</td>
<td>-</td>
<td>12.19</td>
<td>-</td>
<td>17.85</td>
<td>14.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The single-vehicle scheme. Because the clustering center is closer to the customer base, especially the clustering area farther away from the hub, the proactive demand quota policy can dispatch the parts to its service from the distributed logistics center faster.

5.5. Algorithm Performance Evaluation. To avoid the stochastic error of the algorithm, the C101 model, the customer distribution shown in Figure 2, and the corresponding customer demand mean in Table 2 are used to form the example data. The stability of the proposed algorithm is verified by using the same tabu search algorithm parameters as those in the above scheduling scheme. At the same time, the genetic algorithm is used to solve the example to verify the relative accuracy; the two algorithms are run 20 times, respectively, and the results are compared as shown in Figure 5. It shows that the optimal solution of tabu search algorithm is 3876.33, and the optimal solution obtained by genetic algorithm is 3973.56, which improves by 2.4%. According to the statistical mean of 20 algorithms, the average value of tabu search algorithm is 3957.99, and the average value of genetic algorithm is 4047.89, which improves by 2.2%. When the variance of the two algorithms is statistically found, the standard deviation of the tabu search algorithm is 62.48, and the standard deviation of the genetic algorithm is 40.52. It can be seen that the stability of the solution is slightly worse than the genetic algorithm, but considering the quality of the solution is better than the genetic algorithm. It is proved that the proposed algorithm is superior to the genetic algorithm in the quality and stability of the solution and can better realize the routing optimization tasks in each stage.
Table 9: Stock quantity and cost analysis of each center in a complete cycle.

<table>
<thead>
<tr>
<th>Cluster number</th>
<th>Stock quantity</th>
<th>Initial phase</th>
<th>Replenishment phase</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Path load</td>
<td>Margin</td>
<td>Path load</td>
<td>Replenishment quantity</td>
</tr>
<tr>
<td>Cluster 1</td>
<td>143.07</td>
<td>125.90</td>
<td>171700</td>
<td>0</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>113.59</td>
<td>95.97</td>
<td>176200</td>
<td>0</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>180.31</td>
<td>172.86</td>
<td>74500</td>
<td>155.01</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>80.59</td>
<td>60.70</td>
<td>198900</td>
<td>91</td>
</tr>
</tbody>
</table>

Table 10: Path costs of other programs.

<table>
<thead>
<tr>
<th>Scheduling scheme</th>
<th>Single-type B207</th>
<th>Single-type C101</th>
<th>PVRP B207</th>
<th>C101</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheduling times</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Mileage</td>
<td>3013.77</td>
<td>4629.83</td>
<td>1314.91</td>
<td>1468.86</td>
</tr>
<tr>
<td>Distance cost</td>
<td>17068.85</td>
<td>17489.49</td>
<td>9574.55</td>
<td>8006.58</td>
</tr>
<tr>
<td>Stock cost</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>122.22</td>
</tr>
<tr>
<td>Leasehold cost of distribution logistics center</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>200</td>
</tr>
<tr>
<td>Total cost</td>
<td>17068.85</td>
<td>17489.49</td>
<td>17903.35</td>
<td></td>
</tr>
</tbody>
</table>

6. Management Insights

Since the problem to be solved comes from the actual application scenario, we can obtain the following inspiration from the optimization results of this article.

Data analysis is an important means for enterprises to obtain profits and tap the potential benefits. It not only helps enterprises to understand the historical operating conditions, but also helps enterprises to analyze the laws of customers and discover the intrinsic dynamics of customer needs. In our forecasting method, based on the data analysis of the past 30 days, we found that customer historical data can be effectively fit with normal distribution and give the distribution parameters of each customer. This is very useful for enterprises to grasp the trend of changes in customer demand, and we also consider that the two indicators of payment speed and customer reliance to comprehensively measure the dynamic customer can be accepted. The data-driven demand forecasting method proposed by the actual enterprise data validation can effectively identify dynamic clients. Because of the unavoidable prediction errors, two dynamic clients are not predicted in the example of this paper with the more conservative customer selection decision, and there is a customer that failed to predict. Therefore, we can make the following judgment: (1) when there are a large number of available historical data, nonparametric test can be used to analyze the distribution characteristics of historical data and find the optimal distribution among them, so as to achieve effective prediction of customer demand. (2) The dynamic customer’s choice is related to the decision preference. When the decision has the risk preference, some clients with lower probability of occurrence may be included in the distribution network. In this way, more customers may be considered in the distribution system, but may also bring additional scheduling costs. When the decision-making is risk-averse, the choice of service customer group will tend to be conservative, and some dynamic customers may be ignored. However, this article effectively prevents this by designing a proactive demand quota strategy.

The combination of customer clustering and demand quotas is equivalent to directing vehicles to areas where customer needs are more concentrated. If the dynamic demand occurs in the local area and the vehicle response to the dynamic demand from the hub center may require higher service and time cost, the vehicle responding to the dynamic demand promptly from the distribution center has the great significance for improving customer satisfaction. Proactive demand quotas allow distribution centers to have more cargo than the current phase of demand. In the event of a failed forecast or a transient customer, the quantity of cargo stocked by the distribution logistics center ensures its ability to adapt to minor fluctuations in regional demand. In addition, when the dynamic demand in the region is too large and the logistics center is short of supply, the replenishment phase will start with the volume of cargo that is not replenished and the replenishment scope will be drastically reduced. The PVRP method designed in this paper is less significant in terms of delivery costs. However, with vehicle collocation and secondary network design, it is well adapted to the current traffic restrictions in urban management. Although the use of larger vehicles has some cost advantages, it is less flexible in complex transport networks and may require additional delivery of customer goods due to traffic limit. When using smaller models, the limited number of customers with one-off service due to the low vehicle load leads to a larger total path length of the vehicle, thereby increasing the workload and working hours of the drivers and lacking of humanistic concern.

7. Conclusion and Outlook

This paper presents a 2E-PVRP problem with four modules: data driven demand forecasting method, customer clustering method, proactive demand quotas and replenishment
strategy, and vehicle routing optimization procedure. The conclusions are as follows: (1) Data driven demand forecasting method is helpful to give full play to the value of the existing historical data in the enterprise to identify the dynamic customers in the distribution process and reduce the impact of the uncertainty in the delivery process on the delivery work. (2) Customer clustering and demand quota design can guide the vehicles in the areas where customers are more concentrated and respond to dynamic customers faster. (3) 2E-PVRP adopts two-level network and multivehicle design, which enhances the adaptability of distribution system to traffic policy and pays more attention to the humanistic concern of drivers. In the future, we will continue to study the site selection mechanism of distribution logistics centers that take social and economic factors into consideration. The case study shows that there is a problem that the real delivery rate of the replenishment vehicles is too low when the regional replenishment volume is small. Ministry will consider the introduction of multivehicle matching strategy in the first-class distribution network to achieve the optimal allocation of distribution resources.

Data Availability

(1) Basic data: the simulation data needs a distribution center and several customers to be distributed. The researcher can use the simulation method to randomly generate the dynamic customers in the text and simulate the simulation network composed of the data in Figure 2 and Table 1. (2) The generation of historical data: the example of this paper comes from the statistical analysis of enterprise data. The researchers can simulate the historical performance of dynamic customers by randomly generating data. The distribution model obtained from the simulated data may not be consistent with the normal distribution of the results of this paper. However, this does not affect the generation of the final result. Regardless of the distribution, the purpose is to forecast the customer demand. It is also advisable for the researcher to obtain other distributions that can effectively reflect the customer's historical performance. (3) Dynamic customer evaluation, using the assessment indicators provided in this paper and formula (30) and formula (31) to obtain the forecast value of the indicator, and then using formula (32) to comprehensively evaluate the three indicators to generate the prospect value of dynamic customers. (4) The customer clustering and demand quotas can be implemented according to the method proposed in the article. Afterwards, the emergency search path optimization program designed by the article can be used to obtain the desired result. Again, the data in this article cannot be disclosed to the public because it involves commercial secrets. If the subsequent researchers need to verify or redevelop the results of this paper, follow the methodologies provided in the third part of the article’s solution framework for data processing to get the desired results. In addition, we will be honored if the researchers can make personal innovations on the basis of the results of this paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study is supported by the National Natural Science Foundation of China (71502021, 71602015), funded by the Ministry of Education of Humanities and Social Science Fund Project (2014YJC6300382015XJC630007), supported by Postdoctoral Science Foundation (2016T90862), funded by the Chongqing Foundation and Frontier Research Project (cstc2016jcyjA0160), Chongqing City Board of Education, Humanities and Social Science Research Project (17SKG073), and the Chongqing Municipal Science and Technology Research Project (KJ1500702).

References


