Study on the Optimization of Vehicle Scheduling Problem under the E-commerce Environment

12 Haoxiong Yang and 1 Jindan Li

1 School of Business, Beijing Technology and Business University, 100048, Beijing, China
1 Research Base of Capital Retail, 100048, Beijing, China

Abstract: Studying vehicle scheduling problem under the e-business environment, is beneficial to reduce distribution costs, improve service quality, promote the development of e-commerce enterprises. In this study, for the new characteristics of logistics distribution such as small demand, many kinds of goods, scattered location, the vehicle scheduling problem under electronic commerce environment is studied, using clustering analysis method to divide distribution area. The VRPTW model is established, using genetic algorithm to solve the model. Through simulation experiment, comparing with the traditional VRP model, it found the optimized cost is lower than the unoptimized cost. This study is verified on the effectiveness of the method for VRPTW optimization model.

Keywords: Electronic commerce, vehicle scheduling, time window, cluster analysis, genetic algorithm

INTRODUCTION

With the rapid development of social progress, economic globalization, information networks, market competition is more intense, e-commerce is increasingly affecting today's economic life. However, the logistics management system is not perfect and can not keep up the pace of development of e-commerce, resulting in the formation of distribution bottleneck. Reasonable arrangements of vehicle scheduling under e-commerce environment will help to improve service quality and will help to alleviate bottlenecks of logistics distribution.

About researches of logistics under e-commerce environment, scholars in our country have studied. China's distribution services lag problem based on B2C e-commerce was put forward and some scholarly proposed distribution optimization solution (He, 2009). An innovative e-commerce distribution mode "distribution win-win" was proposed (Li, 2007). The idea to improve the responsibility tracing system subject of electronic commerce in China was put forward (Zhu, 2012).

The problem of classic vehicle distribution route was proposed (Dantzig and Ramser, 1959). Some Heuristic algorithm emerged (Clarke and Wright, 1964; Gendreau et al., 1991; Lawrence and Mohammad, 1996). Genetic algorithms optimization problem was analysed based on current Heuristic algorithm (Xie et al., 2010; Gong and Cheng, 2006). Researchers have solved the problem for vehicle scheduling optimization with time window constraints (Tang, 2011).

Nowadays, establishing distribution system in the traditional way can not fully meet the needs of e-commerce. Distribution scope of electronic business become increasingly wide, so the thesis uses the clustering method to improve vehicle scheduling problem. For new requirements "limited time logistics", it establishes Vehicle Routing Problem with Time Windows (VRPTW) model and uses genetic algorithm for finding the optimal solution to obtain satisfactory results, solving the problem of e-commerce logistics and distribution.

CONSTRUCTIONS FOR VEHICLE SCHEDULING MODEL

Dividing delivery area based on cluster analysis method:
According to vehicle load and other constraints, the distribution area is divided into several sub-regions and then seeking the optimal delivery routes in the sub-region, by using the cluster analysis method to narrow the search space for each optimization and thus greatly reducing the amount of calculation and improving the solving speed.

The characteristics of e-commerce logistics and distribution:

- Delivery area is wide
- Delivery time requirements are different
- Demandment of commodities for each customer is relatively small

Rapid sample cluster
determination of clustering evaluation index: In the e-commerce distribution, the customer's location directly

Corresponding Author: Jindan Li, Research Base of Capital Retail, 100048, Beijing, China

7827
affects the cost of transportation. Customer demand for time of the vehicle, producing time cost. Therefore, the customer location coordinates, the demand and the service time windows are as rapid sample clustering evaluation indexes.

Determination of the k value: Assuming one regional distribution for one vehicle, namely the number of vehicles for distribution is equal to the number of clusters of rapid sample cluster. The number of vehicles sent is calculated by the formula:

\[ k = \left[ \frac{\sum q_i}{\alpha Q_k} \right] \]

“[ ]” indicates the integral function, \( q_i \) indicates the demand for customer \( i \), \( Q_k \) represents the maximum load of vehicle \( k \), \( \alpha \) means a known factor, \( 0<\alpha<1 \). By formula to determine the value of \( k \) for clusters.

Basic steps of rapid sample clustering: Algorithm is processed as follows:

- From the entire sample \( n \) namely containing \( n \) objects in the order, choose \( k \) objects as the initial cluster centers \( m_i (i = 1, 2, \ldots, k) \)
- Using Eq. 1, the customer's location coordinates, the demand and service time windows as clustering evaluation index, calculated the distance \( d(p, m_i) \) from each object \( p \) to \( k \) clusters centers in the dataset
- Finding the minimum distance \( d(p, m_i) \) for each object \( p \), the \( p \) is classified into the same cluster \( m_i \)
- Calculating all objects, using Eq. 2 to recalculate the value of the cluster center \( m_i \) as a new cluster center
- Re-assign most similar clusters for objects of the entire dataset. This process is repeated until the minimum squared error criterion

**Equation 1:**

\[ d(i, j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \cdots + (x_{in} - x_{jn})^2} \]

\[ i = (x_{i1}, x_{i2}, \ldots, x_{in}) \]

\[ j = (x_{j1}, x_{j2}, \ldots, x_{jn}) \]

which are two \( n \)-dimensional data objects:

**Equation 2:**

\[ m_i = \frac{\sum x_i}{n} \]

\( m_i \) is the cluster center for \( k^b \), \( n \) is the number of data objects for \( k \)th cluster.

Squared error criterion tries to make the clustering results as compact as possible, that the similarity of objects within the cluster as high as possible. Defined as follows:

\[ E = \sum_{i=1}^{n} \sum_{|p-m_i|} \]

\( E \) is the sum of square errors of all objects, \( p \) is objects in space, \( m_i \) is the average of the cluster \( C_i \).

These attributes such as the position coordinates of all customers, demandment, service time, are as cluster variables. \( k \) objects in the \( n \) objects (clients) are chosen arbitrarily as initial cluster centers, through SPSS software repeat (2-4) process until clustering center no longer changed. Some customers with similar characteristic clustered together, distribution area is divided into different small areas, reducing the computation of the optimization process.

Vehicle scheduling in the distribution region

**Description of the problem:** Since, the time of distribution under e-commerce environment is different, the time of vehicle scheduling is unreasonable, sooner or later leading to the generation of time cost. Establishing vehicle scheduling with time windows model (VRPTW) is needed, making the delay time as little as possible, the time costs as small as possible and transport distances as small as possible, which are considered as the primary factors of objective function.

**Assumptions of VRPTW model:** In order to facilitate the description of the problem and construction of the corresponding mathematical models, this study proposed the following assumptions:

- Flow direction of goods is one-way, that is purely delivery
- There is only one distribution center and the position of beginning and ending for each line is in the distribution center
- The position coordinates of distribution center and the customer points are known
- The load of vehicle in the distribution process may not exceed the rated load
- Each vehicle is the same kind and the capacity is known
- The demandment of customers are known. The cargo in that position is only finished by one car and all customers should be serviced

**Construction of VRPTW model:**

- Description of symbol:
\[ q_i = \text{Demandment of customer i} \]
\[ Q_k = \text{Maximum load for the K car} \]
\[ c_{ij} = \text{Transportation costs from customer i to customer j} \]
\[ n = \text{Total number of customers who will be serviced} \]
\[ t_{ij} = \text{Time from customer i to customer j} \]
\[ t_i = \text{Travel time from distribution centre to customer i} \]
\[ s_i = \text{The service time of vehicle for customer i} \]
\[ ET_i = \text{The allowed earliest service time} \]
\[ LT_i = \text{The allowed latest service time} \]
\[ M = \text{A very large positive constant} \]
\[ a = \text{The penalty coefficient when service time is earlier than } ET_i, \text{ namely a yuan per hour as punishment} \]
\[ b = \text{The penalty coefficient when service time is later than } LT_i, \text{ namely b yuan per hour as punishment} \]

\[ y_{ij} = \begin{cases} 1 & \text{The vehicle runs from point i to point j} \\ 0 & \text{or else} \end{cases} \]

\[ x_k = \begin{cases} 1 & \text{task is accomplished by the vehicle k} \\ 0 & \text{or else} \end{cases} \]

- **VRPTW model:**

\[
\begin{align*}
\min f &= \sum_{k=1}^{K} \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} y_{ij} + a \sum_{i=1}^{n} \max\{0, ET_i - t_i\} + b \sum_{i=1}^{n} \max\{0, t_i - LT_i\} \\
\text{s.t.} &= \sum_{k=1}^{K} q_i x_k \leq Q_k, \forall k \in K \\
&\sum_{i=1}^{n} y_{ij} = x_k, \forall k \in K \\
&\sum_{j=1}^{n} y_{ij} = x_k, \forall k \in K \\
&\sum_{k=1}^{K} x_k = 1, i = 1, 2, \ldots n \\
&0 \leq \sum_{k=1}^{K} x_k \leq n \\
&\text{subject to} \quad t_i + s_i + t_j - M(1 - y_{ij}) \leq t_j 
\end{align*}
\]  

Equation 2 ensures that the sum of demand for customers on each path is no more than delivery of the vehicle load; Eq. 3-5 ensure that the needs of each customer point is served by only one vehicle; Eq. 6 ensures that the number of customers served by the k-th vehicle is no more than the total number n; Eq. 7 ensures the total time of the vehicle's route does not exceed a pre-set value, in order to meet delivery time requirements of customers.

**Genetic algorithm for VRPTW:** The route of each distribution area is optimized using genetic algorithm. Equation 1 is as the objective function, obtaining the optimal path for each distribution area.

Basic steps of genetic algorithm as follows:

- Selection of an encoding, defined a fitness function in the search space \( U \), given population size \( N \), crossover rate \( P_c \), and mutation rate \( P_m \), algebra T
- Randomly generated the \( N \) individuals in \( U \), \( S_1 \), \( S_2 \), ..., \( S_N \), composed of the initial population \( S_t = \{S_1, S_2, ..., S_N\} \), the counter is set algebra \( t = 1 \)
- Calculating fitness function of each individual \( S_t \) in \( S \), \( f = f(s_t) \)
- If the rules of the algorithm terminates, the algorithm will end, choosing the individual of maximum fitness in \( S \) as result of the request, or to calculate probabilities:

\[
P(s_t) = \frac{f_t}{\sum_{j=1}^{N} f_j} \quad t = 1, 2, \ldots, N
\]

And in accordance with the probability distribution above-mentioned determined the selected opportunities, randomly selected an individual from the \( S \) and replicated its chromosomes, a total of \( N \) times and then the resulting copied \( N \) chromosomes compose population \( S_t \):

- The number \( c \) of chromosome is determined by the crossover rate \( P_c \), the participants cross, randomly determining \( c \) chromosomes from \( S_t \), cross-matching operation and use the resulting new chromosomes instead of the original chromosomes, getting population \( S_t \)
- Mutation frequency \( m \) is determined by mutation rate \( P_m \), randomly determined \( m \) chromosomes from \( S_t \) which were the mutation operation and use the resulting new chromosomes instead of the original chromosomes, getting population \( S_t \)
-population \( S_t \) is as a new population, that is instead of \( S_t \), \( t = t + 1 \), return Eq. 3

7829
SIMULATION AND ANALYSIS

Case overview: Description of the problem as follows: randomly generated position coordinates and demand of 100 customers, assuming there is one distribution center and its location coordinate is (0, 0). Everyday the distribution center needs to serve 100 customers, whose geographic locations, demand and delivery time requirements are different. The vehicles in distribution centers are the same type, the same load, which are 200 units.

Case solution

Result of cluster analysis: Arbitrary choosing k objects from 100 customers as initial cluster centers, through SPSS software repeat (2-4) process until the cluster center is no longer changed. Customers with a high similarity are together as a same category and the distribution area is divided into different small regions.

\[ \alpha = 0.85, Q_k = 200, \text{by formula:} \]

\[ k = \left[ \sum_{i=1}^{n} q_i / \alpha Q_k \right] + 1 \]

calculate the number of clusters \( k = 10 \), namely the distribution area is divided into 10 small regions. Through SPSS software, the classification result of customers point is shown in Fig. 1.

Result of genetic algorithm: Using matlab genetic algorithm toolbox optimizes path for each region. The established VRPTW model is solved and finally getting the result via., MATLAB software.

Corresponding parameters of Genetic Algorithm are set as follows.

Penalty coefficient \( a = 5, b = 5 \), crossover rate \( P_c = 0.8 \), mutation rate \( P_m = 0.02 \), population size \( N = 1500 \), the maximum number of iterations \( T = 100 \):

- **Encoding:** This study uses natural number coding method to construct chromosome

- **Initial population:** Population size 1500 is given in this study. Through initial population function randomly generated 1500 individuals constituting the initial population, randomly generating 1500 initial feasible paths

- **Evaluation of fitness:** Through fitness function that is the objective function, to evaluate the fitness function value of each chromosome

- **Selection of operator:** This competition will be randomly selected as the selection operation operator selected by choosing individuals with high fitness function until the election is full

- **Crossover:** Cross-chromosomes crossover is performed by the crossover function, producing offspring

- **Mutation operator:** In this study, uniform mutation method, by mutation function, select the mutation probability of 0.02 for the parent chromosome mutation

- **Updates and termination of populations:** To ensure the optimal solution is getting better depend on the increase of the number of calculation iterations. In this case, the optimal solution in the last time after variation is added to population, preventing cross or mutation from losing the optimal solution, generating degeneracy phenomenon

MATLAB running interface for genetic algorithm is shown in Fig. 2.

The 10 routes in clustering regions are optimized, running MATLAB software 10 times. VRPTW model based on genetic algorithm solution obtained as shown in Table 1.

The results are: The total cost of 10 paths is 5784.7979, the time cost is 1217.5, transportation cost is 4567.2979.

Only the shortest path is considered as the objective function, under the same parameter settings, the result shown in Table 2

The results are: The total cost of 10 paths is 6245.5432, the time cost is 1900, transportation cost is 4345.5432.

![Fig. 1: Results of K-means clustering](image-url)
Fig. 2: Running interface of MATLAB

<table>
<thead>
<tr>
<th>No.</th>
<th>Customer delivery order</th>
<th>Time cost</th>
<th>Total cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0-21-22-33-32-51-76-73-72-67-77-0</td>
<td>108.0</td>
<td>582.6620</td>
</tr>
<tr>
<td>2</td>
<td>0-58-64-82-87-86-0</td>
<td>30.0</td>
<td>394.9524</td>
</tr>
<tr>
<td>3</td>
<td>0-20-16-11-4-7-24-94-71-78-2-99-0</td>
<td>167.5</td>
<td>556.0772</td>
</tr>
<tr>
<td>4</td>
<td>0-54-92-53-90-74-0</td>
<td>95.0</td>
<td>505.9761</td>
</tr>
<tr>
<td>5</td>
<td>0-1-3-19-31-13-6-10-15-25-8-12-1-7-9-0</td>
<td>112.5</td>
<td>471.6856</td>
</tr>
<tr>
<td>6</td>
<td>0-46-34-48-56-66-35-26-98-07-0</td>
<td>95.0</td>
<td>415.9799</td>
</tr>
<tr>
<td>7</td>
<td>0-28-32-44-48-81-34-41-95-65-0</td>
<td>65.0</td>
<td>450.0860</td>
</tr>
<tr>
<td>8</td>
<td>0-70-61-60-93-75-59-69-57-56-0</td>
<td>197.5</td>
<td>680.4058</td>
</tr>
<tr>
<td>9</td>
<td>0-49-83-89-79-62-45-47-52-85-84-63-0</td>
<td>215.0</td>
<td>876.9427</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No.</th>
<th>Customer delivery order</th>
<th>Time cost</th>
<th>Total cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0-22-51-35-72-73-76-67-72-37-21-0</td>
<td>400.971</td>
<td>265.0</td>
</tr>
<tr>
<td>2</td>
<td>0-58-86-87-82-64-0</td>
<td>362.5551</td>
<td>40.0</td>
</tr>
<tr>
<td>3</td>
<td>0-68-99-23-20-16-2-78-94-24-71-74-11-0</td>
<td>407.4461</td>
<td>292.5</td>
</tr>
<tr>
<td>4</td>
<td>0-54-92-53-90-74-0</td>
<td>410.9761</td>
<td>95.0</td>
</tr>
<tr>
<td>5</td>
<td>0-3-1-9-8-4-7-13-11-31-12-19-25-10-6-15-0</td>
<td>325.6000</td>
<td>207.5</td>
</tr>
<tr>
<td>6</td>
<td>0-26-97-98-66-50-48-34-35-46-0</td>
<td>307.5732</td>
<td>185.0</td>
</tr>
<tr>
<td>7</td>
<td>0-42-43-44-81-95-65-41-40-29-0</td>
<td>355.9592</td>
<td>167.5</td>
</tr>
<tr>
<td>8</td>
<td>0-60-56-57-69-55-75-91-70-61-0</td>
<td>470.6237</td>
<td>172.5</td>
</tr>
<tr>
<td>9</td>
<td>0-63-84-19-45-62-79-80-84-84-52-47-0</td>
<td>639.2384</td>
<td>245.0</td>
</tr>
<tr>
<td>10</td>
<td>0-18-37-39-36-38-38-7-7-28-100-5-59-89-36-88-14-30-0</td>
<td>665.7777</td>
<td>290.0</td>
</tr>
</tbody>
</table>

Table 2: Final outcome of path optimization based on the objective function VRPTW

Analysis for result: Compared to Table 1 and 2, the optimization results of VRPTW model saves the total cost 460,7453. The result of VRPTW model increases transportation cost 221,7547 but the time cost is relatively small, saving time cost 682.5. It is indicated that this method is more time-sensitive than the traditional VRP model.

CONCLUSION

This study analyzes the characteristics of e-commerce distribution. Using cluster analysis improves efficiency of optimization. Aimed at the stringent time requirements of e-commerce, VRPTW model is established, total distance traveled by the vehicle as little
as possible and vehicle delay time minimized as the main considerations of the objective function. But many aspects will be further deepened studied. The future research could focus on considering the volume of the vehicle, adding the constraints of the volume of the vehicle and comprehensively considering of various algorithms to achieve optimal objective.

ACKNOWLEDGMENTS

We would like to thank the reviewers for their helpful comments. This article is supported by the Social Science Foundation of China (Grant No. 11CGL105) and Backbone young talent training program in Beijing Education Commission of China (Grant No. PHR20110877).

REFERENCES