Solving the Large-Scale Vehicle Routing Problem by Location Based Heuristic & Genetic Algorithm

Cao Er-bao\textsuperscript{1,2}, Lai Ming-yong\textsuperscript{1,2}
1 School of Management, Hunan University, P.R.China, 410082
2 Hunan Provincial Laboratory of logistics Information and Simulation Technology, P.R.China, 410079

Abstract: The paper considers more than 10440 cigarettes shopkeepers of some logistics enterprise, and minimizes the total transit expenses. The large-scale vehicle routing problem (VRP) partitions to two sub-problems, one is generalized assignment problem and the other is vehicle routing problem intra-region after partitioning. The first problem is solved by the improved location-based heuristics (ILBH), and the hybrid genetic algorithm is presented for solving the second problem. The presented improved two-phase algorithm can obtain a good result effectively for solving the large-scale distribution transition problem in allowable time. The simulation results show that the algorithm is both effective and applicable, and it may extend to other cases.

Keywords: heuristics; the large scale vehicle routing problem; the logistics distribution; hybrid genetic algorithm

1. Introduction

Along with the development of the socialism market economy, the influence of the logistics--"the third profit headsprings"-- upon the economic activities is increasing obviously, drawing more and more attention, and has become “the most important competitive realm” at present. Logistics involves the movement of physical goods from one location to another and third-party transportation companies (such as JB. Hunt) provide a substantial portion of this service. Some of the earliest documentation of the use of logistics can be traced to the military. Several business groups have recently defined logistics for the private sector. American Production and Inventory Control Society (APICS)\textsuperscript{[1]} defines logistics as: “In an industrial context, the art and science of obtaining, producing, and distributing material and product in the proper place and in proper quantities”. The Council of Logistics Management (CLM)\textsuperscript{[2]} defines logistics as “The process of planning, implementing, and controlling the efficient, cost effective flow of raw materials, in-process inventory, finished goods and related information from the point of origin to the point of consumption for the purpose of conforming to customer requirements”. All of these definitions involve the movement of goods from one point to another. The business of goods movement or logistics has created the need for substantial infrastructures such as railroads, highways, river ports, seaports and airports. Cities and towns have emerged and grown along active logistics routes, demonstrating the importance and power of logistics.

Distribution is a partial logistics, and an integration of a big logistics within a small scope, which is the terminal of the logistics system. Distribution service activity of logistics confront the customers directly. So the degree of distribution affects satisfied degree of the logistics system directly. There exist many optimization decision problems for distribution, but among them, the VRP is a core problem, which is a complicated combination optimization problem. It needs to be solved with multi-objectives and within multi-restrictions, belonging to the NP hard problem. The current VRP research builds up mostly in suppose of abstract network diagram, does not base on the real city street environment. Although in the research, we need abstract the real street environment to the network topology diagram, in the converting process, we usually lose some useful information, which is related with concrete environment, and is the very thing concerned by the business enterprise. Though distribution has important function and significance, compared to the researches made abroad, the research situation in China is still stay around the academic aspect mostly. In actual application, the reports of using the

The research is supported by the national 15 significant projects-small city modern circulation essential engineering research and development, PR China [2003BA808A13].

Lai Ming-yong, Professor. Economics and trade college of Hunan University, 410079,laimingyong@hnu.net.
intelligence optimization system are rather few. The distribution characteristics of the merchandise retail trade, the catena industry, the third part logistics is as this: a large number of nodes with rather complicated location and disperse distribution but a small demand in amount and a rather high demand in time. So our main target is to get the allowable solution or the satisfaction solutions that can operate in allowable time.

When the scale and restrictions of VRP are smaller, the existing method can solve it well, especially when the total number of customers is under 100. In large-scale distribution situation, the model faces wider scope, thousands of customers, and a lot of restrictions in time and space, the existing theoretic algorithm is hard to resolve the actual problem satisfactorily. But this kind of large-scale routing problem is widespread in the daily necessity industries, such as: garbage collection, the milk collection and distribution, beverage distribution, cigarette distribution, etc. If talking about the model according to the general diagram of the vehicle routing problem, these customers directly are treated as nodes, the street in the city the arc, the scale of the problem will be very large, the difficulty of solving the problem will become greater, the credibility of calculation lower, and the calculation time longer. It can not be applied to the actual work. Therefore, we need to carry on the regional integration to the customers, reduce the scale of problems. Currently, most of the business enterprises, especially the Chinese business enterprises, to a large extent, solve the problem empirically. At this realm, only some few of scholars did the related research \[3,4,5\]. According to the different position, different quantity, and different demand of the customer network, in this paper, the large-scale vehicle routing problem (VRP) partitions two sub-problems: one is generalized assignment problem and the other is vehicle routing problem intra-region after partitioning. The first problem is solved by the improved location-based heuristics (ILBH); the hybrid genetic algorithm is presented for solving the second problem. The presented two-phase algorithm can obtain a good result effectively for solving the large-scale distribution transition problem in allowable time.

2. Model analysis

The vehicle routing problem (VRP) is a well-known NP hard problem. The basic vehicle routing problem (VRP) consists of a large number of customers, each with a known demand level, which must be supplied from a single depot. Delivery routes for vehicles are required, starting and finishing at the depot, so that all customer demands are satisfied and each customer is visited by just one vehicle. Vehicle capacities are given, and frequently, there is a maximum distance that each vehicle can travel. In the latter case, a drop allowance may be associated with each customer, which is added to the total distance traveled by the vehicle to which the customer is assigned. Thus, a vehicle that visits many customers will not be able to travel as far as a vehicle that visits relatively few customers. Possible objectives may be to end a set of routes that minimizes the total distance traveled, or minimizes the number of vehicles required and the total distance traveled with this number of vehicles. Various mathematical formulations of the VRP are given. In this paper, there is a distribution center has k vehicles which own the capacity as q, 0 stands for the distribution center and i stands for the customers (i=1,2,…l). \( y_{ui} \) is 1, if the vehicle k service for customer i, or else is 0; \( x_{ij} \) is 1, if the vehicle k from i to j, otherwise is 0. The paper improves the method by Julien Bramel and David Simchi-Levi\[6\], we divide the VRP into a general assignment problem p1 and a vehicle routing problem intra-region after partitioning p2.

P1: \[ \min z = \sum_k f(z_k) \]  
Subject to:
\[
\begin{align*}
\sum_i g_{ik} & \leq q, \forall k \\
\sum_k y_{ik} & = 1, i = 1, \ldots, l \\
y_{ik} & = 0 \text{or } 1, \ i = 0,1,\ldots l; \forall k
\end{align*}
\]

\( f(z_1) \) is decided by p2

\( \text{P2: } \min f(z_k) = \sum_{i} \sum_{j} c_{ij}^g x_{ijk} \)

Subject to:

\[
\begin{align*}
\sum_j x_{ijk} & = y_{ik}, f = 0,1,\ldots, n; \forall k; & (1) \\
\sum_j x_{ijk} & = y_{ik}, i = 0,1,\ldots, n; \forall k; & (2) \\
x_{ijk} & \in \{0,1\}, i, j = 0,\ldots, n; \forall k \\
\sum_{i,j} x_{ijk} & \leq |S| - 1, S \subseteq \{1,\ldots, n\}, 1 \leq |S| \leq n; & (3)
\end{align*}
\]

Constraints (1), (2) explain the relations between two variables, for each customer \( i \), only has two customers connect with it, the one reaches it directly, the other leaves directly; Restrictions (3) getting rid of the sub-back track solution.

3. An improved two-phase hybrid algorithm for the VRP

3.1 Basic definitions

Define 1. In each distribution district, there is a key node in each crossing.

Define 2. The street where the customers between two key-points lie is a key chord.

According to the inquisition, most of logistics enterprises although adopt the mode of the separation between visiting and sending, there exists the serious unbalance of workload in different routing. In some routing, the workload is heavy, but in some routing the mission is easy to finish, which affect seriously the employee’s enthusiasm and satisfaction. So we present the general workload to balance circuit workload.

Define 3. The general workload of the \( i \) routing is \( w_i \), which is the weighted mean value of traveling distances \(( k_i \)\), number of customers \(( x_i \)\), quantity of demands \(( y_i \)\), i.e.:

\[
w_i = m_i k_i + m_2 x_i + m_3 y_i, 0 < m_i < 1and \sum m_i = 1.
\]

For some presented in advance \( w_0 \), if the workload is balance in all routing, then for any \( \varepsilon \), we have \( |w_i - w_0| < \varepsilon \). Although \( w_0, w_i, \varepsilon \) have different units, they are comparable.

Define 4. In a circular district, if the shorter mileage can finish more nodes and orders, then we call that district as the key cluster, i.e., the key cluster should satisfy the following condition at the same time:

\[
\sum x_i \geq x_0, \sum y_i \geq y_0, \sum k_i \leq k_0 \text{ the parameter } x_0, y_0, k_0 \text{ are presented in advance according to the history.}
\]

3.2 First search phase: the improved location-based heuristics (ILBH)

In operational process, the distribution mission of the customers further from the distribution center is affected by more restrictions. The nearer the customer is, the fewer restrictions there are. But in order
to maintain and expand the business, the enterprise has to serve those further customers and have the initiative to consider the furthest customer from the distribution center. Making use of the geographic information system (GIS) in electronic map of the distribution district, we can know the shortest distance between two arbitrary nodes, and find the left and right key node for any node. Our idea is as follows: (1) According to the improved location based heuristics (ILBH), we need at least \[
\left\lfloor \sum_{i=1}^{q} g_{i} \right\rfloor + 1
\]
locations \([f] \) means the most integer that does not exceed \( f \), and have the initiative to consider the furthest node from the set including the distribution center and the already decided locations each time, establish it as location (the seed node), and make use of exploration to obtain the nearest key node from the location in turn, then the key nodes come into being key cluster (the parameter of the key clusters are presented in advance according to the actual demand), (2) take the key cluster as the center to expand to form the distribution circuit. (3) When the reserved workload is attained, stop expanding. (4) gathering the nodes that don’t lie in the circuit until all nodes are on line in the circuit. The practical operation adopts the principle of “close together merger, nodes form cluster, expand from the cluster”. To carry on “close together merger” means to connect two key nodes that belong in the same street, and store information in the key node, for example the key node \( A \), considering it as an array \( A \) (number of nodes, the gross demand quantity, the distance to the next key node; \( i = 1, 2, 3, 4 \) stand for the east, south, west, north four directions respectively); “nodes form cluster” means to make use of exploration in the diagram theory to obtain the from the locations, forming the key cluster that regards the locations as the center of a circles; “expand from the cluster” means to take key cluster as the center, make use of the exploration to search close nodes in order to expand outward, until the reserved workload of the vehicle is attained, and that key cluster and its surroundings nodes become a circuit finally, the same as other circuits. Obviously the number of circuits equal the number of locations.

3.3. Second search phase: the hybrid genetic algorithm

The optimization algorithm of the single circuit is as follows: (1) Using the nearest neighbor algorithm to obtain the first feasible solution, the first circuit set is formed; (2) Using genetic algorithm to optimize the first feasible solution, then carrying on the neighborhood operation. The genetic algorithm has a good global searching capacity, but a poor local searching ability. The neighborhood searching has effective local searching capacity. So we combine the respective advantage of them to form hybrid genetic algorithm. In this paper, we mainly consider a set of neighborhood structures \( NS = \{ N_{insert}, N_{opt}, N_{change} \} \).

When we obtain the first feasible solution, we only encode the key nodes, which can decrease the scale of problem, and then we have the first individual, then stochastically encode to obtain the first colony whose scale is \( N \). When we carry on the genetic operation, we consider the restriction checkout mechanism, i.e. no matter it is crossover operation and mutation operation, whenever we obtain a new individual we must check out the relative restrictions, such as: general workload, capacity and so on. We only withhold the solution which satisfies all the relative restrictions. The select operation combines the tournament selection model and elitist model.

4. Verification of case

Some logistics enterprise serves 10446 customers in 10000km\(^2\). According to the retails’ inventory, sale ability, and revolving funds, the enterprise’s distribution strategy is to carry on visiting and selling based on the fixed routing, then obtain the orders of customers. The vehicle routing problem is optimized according to the orders everyday. The result after optimizing is guidance for sorting and packing. The current distribution scheme is as follows: Dividing the customers of the whole city into 48 vehicle service districts; the capacity of the distribution vehicle is 150 unit merchandises, and the
longest work time are 8 hours. The enterprise wants to optimize the distribution strategy in order to
decrease the cost and improve the benefit when maintain the former distribution strategy. We apply the
above frame to program, solve the problem, and the solution is realized by computer. The brief step is as
follows: 1) The analyse of the foundation data: the enterprise provide the geographical site of customers
by geographic information system, obtain the position of each customers, and get the average demand.2)
the integration of the customers: aiming at the cantonal road condition and retailer’s dense degree,
inintegrating the customer to the corresponding routing by the improving laction based heuristics(ILBH).
3) The optimization of the vehicle routing problem: the paper adopts the hybrid genetic algorithm to
optimize the routing intra-region. Making use of the above frame to re-program, the enterprise’s
distribution cost descends significantly, and 25% of the transportation cost is saved, the reaction speed
to the order is raised by 30%, the annual benefit obtained directly can exceed to 6,000,000 dollars.

5. Conclusion
Making use of the improved two-phase algorithm, each district layout after dividing will be more
reasonable. The electronic map incarnates each distribution routing’s optimization, and with artificial
intervention adjustment sometimes. The optimum result can guide the sorting, packing, and distribution,
lowering the logistics cost. The comparative concentration of the geographic position can reduce the
total vehicle by 20% or so, and realize the goods-carrying rate to 85%. The workload of each circuit is
basal equilibrium, which can reduce the employee's complaint, promote their enthusiasm and
satisfaction so as to finish the work better. The logistics enterprise can realize really the separation
between visiting and sending with the above method. The presented two-phase algorithm can obtain a
good result effectively for solving the large-scale distribution transition problem in allowable time. The
simulation results show that the algorithm is both effective and applicable, and it may extend to other
cases.

References
[8] Alant Murray, Ross A.Gerrard. Capacitated service and regional constraints in location-allocation