Vehicle Routing for the Last-Mile Logistics Problem

Harshavardhan Desai, Liuhui Zhao, IEEE Member, Andreas Malikopoulos, IEEE Senior Member, Jackeline Rios-Torres, IEEE Member

Abstract—Fuel consumption is the major contributor associated with the large and growing part of transportation cost in logistics. Optimal vehicle routing approaches not only can provide solutions to reduce their operation costs but can also aim at addressing energy and environmental concerns. This paper outlines a solution to the single-depot capacitated vehicle routing problem with the objective of minimizing daily operation cost with a homogeneous fleet of delivery vehicles. The problem is solved using Simulated Annealing, to provide optimal routes for the vehicles traveling between the depot and destinations. Simulation results demonstrate that the proposed approach is effective to recommend optimal route and reduce operation cost.

I. INTRODUCTION

A. Motivation

Last mile logistics refers to the last portion of a supply chain involving the transportation of people or goods from the last transportation hub to the final destination. With the meteoric rise of the e-commerce industry, last mile delivery, especially parcel delivery has attracted considerable attention [1]. On the other hand, due to overcrowding and congestion in many metropolitan cities, more and more people are looking for ride-sharing as a viable form of transport for their daily commute [2]. In this highly competitive environment, third party logistics and last mile delivery firms must not only be able to meet ever increasing fulfilment deadlines but do so as efficiently as possible to maximize profits. Some major hurdles in maximizing profits are ensuring the utilization of the vehicle inventory to its maximum potential and cost-effective routing of vehicles owing to growing emission constraints coupled with the strive to reduce fuel consumption.

In this paper, we present a last-mile logistics system which combines the transport of people using ride-sharing with parcel delivery.

B. Literature Review

The vehicle routing problem (VRP) is a variation of the extensively investigated traveling salesman problem. There has been significant amount of work done in the area of eco-VRP. Among different approaches that have been reported in the literature, time-dependent VRP (TD-VRP) is based on the notion that the travel time between any pair of points (customers and depots) depends either on the distance between the points, or on the time of day (e.g., rush hours). The feature of fluctuating travel duration enables VRP to account for the actual conditions such as urban congestion, where the traveling speed is not constant due to variation in traffic density. Therefore, TD-VRP is a relevant and useful model to reveal recurring traffic congestion and to explore approaches to avoid it. In the model described by Qian and Egles [3], the speed of the traffic on the underlying road network is time dependent and the path used by a vehicle between a pair of customers is the decision variable. The authors proposed a Tabu based algorithm to solve the problem and concluded that allowing a specified waiting time at customer nodes, vehicles can avoid being caught in congestion, thus leading to overall fuel consumption reduction. Yao et al. [4] further explored the TD-VRP by introducing the concept of alternative stop points assuming that a delivery vehicle could temporarily stop at the opposite side of the client and then the deliveryman walks across the road to serve the client. By enabling alternative stops, detouring could be avoided in vehicle routing. As a result, the study showed that vehicle miles traveled as well as total fuel consumption were reduced in the network. Huang et al. [5] presented an approach with the objective of finding an optimal routing solution such that vehicle arrival times at nodes meet the deadlines specified by the clients.

Bent and Hentenryck [6] considered the partially dynamic VRP with time windows. The goal was to serve as many customers as possible given a fixed number of vehicles. Moreover, stochastic information was assumed to be available on the dynamic customers. To tackle this dynamic stochastic

This research was supported in part by the Laboratory Directed Research and Development Program of the Oak Ridge National Laboratory, Oak Ridge, TN 37831 USA, managed by UT-Battelle, LLC, for the US Department of Energy (DOE), and in part by the Delaware Institute (DEI).

Desai, Zhao and Malikopoulos are with the Department of Mechanical Engineering, University of Delaware, Newark, DE 19716 USA (e-mail: harshd@udel.edu; lhzhao@udel.edu; andreas@udel.edu)

Rios-Torres is with the Energy and Transportation Science Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831 USA (e-mail: riostorres@ornl.gov)

This manuscript has been co-authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doepublic-access-plan).
vehicle routing problem, the authors proposed a multiple
scenario approach, which involves continuously generating
and solving scenarios with different static and dynamic
requests, thereby generating an optimal routing plan.

In the era of sharing mobility, the ride-sharing problem has
been widely studied. Furuhata et al. [7] reviewed that key
aspects of existing ride-sharing systems along with some of
their key challenges (e.g., the design of attractive and price
effective mechanisms), are a concierge like ride-arrangement
customer preferences, multi-modal rides, and building of trust
among unknown travelers. Agatz et al. [8] surveyed the
characteristics, objectives and optimization challenges of
different classes of operations research models related to ride-
sharing. The cost of the ride-share trip should be
proportionately divided among the participants, roughly
proportional to vehicle-miles travelled. Agatz et al. [9]
proposed a framework allocating the costs of a joint trip
proportional to the distances travelled if separate trips were
taken. Kleiner et al. [10] proposed an auction-based
mechanism to determine the driver’s compensation.

Wang et al. [11] presented a model which utilizes crowd-
workers for last-mile parcel delivery. The concept was such
that each parcel will be sent to a pick-own-parcel station
nearest to its consumer’s address, then assigned to the crowd-
workers via a mobile app. Each worker was associated with a
travel pattern (e.g., the driver’s daily commute route) and the
reward was proportional to the distance of the detour taken
to make a delivery. The objective was to assign the parcels to the
most convenient workers to minimize the total reward paid by
the company. Compared to traditional delivery methods, the
crowdsourcing approach resulted in a higher level of
parallelism in job execution since the fleet size of delivery
vehicles is much larger and each worker handles only a small
number of parcels. Communication between workers and
customers is also more effective. Due to the elimination of
vehicles specifically used for last-mile delivery, this approach
helped reduce operations costs as well as carbon emissions,
since the number of vehicles in the network were reduced.

A. Contribution of this paper

Both ride-sharing and last-mile delivery are topics that have
been extensively studied with either static or dynamic settings.
However, the concept of shared logistics is this area that has
not been fully investigated. The main objective of this paper is
to determine how a fleet of vehicles can be dispatched most
efficiently to maximize the daily profit of a company which
aims at fulfilling the last/first-mile connection of passengers
and simultaneous parcel deliveries from a major transport hub.

B. Organization of the paper

The structure of the paper is organized as follows. In Section
II, we formulate the problem. In Section III, we provide
the solution and simulation results, and finally, in Section IV, we
draw concluding remarks.

II. PROBLEM FORMULATION

Consider a major transit hub in a city, which serves as a
consolidation center for parcels to be delivered to residents in
the neighbourhoods. We consider a first/last mile
transportation service for people stationed around the hub as
well as for the delivery of parcels in the same area. Electric
autonomous vehicles owned by the taxi/ride hailing company
are used for simultaneous transport and delivery. A
representation of the problem is shown in Fig. 1.

A. Problem Description

Transportation of People: At the beginning of the operation
time, all vehicles are located at the hub where parking is free
for the service company. Similarly, vehicles will all come back
to the transport hub at the end of the operation period. This
guarantees that the vehicle fleet has enough time (over night)
to fully charge the battery and any required maintenance can
be performed.

Parcel Delivery: With the model presented below, the service
company only needs to handle the delivery of parcels to the
hubs nearest to the consumer’s address after which, the parcels
at the hub will be assigned to the vehicle and eventually reach
the consumers. A parcel will be assigned to the vehicle which
will incur the least cost to deviate from its assigned route to
make the delivery.

Figure 1: Last-Mile City Logistics Problem.

For example, consider a vehicle travelling from node A to
node B. For simplicity, assume that the cost, \( r \), of deviating
from the path and making a delivery at node D on the way, is
proportional to the additional travel distance

\[
r = d(A, D) + d(D, B) - d(A, B)
\]
where \( d(\cdot, \cdot) \) is the Euclidean distance between two locations, \( d(A, D) + d(D, B) \) is the distance incurred for the delivery, and \( d(A, B) \) is the travel distance if the vehicle does not take the task.

**Routing:**

Consider a network with \( N \) nodes and an operation period with \( T \) time instances:

\[ N = \{1, \ldots, i, \ldots, N\} \]: Set of passenger nodes.

\[ T = \{1, \ldots, t, \ldots, T\} \]: Set of time instants in the service period. The time between two consecutive time instants is considered to be one time step where the number of time steps in a day is \( T \).

\( \delta_{ij} \): Travel time between node \( i \) and node \( j \), \( \forall i, j \in N, i \neq j \);

\( \delta_i \): Travel time between hub and node \( i \);

\( \delta_{\text{max}} \): the maximum travel time from the transport hub to any of the nodes:

\[ \delta_{\text{max}} = \max(\delta_i) \ \forall i \in N. \]

**Routes:** A route indicates a routine for a vehicle that starts at the hub, serves some number of customers, and then returns to the depot. Formally, a route is a sequence \([0, v_1, \ldots, v_n, 0]\), where \( 1 \leq v_i \leq N \) and all \( v_i \) are distinct. The demand of a route is denoted by \( Q(r) = \sum_{i=1}^{n} Q_i \). The travel cost of a route \( r \) is denoted by \( c(r) \) and is the cost of visiting all of its customers, i.e., \( c(r) = c_{0v_1} + c_{v_1v_2} + \cdots + c_{v_{n-1}v_n} + c_{v_n0} \).

**Routing Plan:** A routing plan, or a plan for short, is a set of routes \( \{r_1, \ldots, r_m\} \) serving each customer exactly once. A routing plan assigns a unique successor and predecessor for each customer. For a plan, \( \alpha \), the successor of customer \( i \) is denoted by \( \text{succ}(r, \alpha) \) and the predecessor is denoted by \( \text{pred}(r, \alpha) \). The travel cost of a plan is denoted by \( c(\alpha) = \sum_{i=1}^{n} c(r) \). We also use \( \text{cust}(r) \) and \( \text{cust}(\alpha) \) to denote the customers of a route \( r \) and a plan \( \alpha \).

In the modelling framework above, the following assumptions are imposed:

**Assumption 1:** A vehicle can accommodate up to two passengers and one parcel, or one passenger and 3 parcels at a time.

**Assumption 2:** The vehicles are fully charged overnight at the transport hub. These assumptions consider the capacity constraint for each vehicle and ensure that the trip is short enough to be completed on a single charge. They also ensure that the vehicles are fully charged at the start of the operational period.

**B. Objective function**

The revenue earned from carrying passengers from the transport hub to their destinations is

\[ P \cdot \sum_{i \in N, t \in T} D_i^{t+\delta_i} \cdot d_i \]

where, \( P \) is the price per driving distance (\$/mile), \( d_i \) is the travel distance between hub and node \( i \), \( \forall i \in N \), \( D_i^{t+\delta_i} \) is the number of trips satisfied from the transport hub to service node \( i \) from time instant \( t \) to time instant \( t + \delta_i \), \( \forall i \in N, \forall t \in T, t + \delta_i \leq T \).

The revenue earned from carrying passengers from their origins to the transport hub can be calculated by:

\[ P \cdot \sum_{i \in N, t \in T} D_i^{t+\delta_i} \cdot d_i \]

where, \( D_i^{t+\delta_i} \) is the number of trips satisfied from service node \( i \) to the transport hub from time instant \( t \) to time instant \( t + \delta_i \), \( \forall i \in N, \forall t \in T, t + \delta_i \leq T \).

Vehicle depreciation costs are calculated by:

\[ C_d \cdot \sum_{i \in N, j \in T} U_{ij}^{t+\delta_i} \cdot d_{ij} \]

where, \( C_d \) is the vehicle depreciation costs per mile (\$/mile), \( U_{ij}^{t+\delta_i} \) is the number of vehicles travelling from node \( i \) to node \( j \) from time instant \( t \) to time instant \( t + \delta_i \), \( \forall i, j \in N, i \neq j, \forall t \in T \).

Vehicle maintenance costs are calculated by:

\[ C_m \cdot F \]

where, \( C_m \) is the maintenance cost per vehicle per day (\$/day), \( F \) is the vehicle fleet size in the system.

Detour cost for delivering parcels is calculated by:

\[ \sum_{i \in N, j \in T} x_{ij} \cdot d_{ij} \cdot C_p \]

where, \( x_{ij} \) is equal to 1 if there is a vehicle travelling from node \( i \) to node \( j \), otherwise 0, \( \forall i \in N \), \( d_{ij} \) is the travel distance between node \( i \) and node \( j \), \( \forall i \in N, C_p \) is the detour cost per mile (\$/mile).

The following objective function maximizes the total profit, \( G \), during a typical day of operations, considering the revenues earned and the aforementioned costs.
\[
\text{Max } G = P \left( \sum_{i \in N, t \in T, t + \delta_i \not\in T} D_i^{t,t+\delta_i} \cdot d_i + \sum_{i \in N \cap T, t + \delta_i \in T} D_i^{t,t+\delta_i} \cdot d_i \right) - C_{\text{di}} \sum_{i \in N, j \in T} U_{ij}^{t,t+\delta_{ij}} \cdot d_{ij} - C_{\text{m}} \cdot F - \sum_{i \in N, i \neq j} x_{ij} \cdot d_{ij} \cdot C_{\text{p}},
\]

where, \( M \) is a large number.

We also need to ensure that if no trip from node \( i \) is satisfied, then that node is not selected, hence

\[
x_i \leq \sum_{t \in T} D_i^{t,t+\delta_i} + \sum_{t \in T} D_i^{t,t+\delta_i} \forall i \in N.
\]

Next, we need to include a constraint that imposes the condition that the number of vehicles travelling between service node \( i \) and the transport hub must be greater than or equal to the number of people travelling on that OD route

\[
D_i^{t,t+\delta_i} \leq U_{ij}^{t,t+\delta_{ij}} \forall i \in N, \forall t \in T, t + \delta_i \leq T.
\]

The following constraint guarantees that the vehicles leaving a service node at time instant \( t \) is less than or equal to the available vehicles at that service node, namely

\[
\sum_{j \in N, i \neq j} U_{ij}^{t,t+\delta_{ij}} \leq V_i^t \forall i \in N, \forall t \in T.
\]

Finally, the next set of the constraints define the domain for the decision variables.

\[
V_i^t \geq 0 \forall i \in N, \forall t \in T.
\]

\[
D_i^{t,t+\delta_i} \geq 0 \forall i \in N, \forall t \in T, t + \delta_i \leq T.
\]

\[
D_i^{t,t+\delta_i} \geq 0 \forall i \in N, \forall t \in T, t + \delta_i \leq T,
\]

\[
x_{ij} \in (0,1) \forall i \in N.
\]

### III. Solution Approach and Simulation Results

Metaheuristics have been introduced into the solutions for VRP in the last two decades, which are generally recognized to fit combinatorial optimizations. Simulated Annealing, Tabu Search, Genetic Algorithm and Ant Colony Optimization have been tried to apply to VRP.

Simulated Annealing (SA) is inspired by the process of annealing in metallurgy. In this natural process a material is heated and slowly cooled under controlled conditions to increase the size of the crystals in the material and reduce their defects. The heat increases the energy of the atoms allowing them to move freely, and the slow cooling schedule allows a new low-energy configuration to be discovered and exploited. Similarly, each configuration of a solution in the search space represents a different internal energy of the system. Heating the system results in a relaxation of the acceptance criteria of the samples taken from the search space. As the system is cooled, the acceptance criteria of samples are narrowed to focus on improving movements. Once the system has cooled, the configuration will represent a sample at or close to a global optimum.

The main idea of a SA algorithm is to occasionally accept degraded solutions in the hope of escaping the current local
The information processing objective of the technique is to locate the minimum cost configuration in the search space. The algorithm’s plan of action is to probabilistically re-sample the problem space where the acceptance of new samples into the currently held sample is managed by a probabilistic function that becomes more discerning of the cost of samples it accepts over the execution time of the algorithm. The pseudocode for the SA algorithms used for the VRP is given in Algorithm 1.

To evaluate the performance of the proposed approach, ten computational experiments were carried out on a network spanning 200 sq. km with varying number of customer locations as well varying fleet sizes. The number of customer locations vary from 8 to 70, out of which two-thirds of the locations are passenger locations and the rest are parcel delivery locations. The capacity of a vehicle is limited to 5 units, where each passenger takes 2 units and each parcel takes 1 unit.

For the baseline scenario, we have considered a combination of two conventional last mile transportation system, which consists of two separate fleets of vehicles, one for public transport (ride-sharing) and one for parcel delivery. The operational costs in this scenario would be the total costs of operating both fleets simultaneously. For simplicity, we consider the vehicle type and capacity of the vehicles of both fleets to be the same. Keeping the vehicle capacity in mind, in the baseline scenario, the ride-sharing fleet size varies from 3 to 24 vehicles, and the delivery fleet size varies from 1 to 5 vehicles, proportionate to the number of customer locations, bring in the total vehicle count to 4 to 29 vehicles.

With the integrated service, the fleet size varies from 3 to 24 vehicles, proportionate to the number of customers (i.e. 8 to 70 customer locations). We use SA to find the cost of every possible route that can be taken for every instance, until the minimum cost (and its corresponding vehicle routes) is obtained. The routing plan which leads to the lowest cost is taken as the optimal solution. The computation of the SA was executed on MATLAB running on an Intel Core i7-4510U 2.0GHz CPU running Windows 10.

The experimental results for the ten instances are presented in Table 1. It can be observed that the total cost of service is roughly proportional to the number of customer locations and fleet size. For example, the cost of serving 25 customer locations using 9 vehicles (instance 5) is higher than the cost of serving 20 customer locations using 7 vehicles (instance 4). But it is also observed that the cost of servicing 10 customers using 4 vehicles (instance 2) is greater than the cost of servicing 14 customers using 5 vehicles (instance 3). This behavior can be due to fact that in instance 2, one of the vehicles serves only one customer, indicating that the fleet is underutilized. The optimal routes for instances 3, 5, 6, 8 and 10 are shown in Fig. 5-9, in which we see that in each route, each vehicle visits two passenger locations for drop-off or pick-up, as well as a parcel delivery location before returning to the depot.

Table 1: Optimal results for ten instances

| Instance | No. of Vehicles | No. of Passenger Nodes | No. of Delivery Nodes | Total no. of Customer Nodes | Baseline Cost ($USD) | Cost with Integrated Service ($USD) | Reduction in Cost (%) |
|----------|----------------|------------------------|----------------------|-----------------------------|----------------------|------------------------------------|----------------------|
| 1        | 3              | 6                      | 2                    | 8                           | 548.26               | 220.16                             | 59.84                |
| 2        | 4              | 7                      | 3                    | 10                          | 454.72               | 292.21                             | 35.74                |
| 3        | 5              | 10                     | 4                    | 14                          | 671.85               | 275.93                             | 58.93                |
| 4        | 7              | 14                     | 6                    | 20                          | 494.06               | 332.58                             | 32.68                |
| 5        | 9              | 17                     | 8                    | 25                          | 639.75               | 371.26                             | 41.97                |
| 6        | 11             | 20                     | 10                   | 30                          | 690.06               | 410.28                             | 40.54                |
| 7        | 14             | 27                     | 13                   | 40                          | 692.06               | 470.97                             | 31.95                |
| 8        | 17             | 34                     | 16                   | 50                          | 738.53               | 506.50                             | 31.42                |
| 9        | 21             | 40                     | 20                   | 60                          | 743.69               | 518.80                             | 30.24                |
| 10       | 24             | 47                     | 23                   | 70                          | 805.02               | 540.95                             | 32.80                |
We observe that compared to the baseline scenario, the integrated service reduces operational costs by a considerable amount. Cost reduction ranges from 32.8% during high demand (instance 10), to 59.84% during low demand (instance 1). Take instance 3 for example: Since the delivery locations are spread out, the sole delivery vehicle must travel a large amount to make all 4 deliveries, resulting a high operating cost ($414.25). The ride-sharing fleet costs amounted to $257.60 to cater to 10 passengers, bringing the total cost of using both fleets to $671.85. The vehicle routes in instance 3 using the baseline scenario is shown in Fig. 1. Comparatively, in the integrated service, when the ride-sharing vehicles are used to make the deliveries, the cost of operating the delivery vehicle is eliminated, and is offset by a marginal increase in the operating costs of the ride-sharing fleet, since they must travel a bit more to make the nearby delivery. The operating cost for this instance was $275.93, which is 58.93% lower than the baseline scenario.

Figure 3: Baseline Scenario for instance 3 (Ride-sharing fleet)

Figure 4: Baseline Scenario for instance 3 (Delivery fleet)

Figure 5: Optimal Routes for instance 3

Figure 6: Optimal Routes for instance 5

Figure 7: Optimal Routes for instance 6

Figure 8: Optimal Routes for instance 8
IV. CONCLUSIONS

The VRP is an interesting problem not only for distribution centers but also for shared mobility services. This problem has been addressed in many research papers but combining the two areas of VRPs has not been fully investigated. This paper proposes an efficient vehicle routing planning scheme with the objective of minimizing the daily cost of operations of a fleet of vehicles, which is used to ferry passengers to their final destination and simultaneously deliver parcels to nearby destinations. The effectiveness of the approach is tested on a benchmarking network using a SA algorithm and the results indicate that the proposed approach is valid, reliable and has good computing performance. It is shown that the integrated service can reduce operational costs by up to 59%, compared to conventional last mile transportation services, depending on the customer demand. In the computational model, only a few factors were considered, namely demand at passenger and delivery locations, vehicle capacity and transportation costs.

As it is known, fuel efficiency is highly sensitive to the driving cycle, this method can be further enhanced with dynamic traffic flow data so that fuel consumption can be reduced. Finally, this approach can be modified to accommodate a heterogeneous delivery fleet (i.e. vehicles with different cargo carrying capacities as well as vehicles with different kinds of powertrains).

REFERENCES

[1] S. Monohan and M. Hu, “US E-Commerce Trends and the Impact on Logistics,” A.T. Kearney.
[2] S. Burgstaller, D. Flowers, D. Tamberrino, H. Terry, Y. Yang, “Rethinking Mobility,” Equity Research, Goldman Sachs, May 2017.
[3] J. Qian, R. Eglese, “Fuel emissions optimization in vehicle routing problems with time-varying speeds,” European Journal of Operational Research, Volume 248, Issue 3, February 2016, Pages 840-848.
[4] E. Yao, Z. Lang, Y. Yang, Y. Zhang, “Vehicle routing problem solution considering minimising fuel consumption,” IET Intelligent Transport Systems (Volume: 9, Issue: 5, 6 2015) Pages: 523 – 529.
[5] Y. Huang, B. Lin, V. Tseng, “Efficient Multi-Destinations Route Planning with Deadlines and Cost Constraints,” Mobile Data Management (MDM), 18th IEEE International Conference, 2017.
[6] R. Bent, P. Hentenryck, “Scenario-Based Planning for Partially Dynamic Vehicle Routing with Stochastic Customers,” Operations Research Volume 52, Issue 6, 2004.

[7] M. Furuhata, M. Dessouky, F. Ordóñez, M. Brunet, X. Wang, S. Koenig, “Ridesharing: The state-of-the-art and future directions,” Transportation Research Part B, Volume 57, Pages 28-46, 2013.
[8] N. Agatz, A. Erera, M. Savelsbergh, X. Wang, “Optimization for dynamic ride-sharing: A review,” European Journal of Operational Research, Volume 223, Issue 2, Pages 295-303, 2012.
[9] N. Agatz, A. Erera, M. Savelsbergh, X. Wang, “Dynamic ride-sharing: A simulation study in metro Atlanta,” Transportation Research Part B, Volume 45, Issue 9, Pages 1450-1464, 2011.
[10] A. Kleiner, B. Nebel, V. Ziparo, “A mechanism for dynamic ride sharing based on parallel auctions,” In: Proc. of the 22th International Joint Conference on Artificial Intelligence (IJCAI), Barcelona, Spain, Pages 266–272, ISBN 978-1-57735-516-8, 2011.
[11] Y. Wang, D. Zhang, Q. Liub, F. Shen, L. H. Lee, “Towards enhancing the last-mile delivery: An effective crowd-tasking model with scalable solutions,” Transportation Research Part E, Volume 93, Pages 279-293, 2016.
[12] X. Liang, G. Correia, B. Arem, “Optimizing the service area and trip selection of an electric automated taxi system used for the last mile of train trips,” Transportation Research Part E, Volume 93, Pages 115-129, 2016.