Routing Strategies for Efficient Deployment of Alternative Fuel Vehicles for Freight Delivery

Final Report

METRANS Project 14-23

Feb 2017

Principal Investigator: Maged M. Dessouky

Ph.D. Candidate: Yihuan (Ethan) Shao

Daniel J. Epstein Department of Industrial and Systems Engineering

University of Southern California

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, and California Department of Transportation in the interest of information exchange. The U.S. Government and California Department of Transportation assume no liability for the contents or use thereof. The contents do not necessarily reflect the official views or policies of the State of California or the Department of Transportation. This report does not constitute a standard, specification, or regulation.

Abstract

With increasing concerns on environmental issues, recent research on Vehicle Routing Problems (VRP) has added new factors such as greenhouse gas emissions and alternative fuel vehicles into the models. In this report, we consider one such promising alternative fuel vehicle, Compressed Natural Gas (CNG). However, due to the limited number of available fueling stations and small fuel tank capacity, CNG trucks face several challenges on their way to replacing traditional diesel trucks. Even though CNG trucks have advantages on less greenhouse gas emissions and cheaper fuel cost, the detours to the fueling station may increase the total travel distance.

We introduce the CNG Truck Routing Problem with Fueling Stations (CTRPFS) to model decisions to be made with regards to the vehicle routes including the choice of fueling stations. Moreover we consider load capacity, fuel tank capacity and the driver's daily traveling distance limitation. We develop a Mixed Integer Programming (MIP) model with preprocessing and valid inequalities to solve the problem optimally. A hybrid heuristic method is also proposed to solve this problem, which combines an Adaptive Large Neighborhood Search (ALNS) with a local search and a MIP model.

In the numerical experiments section, we solve a set of small instances to show the efficiency of our preprocessing and valid inequalities. The optimal values from MIP models are used as benchmarks to show our ALNS can achieve high quality solutions. We run experiments on large instances to explore some insights on the number of CNG fueling stations and tank capacities. Based on the estimation of daily goods delivery data from the Ports of Los Angeles and Long Beach, we conduct experiments comparing the total traveling distance using CNG trucks with the current use of diesel trucks for local delivery.

Disclosure

The project was funded in entirety under U.S. Department of Transportation. The total funding amount was $90,000. The Principle Investigator of the project was Professor Maged M. Dessouky of the Viterbi School of Engineering at the University of Southern California and was carried out during the period: 05/20/2014 - 12/31/2016.

Acknowledgement

We would like to thank METRANS for funding this research.

Contents

List of Figures

List of Tables

1 Introduction

Transportation plays an important role in a modern society. It is essential for economic development. Globalization and materialization of the economy have resulted in an increasing demand in transportation. However transportation also brings in some problems such as noise and congestion, with resulting negative impact on the environment. In the United States, in 2013, the transportation end-use sector accounted for a large part of CO_2, CH_4 and N_2O emissions from fossil fuel combustions. Among various transportation sources, light duty vehicles was a main factor. It increased 35% in the number of vehicle miles traveled and represented 60% of CO_2 emissions from fossil fuel combustions in 2013 [5]. All these factors call for a better planning on transportation.

Transportation has many facets at multiple decision making levels, with one of the most famous problems is the well-known Vehicle Routing Problem (VRP) [4]. VRP first appeared in a paper by Dantzig and Ramser in 1959 [3]. It solves a routing problem for trucks to meet the demands from customers located in different places while minimizing the total mileage traveled by the fleet. Nowadays, with increasing concerns on the sustainability of the logistics system, current research on VRP has received more interest on environmental issues, often referred as the Green Vehicle Routing Problem (GVRP). GVRP is an emerging area and there are growing publications on it in recent years [4] [12]. The efforts made on GVRP are mainly divided into two directions, changing people's behaviors and exploring alternative fuels [6]. The first direction aims at reducing the number of vehicle miles traveled. It covers topics such as ridesharing and bus scheduling. The second direction focuses on reducing emissions per traveled mile. It introduces clean energy technologies like biofuel, electric vehicles etc.

This research focuses on the second direction. So far, various kinds of green energies like electricity, ethanol and hydrogen have been used for alternative fuel vehicles. However all of them have to face different kinds of problems such as short driving range, limited fueling stations and slow refueling speed. Among those alternative fuel vehicles, the Compressed Natural Gas (CNG) truck has a promising future since it has low greenhouse gas emissions and fuel cost. With the increased emphasis on sustainability, several logistics companies have purchased alternative fuel vehicles like CNG trucks to support their delivery operations. However CNG trucks still face several challenges on their way to replacing traditional diesel trucks. One of the biggest challenges is the sparse, uneven allocation of CNG fueling stations, which may lead to a long detour for refueling. In fact, the detour to a CNG fueling station could increase overall energy consumption and finally offset the low emission benefit from the CNG fuel. Additional challenges exist like limited tank capacity and small load capacity [6] [15]. Indeed, many of the local freight delivery operations are performed by independent truckers who are too small to manage their own fueling station and the public fueling stations are the only option for them. Empirical research shows drivers do consider station locations and convenient location is the most important reason for choosing a fueling station [9]. Thus, it becomes imperative to develop efficient deployment strategies for refueling for these trucks that will have minimal impact on their routing operations. We introduce the CNG Truck Routing Problem with Fueling Stations (CTRPFS) to model decisions to be made with regards to the vehicle routes including the choice of fueling stations. The CTRPFS is a special case of the well-known VRP. It combines classical restrictions such as capacity, daily traveling distance as well as new constraints for CNG trucks, such as fuel tank capacity.

This report is organized as follows. In Section 2 a review of the related literature and our contribution is presented. Section 3 describes the CTRPFS problem and a general MIP formulation for solving the problem. Section 4 gives a heuristic solution method to solve the problem. Section 5 presents the computational results of benchmark instances as well as an estimated data set from the Southern California Association of Government for the Ports of Los Angeles and Long Beach. We conclude the report and give future research directions in Section 7.

2 Literature Review

Since first proposed by Dantzig and Ramser in 1959 [3], VRP has been a hot topic in combinatorial optimization in the last decades. It solves the problem of designing optimal delivery routes from one or several depots to a number of scattered cities or customers [11]. The classical VRP has been extended to a number of variants with respect to different real-world constraints. Capacitated VRP (CVRP) is a basic variant of VRP, where each city or customer has a deterministic, non-negative and indivisible demand to be satisfied, and each vehicle has an identical load capacity as well as a route length constraint [18]. Our CTRPFS belongs to the category of VRP for alternative fuel vehicles. Compared with the traditional diesel or gasoline vehicles, the sparse allocation of fueling stations and small fuel tank capacity of alternative fuel vehicles call for an extra attention on the refueling process. Traditional vehicle routing problems assume that the fuel is adequate for covering the whole tour and the vehicles can be refueled any time. However, such an assumption may not hold for CNG trucks, or other alternative fuel vehicles like electric vehicles. Since VRP for alternative fuel vehicles is a relatively new area, the existing research about this topic is limited. Most of the studies so far use heuristic methods and some of them give MIP formulations as their exact solution methods. Even for those who give MIP formulations, only cases with very small size can be solved optimally. Based on the objectives of the problems, current research on this topic can be divided into two groups. One group of research is to solve a routing problem while minimizing CO_2 emissions or fuel consumption. The other group of research is to solve the recharging or refueling problem.

Jemai et al. [8] propose a method to find routes for green vehicles to serve a set of customers while minimizing the total traveling distance and the CO_2 emissions. A heuristic method called NSGA-II is used to search the Pareto frontier of the bi-objective problem. Xiao et al. [19] develop a Simulated Annealing method with a hybrid exchange rule to find a route with lower fuel consumption.

Erdogan and Miller-Hooks [6] extend CVRP to GVRP where the routing problem for alternative fuel-powered vehicle fleets is solved. Limited by its tank capacity, an alternative fuel vehicle has to visit one or more times the fueling stations during its tour to serve the customers. Two construction heuristic methods and a customized improvement technique are developed to minimize the total distance traveled by the vehicles. Schneider et al. [16] present a hybrid heuristic that combines a variable neighborhood search algorithm with a Tabu Search to solve an electric vehicle routing problem. A feasible route for this problem needs to satisfy the recharging, time windows and load capacity constraints. The hybrid heuristic is tested on newly designed instances based on the traditional cases, as well as special benchmark cases from other papers like Erdogan and Miller-Hooks [6]. These tests show that hybridization has a positive effect on the solution quality. Hiermann et al. [7] propose a more complex model which combines the use of electric vehicles and the optimization within a heterogeneous fleet. They develop a hybrid heuristic which combines an ALNS with an embedded local search and labeling procedure for intensification.

In this research, our contribution for the alternative fuel vehicle routing problem is we first introduce the CTRPFS to model decisions to be made with regards to the vehicle routes including the choice of fueling stations. A fixed fueling cost, which may represent the waiting time at the fueling station, is considered for the fueling procedure. This fixed refueling cost is unique for CNG trucks and few papers add it into their models. Although only few researchers touch this topic, the motivation for fixed refueling cost comes from the real world. Users may wait for hours in long queues to get their trucks refueled at CNG fueling stations [10]. Meanwhile the fixed fueling cost can have a large influence on the solution method. A fueling station with a shorter detour may not lead to a better solution. We need to balance the detour and the frequency of visiting fueling stations. In order to find exact solutions for the problem, we provide a MIP model. Our contribution is to introduce a new decision variable setting method with a variable elimination method and valid inequalities. Small instances are solved via CPLEX and serve as benchmarks for the heuristic method. To deal with instances in realistic size, we propose a new hybrid heuristic method which combines an ALNS with a local search and MIP model. Compared with other methods, our unique embedded MIP model can solve a problem for a more general situation. In computational experiments, we modify benchmark instances from Augerat et al. [1] with constraints from CNG trucks. The computational experiments explore some insights. Based on the estimation of daily goods delivery data from the Ports of Los Angeles and Long Beach, we conduct experiments comparing the total traveling distance using CNG trucks with the current use of diesel trucks for local delivery. The purpose of these experiments is to show how much the losses of route efficiency for CNG trucks can be reduced with an efficient deployment.

3 Problem Statement and Formulation

This section first describes the problem, and then a MIP model is presented to optimally solve the problem. In order to improve the solving speed, variable elimination and valid inequalities are provided.

3.1 Problem Statement

CTRPFS is designed to solve the vehicle routing problem for a distributor that employs a fleet of CNG trucks to daily deliver products to customers. Customers are widely distributed in a region and demands from customers are known ahead of time. We assume the distributor's depot does not own CNG fueling stations by itself, which means its trucks have to use the sparsely distributed public CNG fueling stations during their delivery routes. Thus we need to develop routing plans for the fleet which minimizes the total traveling distance while meeting the fuel tank capacity constraints. A typical route for the fleet may start from a departure depot, visit some customers and CNG fueling stations if necessary, and return back to the departure depot at the end of the day. The main objective for this problem is to minimize the total cost, while keeping the fleet size as small as possible.

The total cost mainly comes from two sources. One is the traveling cost, which is directly proportional to the fleet's daily traveling distance. The other is the fixed refueling cost from the fueling process. It occurs every time when the truck goes to a fueling station. As previously mentioned, this fixed fueling cost could be the waiting time for the availability of the pump. In some very special case, the optimal route for a truck may contain two successive visits of fueling stations. We ignore such a situation and assume a truck will not visit fueling stations successively. For the fueling time, we assume it is a linear function of the volume to be fueled. Thus the fueling time cost can be included in the distance cost. To simplify the model, we do not explicitly separate the variable fueling time from the driving time since we assume fuel consumption is a linear function of traveling distance. Thus in order to measure how much fuel is left for a truck, we only need to calculate the cumulative distance after its last refueling station. Initially, all trucks are located at the depot and their tanks are assumed full. Another assumption is all of the trucks in the fleet are homogeneous and each truck has three kinds of constraints which are load capacity, total daily traveling distance and fuel tank capacity constraints.

3.2 Mixed Integer Programming Model

An instance network of CTRPFS consists of a set of vertices $\mathbb{N}$ and a set of arcs $\mathbb{A}$. Vertices in set $\mathbb{N}$ are marked from 0 to n. Vertices v_0 and v_n are the depot and the rest of the vertices are divided into two parts, customers' vertices $\mathbb{C}$ and fueling stations vertices $\mathbb{F}$. Though v_0 and v_n are different vertices in our model, in fact they could be the same location. An arc $a^k_{i,j}$ represents a route from vertex v_i to vertex v_j for truck k in fleet $\mathbb{M}$. The fleet has m homogeneous trucks and all of them need to depart from vertex v_0 and arrive at vertex v_n at the end of a day. A truck needs to load the customer's cargo when departing from the depot if it passes through the customer's vertex. A truck is allowed to go to any fueling station at any time within its route. When a truck goes through a fueling station vertex, a fixed refueling cost occurs and the truck is fully refueled.

In the MIP model, binary variable $x^k_{i,j}$ is equal to 1 if arc $a^k_{i,j}$ is traveled by truck k in the solution and 0 otherwise. Another binary variable $y^k_{i,f,j}$ is an indicator for the fueling decision. It is 1 if truck k goes to CNG fueling station f when traveling from customer vertex v_i to another customer vertex v_j and 0 otherwise. In many papers, fueling stations are set as nodes in the network and have some replicates if the fueling stations are allowed to be visited multiple times by the same truck. However it is hard to decide how many replicates are enough for the model and too many replicates will dramatically increase the size of the network. Compared with their methods, our model does not have a limitation on the number of times a truck can visit the same fueling station. Continuous variables d^k_i, t^k_i and c^k_i keep track of the traveling distance from the departure depot, traveling distance since last fueling station and current delivered load of truck k at vertex v_i, respectively.

Constant R_1 is the per mile cost for each truck. Constant R_2 is the fixed refueling cost for a truck's single visit to a fueling station. Constant R_3 is the fixed cost if a truck is used for delivery. It can be considered as the price for a truck. M is a big constant number. Constants D and T are the limitations for daily traveling distance and the distance a truck can travel for a CNG fuel tank. Constant L is the limitation of a truck's cargo load. $\delta_{i,j}$ is the distance from vertex v_i to vertex v_j. $\delta^+_{i,f,j}$ is the distance from customer vertex v_i to CNG fueling station f and $\delta^-_{i,f,j}$ is the distance from this same CNG fueling station to customer vertex v_j. Figure 1 shows the relationship between $\delta_{i,j}$, $\delta^+_{i,f,j}$ and $\delta^-_{i,f,j}$. Note that we assume a truck will not visit fueling stations successively. These three kinds of parameters can be calculated before solving the MIP model. σ_i is the demand for customer i, which is known in our model.

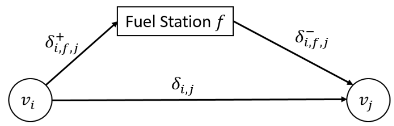


Figure 1: An example of $\delta_{i,j}$, $\delta^+_{i,j}$ and $\delta^-_{i,j}$

3.2.1 Basic Formulation

The basic MIP model for CTRPFS is as follows.

$$\min\ R_1 \sum_{k\in\mathbb{M}} d_n^k + R_2 \sum_{i\in\mathbb{C}\cup\{0\}} \sum_{j\in\mathbb{C}\cup\{n\},i\neq j} \sum_{f\in\mathbb{F}} \sum_{k\in\mathbb{M}} y_{i,f,j}^k + R_3 \sum_{i\in\mathbb{C}} \sum_{k\in\mathbb{M}} x_{0,i}^k \tag{1}$$

$$\text{s.t.} \quad \sum_{i\in\mathbb{C}\cup\{0\},i\neq j} x_{i,j}^k - \sum_{i\in\mathbb{C}\cup\{n\},i\neq j} x_{j,i}^k = 0 \qquad \forall j\in\mathbb{C}, k\in\mathbb{M} \tag{2}$$

$$\sum_{i\in\mathbb{C}\cup\{0\},i\neq j} \sum_{k\in\mathbb{M}} x_{i,j}^k = 1 \qquad \forall j\in\mathbb{C} \tag{3}$$

$$\sum_{i\in\mathbb{C}\cup\{n\}} x_{0,i}^k = 1 \qquad \forall k\in\mathbb{M} \tag{4}$$

$$d_i^k + \delta_{i,j} + \sum_{f\in\mathbb{F}} \left(\delta^+_{i,f,j} + \delta^-_{i,f,j} - \delta_{i,j}\right) y_{i,f,j}^k + M\left(x_{i,j}^k - 1\right) \leq d_j^k$$

$$\forall i\in\mathbb{C}\cup\{0\}, j\in\mathbb{C}\cup\{n\}, i\neq j, k\in\mathbb{M} \tag{5}$$

$$0 \leq d_i^k \leq D - \delta_{i,n} \qquad \forall i\in\mathbb{N}\backslash\mathbb{F}, k\in\mathbb{M} \tag{6}$$

$$t_i^k + \delta_{i,j} - M\sum_{f\in\mathbb{F}} y_{i,f,j}^k - M\left(1 - x_{i,j}^k\right) \leq t_j^k$$

$$\forall i\in\mathbb{C}\cup\{0\}, j\in\mathbb{C}\cup\{n\}, i\neq j, k\in\mathbb{M} \tag{7}$$

$$t_i^k + \delta^+_{i,f,j} y_{i,f,j}^k \leq T$$

$$\forall i\in\mathbb{C}\cup\{0\}, j\in\mathbb{C}\cup\{n\}, i\neq j, f\in\mathbb{F}, k\in\mathbb{M} \tag{8}$$

$$\delta^-_{i,f,j} y_{i,f,j}^k \leq t_j^k \qquad \forall i\in\mathbb{C}\cup\{0\}, j\in\mathbb{C}\cup\{n\}, i\neq j, f\in\mathbb{F}, k\in\mathbb{M} \tag{9}$$

$$0 \leq t_i^k \leq T \qquad \forall i\in\mathbb{N}\backslash\mathbb{F}, k\in\mathbb{M} \tag{10}$$

$$c_i^k + \sigma_i - M\left(1 - x_{i,j}^k\right) \leq c_j^k$$

$$\forall i\in\mathbb{C}\cup\{0\}, j\in\mathbb{C}\cup\{n\}, i\neq j, k\in\mathbb{M} \tag{11}$$

$$0 \leq c_i^k \leq L \qquad \forall i\in\mathbb{N}\backslash\mathbb{F}, k\in\mathbb{M} \tag{12}$$

$$x_{i,j}^k, y_{i,f,j}^k \in \{0,1\}$$

$$\forall i\in\mathbb{C}\cup\{0\}, j\in\mathbb{C}\cup\{n\}, i\neq j, f\in\mathbb{F}, k\in\mathbb{M} \tag{13}$$

Objective function (1) has three parts. The first part is the total traveling cost, which is calculated by the unit cost R_1 multiplied by the total traveling distance $\sum_{k \in \mathbb{M}} d_n^k$. Similarly unit cost R_2 multiplied by the total number of times the fueling stations are visited $\sum_{i,j \in \mathbb{N}, i \neq j, k \in \mathbb{M}} y_{i,j}^k$ is the second part, the total fixed refueling cost. The last part is the cost of purchasing trucks. The weight for it is set to be R_3, a very big number. Namely minimizing the fleet size is our primary objective. The initial number of trucks will be sufficient so that the problem will always be feasible. However any redundant trucks will directly go from node 0 to node n without causing any cost increase in the objective function. Given that R_3 is a big number, only the ratio of R_2/R_1 will influence the fleet's routing decisions. Constraints (2) to (4) ensure each customer is visited by a truck and all routes are feasible. Constraints (5) and (6) are distance constraints. The daily total traveling distance for each truck is bounded by D. The fuel tank capacity constraint is covered by constraints (7) to (10). If a truck goes into a fueling station between customer vertex v_i and v_j, its cumulative traveling distance since the last CNG fueling station will decrease to 0, namely it is fully refueled. Constraint (8) ensures that a truck always has enough fuel to go to the nearest fueling station, if it needs to refuel. Constraint (9) resets the cumulative traveling distance level if the truck goes into a fueling station when traveling between vertex v_i and v_j. Constraint (10) ensures the fuel tanks are never empty. Constraints (11) and (12) guarantee the demand for each customer can be fulfilled by the assigned truck.

3.2.2 Preprocessing and Valid Inequalities

Variable elimination. Though the setting of variable $y_{i,f,j}^k$ solves the problem of visiting the same fueling station multiple times, it also brings in a number of binary variables, which makes the problem hard to solve. In fact, most of the CNG fueling stations are dominated by a few candidates for a given i, j pair and those corresponding $y_{i,f,j}^k$ can be deleted. We say a CNG fueling station f_1 dominates another f_2 if equations (14) and (15) are satisfied.

$$\delta_{i,f_1,j}^{+} \leq \delta_{i,f_2,j}^{+} \tag{14}$$

$$\delta_{i,f_1,j}^{-} \leq \delta_{i,f_2,j}^{-} \tag{15}$$

The valid inequalities are mainly from the distance and load capacity constraints.

Big M selection. In theory the big M in constraints (5), (7) and (11) can be any number as long as it is sufficiently large. However, an arbitrary big M may bring in some numerical problems and weaken the bound from the linear relaxation of the MIP. We need to choose M as small as possible to accelerate the computation speed. The minimal M for constraints (5), (7) and (11) are $D - \delta_{j,n}$, T and 1.

Constraints combination. Considering the fact that, given k, i, j, at most one of $y^k_{i,f,j}$ can be 1, constraints (8) and (9) can be combined as

$$t^k_i + \sum_{f \in \mathbb{F}} \delta^+_{i,f,j} y^k_{i,f,j} \leq T \qquad \forall i \in \mathbb{C} \cup \{0\}, j \in \mathbb{C} \cup \{n\}, i \neq j, k \in \mathbb{M} \tag{16}$$

$$\sum_{f \in \mathbb{F}} \delta^-_{i,f,j} y^k_{i,f,j} \leq t^k_j \qquad \forall i \in \mathbb{C} \cup \{0\}, j \in \mathbb{C} \cup \{n\}, i \neq j, k \in \mathbb{M} \tag{17}$$

The combination of the constraints is likely to give a better bound in solving the linear relaxation of the MIP.

Distance inequalities. If the total distance of a subset of arcs $\mathbb{A}'$ is bigger than D, a truck can not go through them within a day. We select those subsets of arcs from one customer vertex to another with cardinality no bigger than 3, which are $|\mathbb{A}'| \leq 3$, and add inequality (18) for all of such $|\mathbb{A}'|$.

$$\sum_{arc(i,j) \in \mathbb{A}'} x^k_{i,j} \leq |\mathbb{A}'| - 1 \qquad \forall k \in \mathbb{M} \tag{18}$$

Load inequalities. If a subset $\mathbb{C}'$ of customers' cargo load is more than a truck's load capacity, those subset of customers can not be served by a truck. We select those subsets with cardinality no bigger than 3, namely $|\mathbb{C}'| \leq 3$, and add inequalities (19) for all of such $\mathbb{C}'$.

$$\sum_{i \in \mathbb{C}'} \sum_{j \in \mathbb{C}', i \neq j} x^k_{i,j} \leq |\mathbb{C}'| - 2 \qquad \forall k \in \mathbb{M} \tag{19}$$

4 Heuristic Solution Method

Adaptive Large Neighborhood Search (ALNS) is an efficient heuristic method for solving the classical vehicle routing problems. It is proposed by Ropke and Pisinger [14] from Large Neighborhood Search (LNS), which is first introduced by Shaw [17] for the routing problem. Hiermann et al. [7] extend it by embedding a local search and labeling procedure to solve a vehicle routing problem with time windows and recharging stations. In this section, we propose a hybrid heuristic algorithm that combines an ALNS with a mixed integer programming (MIP) model, which is different from the previous research, to solve the CNG truck routing problem with fueling stations. Note that in our model, we have a daily traveling distance constraint as well as a fuel tank capacity constraint. Preliminary experiments show that it is hard to get a feasible solution after moving into an infeasible region, even from the initial solution. Hence in our method, the search is conducted within the feasible solutions. This procedure has an advantage that whenever the method is stopped, we will always get a feasible solution for the fleet.

Based on the settings in the problem statement, we introduce some more notation. Let r be the routes for the fleet, and r^k be the route for truck k. $r^k_{(i)}$ is the ith node in route k. Figure 2 shows an example of route k. From the problem statement in the previous section, the first node $r^k_{(0)}$ and the last node $r^k_{(j)}$ are the depot vertices v_0 and v_n. Within the route, some nodes, for instance $r^k_{(i+1)}$, may be fueling stations. Let $Dist\left(r^k\right)$ be the total traveling distance for route k and $Load\left(r^k\right)$ be the total load for route k. Let $Dist\left(r^k_{(i)}\right)$ and $Load\left(r^k_{(i)}\right)$ be the traveling distance and cargo load from the start depot until current node $r^k_{(i)}$. Let $CNG\left(r^k_{(i)}\right)$ be the cumulative traveling distance since the last CNG fueling station. Note $Dist\left(r^k\right)$ and $Load\left(r^k\right)$ should be bounded by D and T. Let $Num\left(r^k\right)$ be the number of times truck k visits the fueling stations. Let c^k be the cost of truck k and C be the fleet's total cost, then equation (20) is the objective function of the hybrid heuristic algorithm. C has two components, the cost from the total traveling distance and fixed refueling cost from the refueling processes. As to the fleet size, though it is not directly represented in the objective function, we try to minimize it in our algorithm. Thus the objective is identical with the MIP model.

$$C = \sum_{k \in \mathbb{M}} c^k = \sum_{k \in \mathbb{M}} \left(R_1 Dist\left(r^k\right) + R_2 Num\left(r^k\right) \right) \tag{20}$$

route.png

Figure 2: An example of route k

4.1 Initial Feasible Solution Construction

The construction of an initial feasible solution is modified from Ropke and Pisinger's parallel insertion heuristic [14]. Figure 3 shows the general steps for the construction process. Let Δd_i^k denote

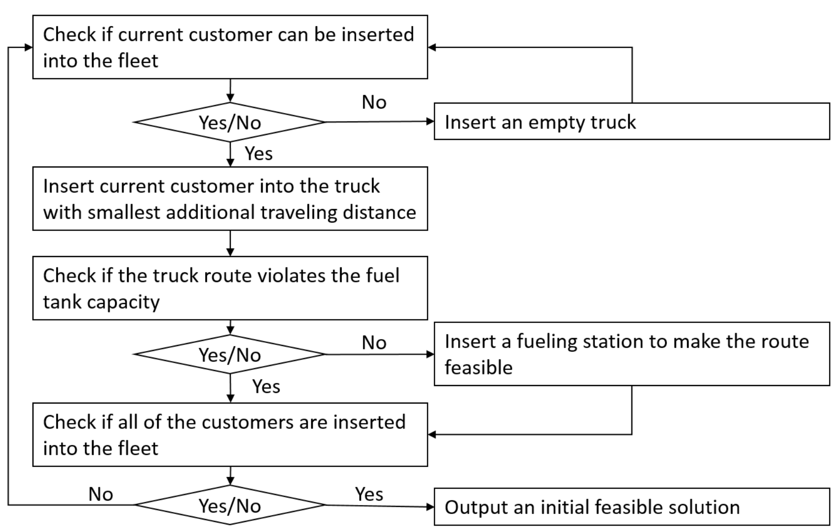


Figure 3: Steps of constructing an initial feasible solution

the change in distance incurred by a valid insertion of customer vertex v_i into route k at the position that increases the distance for truck k the least. Suppose customer vertex v_i is going to be inserted into truck k's route, after the jth node. The change, or increase in distance, is simply measured by $\delta_{r^k_{(j)},v_i} + \delta_{v_i,r^k_{(j+1)}} - \delta_{r^k_{(j)},r^k_{(j+1)}}$. A valid insertion means it meets four constraints, which are distance, fuel tank capacity, load capacity and location constraints.

Distance constraint. Equation (21) shows the criterion for the distance constraint. The daily traveling distance can not exceed D.

$$Dist\left(r^k\right) + \delta_{r^k_{(j)},i} + \delta_{i,r^k_{(j+1)}} - \delta_{r^k_{(j)},r^k_{(j+1)}} \leq D \tag{21}$$

Fuel tank capacity constraint. For the fuel tank capacity constraint, equations (22) to (24) show its two criteria. Equation (22) ensures the truck has enough fuel to go to the fueling station and equation (23) ensures the truck still meets the daily traveling distance constraint if it needs to refuel. In equation (24), the first part $\delta^-_{r^k_{(j)},i} - CNG\left(r^k_{(j+1)}\right)$ is the decrease in the cumulative traveling distance after visiting the fueling station. The second part $CNG\left(r^k_{(j+1)^*}\right)$ is the cumulative traveling distance at node $r^k_{(j+1)^*}$ where node $r^k_{(j+1)^*}$ is the first fueling station after node $r^k_{(j+1)}$, or the last node if there is no fueling station after it. It eliminates a special case in which the distance from the fueling station back to vertex v_i is so large that inserting a fueling station into the route still can not make it feasible.

$$CNG\left(r^k_{(j)}\right) + \delta^+_{r^k_{(j)},i} \leq T \tag{22}$$

$$Dist\left(r^k\right) + \delta^+_{r^k_{(j)},i} + \delta^-_{r^k_{(j)},i} + \delta_{i,r^k_{(j+1)}} - \delta_{r^k_{(j)},r^k_{(j+1)}} \leq D \tag{23}$$

$$\delta^-_{r^k_{(j)},i} - CNG\left(r^k_{(j+1)}\right) + CNG\left(r^k_{(j+1)^*}\right) \leq T \tag{24}$$

Load capacity constraint. Load capacity constraint is explained in Equation (25). The load of truck k can not be more than L.

$$Load\left(r^k\right) + \sigma_i \leq L \tag{25}$$

Location constraint. The last constraint for a valid insertion is the location constraints. Given the assumption that fueling stations can not be visited successively, node $r^k_{(j)}$ can not be a fueling station vertex.

$$r^k_{(j)} \notin \mathbb{F} \tag{26}$$

Initially Δd^k_i is set to be $+\infty$ and Δd^k_i is updated by equation (27).

$$\Delta d^k_i = \min_{j \in \mathbb{J}^k_i} \left(\delta_{r^k_{(j)}, v_i} + \delta_{v_i, r^k_{(j+1)}} - \delta_{r^k_{(j)}, r^k_{(j+1)}} \right) \tag{27}$$

where $\mathbb{J}^k_i$ is the set of valid insertion positions in truck k for customer vertex v_i. Finally we choose truck k that minimizes Δd^k_i and insert customer vertex v_i at the minimum cost position in that truck. If $\Delta d^k_i = +\infty$ for all k, there is no valid insertion for customer vertex v_i and we need to use another truck route.

After inserting customer vertex v_i into truck k, we need to process an inspection on the truck's fuel tank capacity constraint. If there exists a node $r^k_{(j)}$ whose $CNG\left(r^k_{(j)}\right)$ is bigger than T, an insertion of the fueling station is required. Though equations (22) to (24) ensure that inserting a fueling station with minimum detour in front of v_i will make the route feasible again, we still want to find a feasible position j so that $CNG\left(r^k_{(j)}\right)$ is less than T while as big as possible. In reality, this strategy means we try to delay the fueling as much as possible. The truck will keep working until its fuel tank is almost empty. Intuitively, this delay will minimize the number of visits to the fueling stations. The search of position j starting from the first node $r^k_{(m)}$ in route r^k whose $CNG\left(r^k_{(m)}\right)$ is bigger than T. We set $j = m$ and if two adjacent nodes $r^k_{(j-1)}$ and $r^k_{(j)}$ satisfy equations (22) to (24), we find feasible position j. Otherwise we set $j = j - 1$ and check again, until we find a feasible position. In the worst case, we will stop at the position where we just inserted the customer vertex v_i. The insertion will stop until all customers are inserted into the fleet.

4.2 Adaptive Large Neighborhood Search

The ALNS is adapted from Hiermann et al. [7]. However, our search is conducted only within the feasible region, which is different from the paper by Hierman et al. [7]. Figure 4 shows the steps of one iteration of the ALNS. We first do local search to intensify the search in each iteration. The next step is the destroy and repair procedures from the general methodology of the ALNS. As our objective is to make the fleet size as small as possible, we try to reduce the fleet size explicitly in the following step. The last one is using a MIP model to find the optimal fueling plan for each route.

alns.png

Figure 4: Steps of our Adaptive Large Neighborhood Search (ALNS)

4.2.1 Local Search

We introduce some well-known local search methods like shift, swap, 2-opt and 2-opt*, to improve the solution quality in each search iteration. These methods are widely used in various kinds of heuristic methods for vehicle routing problems [13] [2] [7]. In our paper, graphs from Hiermann et al. [7] are used for illustration purposes.

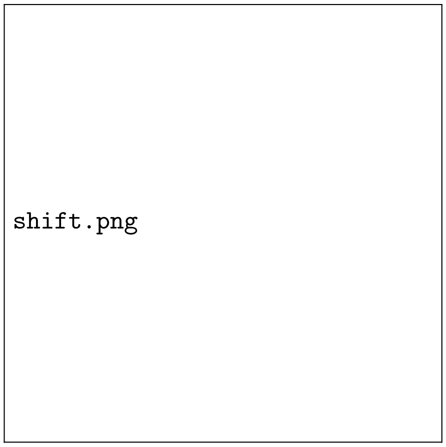
Shift. Figure 5a shows the shift movement. We randomly select a node, for instance the ith node from route k, and insert it into another randomly selected route k'.

Switch. In the switch movement, two nodes i and j are randomly picked from two routes k and k', and switch positions, as shown in Figure 5b.

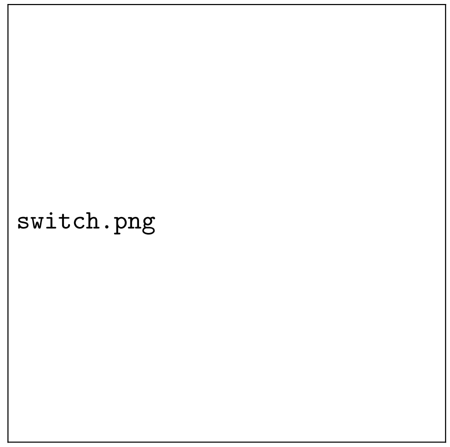
2-Opt. In this movement, we first randomly select two positions i and j, reverse the order of nodes between i and j, and insert into the original position. Figure 5c shows this movement.

2-Opt*. 2-Opt* movement, also called 2-Exchange, is a movement within two routes. Figure 5d is an example of a 2-Opt* movement. Two routes k and k' exchange the rest of the nodes at positions i and j.

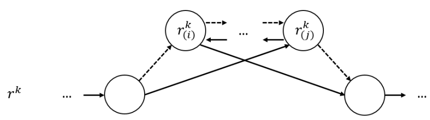
To select the local search methods to use, we adapt the weighted roulette wheel selection method and adaptive weight adjustment from Ropke and Pisinger [14]. Similar to their work, we define a segment as a number of iterations of the local search; here we define a segment as 10,00 iterations. Let $w^s_{i,t}$ represent the weight for local search method i at segment t. We select method i with



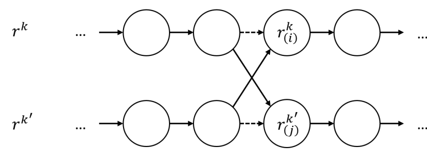
(a) Shift



(b) Switch



(c) 2-Opt



(d) 2-Opt*

Figure 5: Local search

probability given by equation (28) at segment t. The weights are updated via equation (29) and initial weights $w^s_{i,0}, i = 1, 2, 3, 4$ for the four methods are equal.

$$\frac{w^s_{i,t}}{\sum_{j=1}^{4} w^s_{j,t}} \qquad i = 1, 2, 3, 4 \tag{28}$$

$$w^s_{i,t+1} = w^s_{i,t}(1 - r) + r^s \frac{\pi_{i,t}}{\theta_{i,t}} \qquad i = 1, 2, 3, 4 \tag{29}$$

In equation (29), r^s is a constant number, $\pi_{i,t}$ is the score of method i obtained during segment t; here we set it to be the decrease in the objective function, and $\theta_{i,t}$ is the number of times we have attempted to use method i during segment t. In our experiments, we select $w^s_{i,0} = 100$ and $r^s = 0.5$.

Note that the search is conducted within the feasible region. The first acceptance criterion is that the new routes after the local search should still be feasible. The feasibility check procedures are similar to equations (21) to (24) in Section 4.1. The second criterion is from the Simulated Annealing method, which is also used by Ropke and Pisinger [14]. To avoid getting trapped in a local minimum, a worst solution may also be accepted with some probability. A solution will be accepted with probability

$$e^{-(C'-C)/T^s} \tag{30}$$

where C' is the cost of the new routes and C is the original one. T^s is the temperature and starts with an initial value T_0^s. In each iteration T^s is updated via $T^s = \alpha^s * T^s$, where $0 < \alpha^s < 1$. We select $T^s = 10$ and $\alpha^s = 0.9999$ for our experiments.

Since local search can be conducted very fast, we run Φ_1 iterations in every ALNS iteration. If after a local search iteration a route is empty, or has fueling stations only, it will be deleted. Based on our test, we set $\Phi = 100,000$ in the experiments.

4.2.2 Destroy and Repair Operators

The destroy and repair operators we introduce in this section are mainly from Ropke and Pisinger [14]. We get a new solution by removing some nodes using the destroy operator, followed by the repair operator. These operators will extend the search space by bringing a larger search neighborhood. We implement three destroy operators and two repair operators. Notice only the nodes representing customer vertices will be considered in the destroy and the following repair operators. We want to generate a feasible solution whenever the algorithm stops. A destroy of a CNG fueling station might make it hard to generate a feasible solution during the repair procedure.

Random destroy operator. Random destroy operator is the simplest operator. Every node representing a customer vertex is chosen with equal probability.

Related destroy operator. The idea of the related destroy operator is to remove similar nodes together, and reinsert them into routes again. Intuitively if the removed nodes are far away from each other, the solution after the repair operators may still be similar to the original one, or even worse. The relatedness $Related\,(i,j)$ for two vertices v_i and v_j is from Shaw [17].

$$Related\,(i,j) = \frac{1}{\delta'_{i,j} + V_{i,j}} \tag{31}$$

In equation (31), $\delta'_{i,j}$ is the normalized distance from vertex v_i to vertex v_j and $\delta'_{i,j} = \frac{\delta_{i,j}}{\max \delta_{i,j}}$, where $\delta_{i,j}$ is the distance between vertex v_i and vertex v_j and $\max \delta_{i,j}$ is the maximal one for all of the distances. $V_{i,j}$ is an indicator variable. It is 1 if vertex v_i and vertex v_j are in the same route and 0 otherwise. When using the related destroy operator, the first node is randomly selected with the random destroy operator, followed by 4 nodes, selected by the relatedness of the first node. The higher value the relatedness is, the higher probability the node will be chosen.

Worst destroy operator The idea of worst destroy operator is if the node leads to a high detour in its route, it is likely to be misplaced and should be removed. This operator is used by Ropke and Pisinger [14]. First we calculate the detour $Detour(i)$ of the ith node in route k via equation (32).

$$Detour(i) = \delta_{r^k_{(i-1)}, r^k_{(i)}} + \delta_{r^k_{(i)}, r^k_{(i+1)}} - \delta_{r^k_{(i-1)}, r^k_{(i+1)}} \tag{32}$$

Then we choose the node to be removed based on $Detour(i)$. The higher the detour is, the higher probability the node will be chosen.

Cheapest route repair operator. This repair operator is exactly the same as the insertion method in Section 4.1.

Cheapest customer repair operator. In this operator, we first calculate the cheapest insertion cost for every customer removed by the destroy operators. The procedure is the same as the cheapest route repair operator and is given in Section 4.1. Next we choose the customer with the lowest overall cost and insert it into the corresponding route using the same weighted roulette selection method as in Section 4.1. The parameters in equations (28) and (29) for the destroy operator are $w^d_{i,t}, i = 1, 2, 3$ and α^d. The parameters for the repair operators are $w^r_{i,t}, i = 1, 2$ and α^r. The segment is redefined as 100 iterations for both the destroy and repair operators. After tuning the parameters, the initial values for $w^d_{i,0}, i = 1, 2, 3$ and $w^r_{i,0}, i = 1, 2$ are all 100, and $\alpha^d = \alpha^r = 0.8$.

Every time before using the destroy and repair operators, we first decide how many nodes are going to be removed. In our case, we randomly generate a number within $\gamma_1|\mathbb{C}|$ and $\gamma_2|\mathbb{C}|$, where γ_1 and γ_2 are constants and $0 < \gamma_1 < \gamma_2 < 1$ and $|\mathbb{C}|$ is the total number of customers. Next we choose one destroy and one repair operator for this iteration. The selection method is similar to the weighted roulette wheel selection method and adaptive weight adjustment described in Section 4.2.1. If a route after the destroy operator is empty, or has fueling stations only, it will be deleted. In our experiments, $\gamma_1 = 0.05$ and $\gamma_2 = 0.15$.

4.2.3 Fleet Size Minimization

A lower bound for the fleet size can be calculated from the customers' cargo load. Let S^* be the lower bound.

$$S^* = \lceil \frac{\sum_{i \in \mathbb{C}} \sigma_i}{L} \rceil \tag{33}$$

If the current fleet size S is bigger than S^*, we select $S - S^*$ routes with the smallest number of customers, delete these routes and put the customers into the request pool again. We insert those customers into the fleet via repeating the construction procedure in Section 4.1. The new fleet will be accepted if its size S' is smaller than S, otherwise we keep the original one. Though the new fleet size is usually smaller, the total cost C is likely to be larger because of the relatively narrow view on the insertion procedure.

Preliminary experiments show the fleet size minimization procedure will dramatically slow down the computing speed and increase the total cost. Thus, we set the frequency of this procedure to be very low. We run it every Φ_2 ALNS iterations. We set Φ_2 to be 200 in our experiments so that the fleet size are usually minimized, while the computing speed is not affected too much.

4.2.4 CNG Fueling Station Optimization

Hiermann et al. [7] also add the recharging stations into their model and implement a special labeling algorithm to explicitly handle these stations. The idea of the labeling algorithm is to use a specific design label to denote all of the possible recharging plans and pick the one with the lowest total cost. The success of the labeling algorithm heavily depends on the elimination of the search space, namely the so-called dominance checks in their paper. Recall in Hiermann et al. [7], they have customer time windows so the search space is significantly reduced, which is not the case for our problem. In our model the dominance checks are weak and the labels will extend exponentially even if we only consider one possible CNG fueling station between every successive customer pair. Thus we develop a MIP model to deal with the CNG fueling stations.

Different from the MIP model in Section 3, the choice of CNG fueling stations for each route is independent now and each route can be solved separately. We take route r^k for instance. We delete all of the existing fueling stations in r^k and leave the nodes with the customer and depot vertices only. Let $r^{k'}$ be the route after deletion and suppose there are n_d nodes left. We introduce a binary decision variable $z^k_{(i),f}$ to denote the fueling decision with respect to fueling station f for route k after the ith node, which is customer node $r^k_{(i)}$. Let $z^k_{(i),f}$ be 1 if the corresponding truck goes to CNG fueling station f after $r^k_{(i)}$ and be 0 otherwise. We follow the notation in Section 3. For simplification, let $z_{i,f}$ represent $z^k_{(i),f}$, i represents customer node $r^k_{(i)}$ and t_i represent t^k_i. The MIP is as follows.

$$\min R_1 \sum_{i=1}^{n_d-1} \sum_{f\in\mathbb{F}} \left(\delta^+_{i,f,i+1} + \delta^-_{i,f,i+1} - \delta_{i,i+1}\right) z_{i,f} + R_2 \sum_{i=1}^{n_d-1} \sum_{f\in\mathbb{F}} z_{i,f}$$

$$\sum_{i=1}^{n_d-1} \delta_{i,i+1} + \sum_{i=1}^{n_d-1} \sum_{f\in\mathbb{F}} \left(\delta^+_{i,f,i+1} + \delta^-_{i,f,i+1} - \delta_{i,i+1}\right) z_{i,f} \le D \tag{34}$$

$$t_i + \delta_{i,i+1} - M \sum_{f\in\mathbb{F}} z_{i,f} \le t_{i+1} \qquad \forall i = 1, \ldots, n_d - 1 \tag{35}$$

$$t_i + \delta^+_{i,f,i+1} \sum_{f\in\mathbb{F}} z_{i,f} \le T \qquad \forall i = 1, \ldots, n_d - 1 \tag{36}$$

$$\delta^-_{i,f,i+1} z_{i,f} \le t_{i+1} \qquad \forall f \in \mathbb{F}, i = 1, \ldots, n_d - 1 \tag{37}$$

$$0 \le t_i \qquad \forall i = 1, \ldots, n_d \tag{38}$$

$$z_{i,f} \in \{0, 1\} \qquad \forall f \in \mathbb{F}, i = 1, \ldots n_d \tag{39}$$

The objective function is to minimize the total traveling cost from the detour to CNG fueling stations, and the total fixed refueling cost. Constraint (34) is the limitation on daily traveling distance. Constraints (35) to (38) are the same as constraints (7) to (10), which ensure the route meets the fuel tank capacity constraint.

In practice, we use the variable elimination method in Section 3.2.2. However, the speed is slow considering that the above MIP problem needs to be frequently solved. Hence, for any two successive nodes in $r^{k'}$, we just select the fueling station with smallest detour as the only candidate. Similar to the fleet size minimization step, we run it every Φ_3 ALNS iterations. In our experiments, we set Φ_3 to be 100.

4.2.5 Acceptance and Stopping Criteria

Similar to Ropke and Pisinger's work [14], in every ALNS iteration we accept the new solution if it is better than the previous one, or with a probability calculated by equation (30) if it is worse. However, the parameters T^a_0 and α^a are different from T^s_0 and α^s. In fact, these two probabilities are updated independently.

The ALNS stops if the improvement of every ϕ iterations is less than a γ percentage, or the total number of iterations reaches Φ. We set $T_0^a = 1000$, $\alpha^a = 0.999$, $\phi = 100$, $\gamma = 0.1$ and $\Phi = 5,000$.

5 Computational Results

In this section we present our computational results. We first run the proposed ALNS on small instances. These small instances are also solved by the MIP formulation using a standard commercial optimization solver CPLEX. The small instances are used to show the efficiency of the preprocessing and valid inequalities presented in Section 3.2.2 in reducing the computation time in finding an optimal solution and serve as benchmark cases to show that our ALNS can get high quality solutions in a short amount of computation time. Also we use our ALNS to solve large instances to explore some interesting insights. In the last part, we conduct experiments on the estimation of daily goods delivery data from the Ports of Los Angeles and Long Beach to show the losses of route efficiency for CNG trucks can be reduced with an effective refueling strategy.

All experiments are performed on a PC with Intel Core i7-4790K CPU and 32 GB RAM, running Linux distribution Ubuntu 14.04 and CPLEX 12.5.

5.1 Results on Small Instances

We first create small instances by randomly generating 8 to 12 customers on a 100*100 map. A depot and 3 or 4 CNG fueling stations are also randomly picked within the map. The demands of customers are generated from a uniform distribution between 0 to 0.6 unit. We solve these small instances with CPLEX 12.5, as well as our proposed ALNS. In order to show the effectiveness of our preprocessing and valid inequalities, three kinds of formulations, basic, improved and full, are tested. The basic formulation only contains objective function (1) and constraints (2) to (13). The improved formulation includes the basic formulation and preprocessed by variable elimination and big M selection. The full formulation includes everything presented in Section 3.2. We set the parameters R_1 and R_2 in both the MIP objective function (1) and the ALNS objective function (20) as 1 and 10 respectively. Namely the cost for traveling distance is \$1 per mile and fixed fueling cost is \$10 per time. Parameter R_3, the fixed cost for operating a truck, in the MIP objective function (1) is set to 200, which is much higher than R_1 and R_2 so that the MIP solution will use as few trucks as possible. Daily traveling distance D and tank capacity T are set to 400 and 200, respectively. Load capacity L for each truck is 1.

Name	CPU Time (s)				Solution Value				
	Basic	Improved	Full	Dif(%)	Opt	Obj	# Opt	$\overline{\text{Obj}}$	Gap(%)
8C3F3T1	724.2	369.8	282.7	23.6%	322.1	322.1	10	322.1	0.0%
8C3F3T2	785.3	358.9	322.3	10.2%	367.4	367.4	10	367.4	0.0%
8C3F3T3	425.2	244.3	196.9	19.4%	425.9	425.9	6	445.5	4.5%
8C3F3T4	676.0	343.2	255.1	25.7%	458.3	458.3	8	467.4	2.0%
8C3F3T5	626.5	283.5	53.8	81.0%	376.1	376.1	10	376.1	0.0%
avg	**647.5**	**319.9**	**222.1**	**30.6%**	**390.0**	**390.0**	**8.8**	**395.7**	**1.5%**
9C3F4T1	8,375.0	3,839.2	2,949.1	23.2%	472.2	472.2	8	473.7	0.3%
9C3F4T2	6,724.6	3,710.4	2,161.7	41.7%	732.0	732.0	10	732.0	0.0%
9C3F4T3	4,647.0	2,615.7	2,273.4	13.1%	523.2	523.2	3	529.5	1.2%
9C3F4T4	2,838.2	1,427.8	859.9	39.8%	376.6	376.6	5	380.5	1.0%
9C3F4T5	1,863.3	1,280.0	831.2	35.1%	529.0	529.0	7	531.4	0.4%
avg	**4,889.6**	**2,574.6**	**1,815.1**	**29.5%**	**526.6**	**526.6**	**6.6**	**529.4**	**0.6%**
10C4F4T1	-	48,894.0	44,799.5	8.4%	588.5	588.5	6	591.7	0.5%
10C4F4T2	-	70,440.6	50,068.0	28.9%	587.3	587.3	7	593.7	1.1%
10C4F4T3	-	60,495.2	43,046.0	28.8%	584.4	584.4	10	584.4	0.0%
10C4F4T4	-	42,867.3	34,269.6	20.1%	462.7	462.7	1	463.7	0.1%
10C4F4T5	42,827.9	26,873.9	16,237.8	39.6%	678.6	678.6	8	681.4	0.2%
avg	**-**	**49,914.2**	**37,684.2**	**24.5%**	**571.3**	**571.3**	**6.4**	**583.0**	**0.2%**

Table 1: Results on small instances

Table 1 presents the results on small instances. In total we conduct experiments with 3 different settings and each setting has 5 random instances. Name 8C3F3T1 means it is the first instance with a setting of 8 customers, 3 CNG fueling stations and 3 trucks. The first part CPU Time has four columns. Colum Basic, Improved and Full are the CPU times (in seconds) used by the basic, improved and full formulations. Column Dif is the improvement on CPU time when comparing columns Improved and Full. The "-" notation in column Basic means CPLEX can not solve the problem optimally within 20 hours of CPU time. These three columns show that our preprocessing and valid inequalities can save CPU time, even for constraints combination and valid inequalities only. The second part Solution Value contains the solution values from the MIP and the ALNS. Since both the MIP and the ALNS solutions use the same number of trucks, we use the sum of the traveling and fueling cost for comparison. Column Opt is the optimal solution value from CPLEX.

For the ALNS, we run 10 times for each instance. Column Obj is the best one and Column # Opt is the number of times the ALNS obtained the optimal solution. Column $\overline{\text{Obj}}$ is the average of the 10 solution values. Column Gap is the gap between the average value from the ALNS, which is column $\overline{\text{Obj}}$, and the optimal value from CPLEX, which is column Opt. The Solution Value part indicates that for each instance, our ALNS can achieve the optimal solution within 10 runs, and the gap between the average value and optimal solution is also very small. In these 15 instances, the average gap is less than 2.0% and at least one run of the 10 can identify the optimal solution for each scenario.

5.2 Results on Large Instances

In order to explore the effectiveness of CNG trucks, we conduct experiments on large instances. The large instances are modified from the test case A-n33-k5 by Augerat et al. [1]. In the original test case, there are 32 customers and 1 depot randomly allocated on a 100*100 map. Each customer is assigned with a random number representing the demand volume. Each truck has identical capacity limitation and delivers cargo to customers with a tour starting and ending at the depot. The optimal solution for this problem is known in the literature and in the optimal solution, 5 trucks are used to deliver products to the customers. We modify the test case by adding tank capacity and daily traveling distance constraints to the trucks so that they are more like CNG trucks. Note that the previous known optimal solution may no longer be feasible with these added constraints. We also randomly select several points within the map to be CNG fueling stations. Note that we assume the trucks' tanks are initially full at the beginning. Another assumption is that there is no CNG fueling station at the depot and all trucks in the fleet need to use the public CNG fueling stations. If we develop routing plans for a single day, all trucks are likely to return to the depot with empty tanks, which will have a large influence on the next day. Thus we treat the customers' demands as daily, rather than a single time, and extend the schedule to be 5 days with the same repeated daily demand. We set the objective function parameters R_1 and R_2 as 1 and 10, which is identical with Section 5.1. For daily traveling distance D, tank capacity T and load capacity L, we set to 400, 250, and 1, respectively.

$\|\mathbb{F}\|$	Instance 1		Instance 2		Instance 3		Instance 4		Instance 5	
	Dist	Fuel	Dist	Fuel	Dist	Fuel	Dist	Fuel	Dist	Fuel
3	3,704.0	26.2	3,949.3	29.1	3,727.5	27.5	3,639.2	27.3	3,992.5	29.7
4	3,677.5	26.2	3,797.8	27.3	3,704.1	26.8	3,604.9	26.4	3,942.3	29.1
5	3,664.4	25.9	3,778.4	26.9	3,668.4	26.5	3,591.6	26.4	3,939.3	29.1
7	3,641.5	25.6	3,755.5	27.0	3,642.3	26.2	3,590.5	26.4	3,852.1	28.1
10	3,603.8	24.9	3,666.1	26.1	3,637.5	26.1	3,554.1	25.2	3,653.3	26.9

Table 2: Results for different number of CNG fueling stations

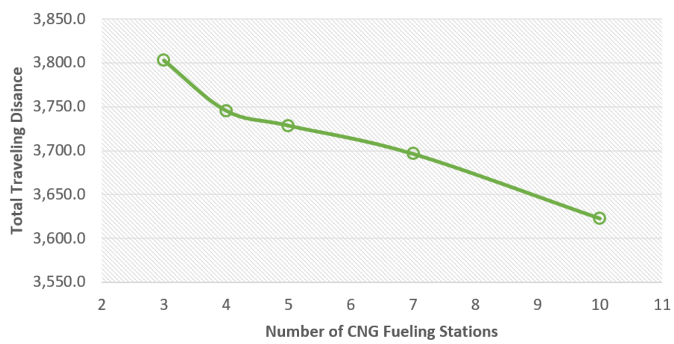


Figure 6: Average total traveling distance with different number of CNG fueling stations

Table 2 presents the detailed results on different number of fueling stations. We run 5 instances and each instance has fueling stations increasing from 3 to 10, by gradually adding extra randomly located CNG fueling stations into the map. The instances along the column in Table 2 differ only in the location of the CNG fueling stations. The first column in Table 2 shows the cardinality of the CNG fueling stations set $\mathbb{F}$, namely the number of CNG fueling stations. For each instance with a certain number of fueling stations, we run the ALNS 10 times and each time the best solution is recorded. Column Dist is the average total traveling distance of these 10 runs. Column Fuel is the average number of times the 5 trucks use a CNG fueling station in the 10 runs. Figure 6 presents the average total traveling distance of these 5 instances. From Table 2 we observe that when increasing the number of CNG fueling stations, the number of times the trucks visit a fueling station decreases. One explanation for this phenomenon is, trucks have more flexibility on the choice of fueling stations when the number of CNG fueling station increases and drivers are more likely to refuel when their tanks are almost empty. Considering the relatively high setup cost for the CNG trucks refueling process, increasing the number of CNG fueling stations will bring in extra benefits. Also the queue in the stations will be shorter, not only because of the allocation of refueling requests, but also the reduction in the total refueling requests. As expected we can observe from Table 2 that for each instance the total traveling distance decreases, if the total number of available CNG fueling stations increases. Figure 6 also presents such a trend. The line is the average total traveling distance of the 5 Dist columns in Table 2. It has a rapid decrease when the number of CNG fueling station increases from 3 to 4, which means the number of CNG fueling station is an important factor when the total number is low, and it also follows the law of diminishing marginal utility. We next compare these solutions with the optimal solution without considering the daily traveling distance and tank capacity constraints. A solution without these constraints can represent a routing plan for a fleet of diesel trucks and is a standard solution to a capacitated VRP problem. The optimal solution for the original instance A-33n-5k can be found in the work from Augerat et al. [1] and the corresponding total traveling distance is 3,305. Compared with the optimal solution, the average total traveling distance with 3 CNG fueling stations is 15.1% more and with 10 CNG fueling stations is only 9.6% more.

T	Instance 1		Instance 2		Instance 3		Instance 4		Instance 5	
	Dist	Fuel	Dist	Fuel	Dist	Fuel	Dist	Fuel	Dist	Fuel
150	3,610.4	24.9	3,671.6	26.0	3,642.8	25.9	3,553.1	25.4	3,657.7	26.8
175	3,545.1	21.6	3,577.6	22.1	3,570.0	22.8	3,484.4	21.0	3,532.4	21.6
200	3,511.4	17.4	3,509.6	18.9	3,495.7	18.0	3,454.2	18.1	3,467.3	18.9
225	3,472.8	14.6	3,487.3	16.2	3,467.0	15.1	3,438.1	15.7	3,443.0	15.8
250	3,453.9	12.6	3,456.5	13.4	3,441.9	13.0	3,425.5	13.1	3,432.0	12.9
275	3,445.5	10.8	3,447.1	11.8	3,425.7	10.9	3,419.0	10.9	3,409.7	10.6

Table 3: Results for different fuel tank capacities

tankcase.png

Figure 7: Average total traveling distance with difference tank capacities

Next we do sensitivity analysis on the tank capacity to explore its influence on the routing plans. We modify the instances by fixing the number of CNG fueling stations at 10 and changing the tank capacity from 150 to 275. All the other parameters are identical with the previous group instances. Again we run 5 instances. For these 5 instances, all of the settings are the same except for the locations of the CNG fueling stations. Within each instance we run the ALNS with different tank capacities and for each tank capacity level, we run 10 times to get the average results. Table 3 and Figure 7 present the influence of the tank capacity. Table 3 is quite similar to Table 2. Figure 7 shows the tank capacity has a large influence on the routing plans when the tank capacity is low. Note that when the tank capacity is set to 275, the average total traveling distance of 3,429.4 is quite close to the total traveling distance in the optimal solution without the daily traveling distance and tank capacity constraints, which is 3,305. Indeed, the average total traveling distance only increased by 3.76%. Note that in the benchmark instance, there is no daily traveling distance limitation or tank capacity constraint. This small gap indicates that refueling is no longer an important factor for CNG trucks under this specific setting, as well as demonstrating the efficiency of our proposed ALNS.

5.3 Results for Data from the Ports of Los Angeles and Long Beach

We run experiments on the estimation of daily goods delivery data for the Ports of Los Angeles and Long Beach. The data is from the report "SCAG Regional Travel Demand Model and 2008 Model Validation", which was conducted by the Southern California Association of Governments (SCAG) in 2012. SCAG commissioned a team to develop a demand model for trucks. Within this model the Southern California region is divided into 4,109 blocks and each block is associated with a pair of longitude and latitude, which represents its location. We calculated the direct distance between every pair of blocks based on their longitudes and latitudes and use it as the distance between them. For every pair of blocks the report estimates the daily demand in terms of truckloads for light, medium and heavy duty diesel trucks based on the information of land use and socioeconomic data. We select the block where the Ports of Los Angeles and Long Beach are located as the depot for our fleet. We use the light duty diesel trucks' truckload trip rates between the depot and the rest of the 4,108 blocks as the daily demands to meet. For the locations of CNG fueling stations, we collect their locations from the Alternative Fuels Data Center under the U.S. Department of Energy and locate them within one of the 4,109 blocks. In total there are 85 public accessible CNG fueling stations in the Southern California region.

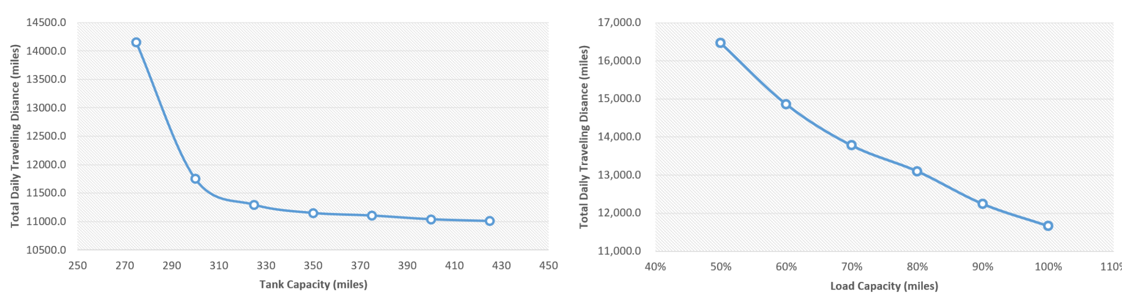
In order to estimate the tank capacity differences between CNG and diesel trucks, we take a model, Chevrolet 2014 Express 3500 as representative for light duty trucks. This model has a CNG engine and a diesel engine version. For Chevrolet 2014 Express 3500, the standard CNG version has a 300 mile range and the CNG load capacity is about 28.4% less than the capacity of the diesel version. Based on the information above, we assume light duty CNG trucks have a tank capacity limitation of 300 miles traveling at normal speed. As to the load capacity, light duty trucks are assumed to be 30% less than that of diesel versions. Namely for one truckload trip of a diesel truck, it requires $\frac{1}{1-0.3} = 1.429$ truckload trips for a CNG truck. Thus compared with diesel trucks, the demand in terms of truckload trips is increased by 42.9% when using CNG trucks. For all trucks, the daily traveling distance limitation is 400 miles. Similar to the setting in Section 5.2, we run 5 days with the same repeated daily demands to minimize the influence of initial full tanks.

	Light duty trucks		
	LTL	FTL	Total
Diesel	10,468.6	209.0	10,677.6
CNG	13,435.1	341.6	13,776.7
Dif	28.2%	63.4%	**29.0%**

Table 4: Total daily traveling distance for light and heavy duty trucks

Table 4 presents the total daily traveling distance for light duty trucks. The values in the table are the average of the best solutions for the 10 independent runs. For any block, if its trip rate is bigger than 1, the integer part will be delivered by full truckload (FTL) shipping and the remaining part may be combined with the trips from other blocks and will be delivered by a less than truck load (LTL) shipping. With 42.9% more demand in terms of truckloads, the total daily traveling distance for CNG trucks will be increased by 42.9% in the worst case. However, for light duty trucks, the total daily traveling distance is increased by 29.0%, which is less than 42.9%.

The extra 29.0% total daily traveling distance for light duty CNG trucks comes from two sources, the tank capacity and the load capacity constraints. In order to explore which factor has a larger influence on the total daily traveling distance, we first run experiments with different tank capacities changing from 275 miles to 425 miles, while the load capacity is identical with the diesel trucks. Then we fix the tank capacity at 300 miles level and run experiments with different load capacities increasing from 50% to 100% of that for diesel trucks.



(a) Tank capacity (b) Load capacity

Figure 8: Total daily traveling distances under different settings for light duty CNG trucks

Figure 8 shows the sensitivity analysis for light duty CNG trucks. For each setting, we run 10 independent replicates and take the average of the 10 best solutions as the value in the figure. In Figure 8a, the load capacity for light duty CNG trucks is the same as diesel trucks and the tank capacity changes from 275 miles to 425 miles. The total daily traveling distance has a rapid reduction when the tank capacity increases from 275 miles to 300 miles after which the reduction is much smaller if the tank capacity keeps increasing. Compared with the total daily traveling distance of 10,677.6 miles for light duty diesel trucks, the increase in the total daily traveling distance for CNG trucks who have the same load capacity as diesel trucks and tank capacity greater than 300 miles is relatively small. In Figure 8b, the tank capacity for CNG trucks is 300 miles and the load capacity increases from 50% to 100% of the diesel trucks. The results indicate load capacity is always an important factor. That is, only when the load capacity of CNG trucks is equivalent to the diesel trucks, the total daily traveling distances between these two types of vehicle are close to being equivalent.

The above computational results illustrate that with the proposed deployment strategies, the influence of the tank capacity from CNG trucks can be reduced to a low level under the current availability of the public CNG fueling stations in the Southern California area. However, the load capacity is a main source for the increase of total daily traveling distance.

6 Implementation

This research proposes an exact algorithm as well as a heuristic solution method to solve the vehicle routing problem for CNG trucks with fueling stations. A typical application of this research is for a small carrier who deploys a fleet of CNG trucks but is not equipped with CNG fueling stations in the depot. By using efficient deployment strategies, the impact of fueling stations can be reduced to a low level even if the sparsely located public CNG fueling stations are the only options. Both the exact algorithm and the heuristic solution method are programmed with C++. In the exact algorithm, we call API from a standard software CPLEX 12.5 to solve the MIP formulation. In the heuristic solution method, the ALNS as well as the embedded local search are implemented with C++, and the MIP formulation for the selection of fueling stations is also solved by calling CPLEX API.

All experiments are conducted on a PC with Intel Core i7-4790K CPU and 32 RAM, running Ubuntu 14.04. Small instances are randomly generated to illustrate the efficiency of our improvement procedures for the MIP model. Large instances are modified from Augerat et al. [1]. We draw some interesting insights from the sensitivity analysis on large instances. In order to show the impact of our algorithm on the cases from the real world, we collect the demand data and the locations of CNG fueling stations in the Southern California area to generate a test case for the ALNS. The demand data is from the report "SCAG Regional Travel Demand Model and 2008 Model Validation". The location information about CNG fueling stations are from the website maintained by the Alternative Fuels Data Center under the U.S. Department of Energy. Results show with an efficient deployment, the impact of small tank capacities can be reduced to a low level.

7 Conclusion and Future Research

In this report, we introduce the CTRPFS to model decisions to be made with regards to the vehicle routes including the choice of fueling stations. We propose a new decision variable setting method for the MIP model. Also we develop some preprocessing steps and valid inequalities to reduce computing time. Experiments with CPLEX show our MIP model works well for this problem and the preprocessing steps and valid inequalities can save 25% to 30% on the CPU time. In order to solve a large instance of realistic size. We propose a hybrid heuristic method combining an ALNS with a local search and MIP model. We run experiments on both small and large instances, and results show that our proposed method can solve the CTRPFS problem in a short amount of computation time with high quality solutions.

We also get some interesting insights from our experiments results. Based on the results from large instances, we find increasing the number of CNG fueling stations can reduce the workload of a single station, not only due to the allocation of workload, but also the reduction in the total number of times trucks use a CNG fueling station. Results show that the tank capacity has an impact on the total traveling distance, especially when the mile range is low. Using data from the Ports of Los Angeles and Long Beach, we determine the extra miles that need to be driven for light and heavy duty CNG trucks over diesel trucks.

Future research may focus on improving the preprocessing and valid inequalities. So far all of the valid inequalities are statically added into the MIP formulation. A heuristic algorithm can be used to select efficient valid inequalities and dynamically add them into the branch-and-bound procedure. Another research direction is to explore the best locations of additional CNG fueling stations in urban areas.

References

[1] Augerat, P., Belenguer, J., Benavent, E., Corberán, A., Naddef, D., and Rinaldi, G. *Computational results with a branch and cut code for the capacitated vehicle routing problem.* IMAG, 1995.

[2] Chen, J.-F., and Wu, T.-H. Vehicle routing problem with simultaneous deliveries and pickups. *Journal of the Operational Research Society* (2006), 579–587.

[3] Dantzig, G. B., and Ramser, J. H. The truck dispatching problem. *Management science 6*, 1 (1959), 80–91.

[4] Demir, E., Bektaş, T., and Laporte, G. A review of recent research on green road freight transportation. *European Journal of Operational Research 237*, 3 (2014), 775–793.

[5] EPA, A. Inventory of us greenhouse gas emissions and sinks: 1990-2013. Tech. rep., EPA 430-R-11-005, 2011.

[6] Erdoğan, S., and Miller-Hooks, E. A green vehicle routing problem. *Transportation Research Part E: Logistics and Transportation Review 48*, 1 (2012), 100–114.

[7] Hiermann, G., Puchinger, J., and Hartl, R. F. The electric fleet size and mix vehicle routing problem with time windows and recharging stations. *Transportation Science* (2014).

[8] Jemai, J., Zekri, M., and Mellouli, K. An nsga-ii algorithm for the green vehicle routing problem. In *Evolutionary computation in combinatorial optimization.* Springer, 2012, pp. 37–48.

[9] Kelley, S., and Kuby, M. Compressed natural gas (cng) fleets in southern california: Variations in vehicles and route types 2. In *Transportation Research Board 94th Annual Meeting* (2015), no. 15-3584.

[10] Khan, M. I., Yasmin, T., and Shakoor, A. Technical overview of compressed natural gas (cng) as a transportation fuel. *Renewable and Sustainable Energy Reviews 51* (2015), 785–797.

[11] Laporte, G. The vehicle routing problem: An overview of exact and approximate algorithms. *European Journal of Operational Research 59*, 3 (1992), 345–358.

[12] Lin, C., Choy, K. L., Ho, G. T., Chung, S., and Lam, H. Survey of green vehicle routing problem: Past and future trends. *Expert Systems with Applications 41*, 4 (2014), 1118–1138.

[13] Montané, F. A. T., and Galvao, R. D. A tabu search algorithm for the vehicle routing problem with simultaneous pick-up and delivery service. *Computers & Operations Research 33*, 3 (2006), 595–619.

[14] Ropke, S., and Pisinger, D. An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. *Transportation science 40*, 4 (2006), 455–472.

[15] Salimifard, K., and Raeesi, R. A green routing problem: optimising co2 emissions and costs from a bi-fuel vehicle fleet. *International Journal of Advanced Operations Management 6*, 1 (2014), 27–57.

[16] Schneider, M., Stenger, A., and Goeke, D. The electric vehicle-routing problem with time windows and recharging stations. *Transportation Science 48*, 4 (2014), 500–520.

[17] Shaw, P. Using constraint programming and local search methods to solve vehicle routing problems. In *Principles and Practice of Constraint Programming CP98*. Springer, 1998, pp. 417–431.

[18] Toth, P., and Vigo, D. *The vehicle routing problem.* Society for Industrial and Applied Mathematics, 2001.

[19] Xiao, Y., Zhao, Q., Kaku, I., and Xu, Y. Development of a fuel consumption optimization model for the capacitated vehicle routing problem. *Computers & Operations Research 39*, 7 (2012), 1419–1431.