

Working Paper Series

A Review of Models for the Alternative-Fuel Station Location Problem

Trung Hieu Tran

Gabor Nagy

Niaz A. Wassan

Kent Business School

A Review of Models for the Alternative-Fuel Station Location Problem

Trung Hieu Tran, Gábor Nagy*, Niaz A. Wassan

Kent Business School, University of Kent, Canterbury, Kent CT2 7PE, United Kingdom

Abstract

The problem of alternative-fuel location is both highly applicable for a sustainable transport infrastructure and challenging to solve. The challenge arises from this problem not fitting into the traditional flow-capturing models, this is because of the scarcity of current infrastructure and the limited driving range of alternative-fuel vehicles. In this paper, we review the models put forward in the literature, with a special emphasis on the mathematical formulations.

Keywords: alternative-fuel station location; sustainability;

1 Introduction

The problem of alternative-fuel station location is a recent, but very applicable research topic within location science. In essence, what make the problem of determining locations of alternative-fuel refuelling stations different from those of petrol stations is the scarcity of current infrastructure. The alternative-fuel industry is suffering from a “vicious circle”: there is little appetite for infrastructure investment as there are not a sufficient number of alternative-fuel vehicles, the automotive industry can only produce these vehicles at high process as there are not sufficient economies of scales due to limited demand, and customers are discouraged from buying such vehicles due to both their price and the limited refuelling infrastructure. For more information on the current state of the alternative-fuel infrastructure, see European Commission (2013). This topic is especially timely in the light of the recent European directive requiring Member States to provide a minimum coverage of refuelling points for alternative fuels (European Commission, 2014). The directive provides a regulatory framework for alternative fuels such as hydrogen, electricity, liquefied natural gas and compressed natural gas. The targets are very ambitious. Compressed natural gas stations and hydrogen stations are to be built along the European TEN-T core network at intervals of 400 and 300 kilometres, respectively. The electricity refuelling network is to be multiplied significantly, from about 12,000 to 800,000. For further information of the European plans, see European Commission (2013). Thus, this is the right time for Operational Researchers to devote their energies to finding optimal or near-optimal locations for alternative-fuel refuelling facilities.

From an Operational Research point of view, the problem is clearly one of location. However, traditional covering models, such as maximum-cover, are not applicable, as these related to demand arising at nodes. In

*Correspondence author: G.Nagy@kent.ac.uk (G. Nagy).

the context of refuelling, the demand arises from vehicle journeys, thus origin-destination flows rather than single points. Since the seminal paper of Hodgson (1981), a steady stream of research has been devoted to such “flow-capturing” problems.

2 Literature Review

2.1 The flow-capturing location model

The flow-refuelling location model (FRLM), introduced by Kuby and Lim (2005), has its origins in the flow-capturing location model (FCLM) of Hodgson (1990). Hodgson originally investigated the concept of “locating facilities on the home-to-work journey”, see Hodgson (1981). He observed that, unlike traditional location models, where facilities are to be sited near customer locations, in some cases it makes more sense to locate facilities near routes that customers already take. The example mentioned is that of locating childcare centres. Hodgson (1990) formalised this, creating the so-called flow-capturing location model. A main theme of this paper is that of cannibalisation, that is, the multiple (and thus unnecessary) capturing of flows. A cannibalising and a non-cannibalising heuristic are compared, showing the superiority of the latter. The author shows that simply observing traffic flows is not sufficient as locating facilities on this basis may lead to cannibalisation. Instead, models should be based on detailed origin-destination flow data. This seminal work spurred a number of flow-capturing papers, some of which we review here. It may be interesting to point out that models that combine both flow and node capturing exist, see for example Hodgson and Rosing (1992) and Berman (1997), who combine flow-capturing and p-median/max-cover objectives.

An important aspect of the FCLM is that any flow (origin-destination pair) is captured by a single facility. This is sensible as one would not, for example, stop at every roadside supermarket on the way home, one stop is sufficient to satisfy one’s shopping needs. However, some variations exist. For example, Hodgson and Berman (1997) considered facilities (roadside advertisement billboards) where one billboard can be considered to “capture” a passing motorist, but seeing the same advertisement again will “reinforce the message” and is thus beneficial. Another interesting example is given in Hodgson et al. (1996), where a facility cannot capture the entire flow (origin-destination trip). The problem at hand was locating inspection stations where drunk drivers or drivers of unlicensed hazardous waste material may be intercepted. The drivers are prevented from causing damage “downstream” from the inspection station but the station cannot prevent accidents or spillages on the previous part of the journey.

It is often assumed that in order to capture a flow, a facility must lie on the origin-destination path. However, it may also be reasonable to assume, especially if the network of facilities is very sparse, that drivers would make some reasonable detours to visit a facility. Berman et al. (1995) allow a flow to be captured if it passes a facility within a specified distance. Variations of the model are also considered, most importantly an assumption whereby a portion of the flow is captured proportional to the distance between facility and path. Berman (1997) considers a median-objective problem with deviation distances.

2.2 The flow-refuelling location model

The main difference of the FRLM from the FCLM is that a single facility may be unable to capture an entire flow. This is due to the issue of “limited range”, namely, that a vehicle may not be able to undertake

a given origin-destination journey with a single refuelling stop. This model is most applicable to vehicles powered by alternative fuels, such as hydrogen or electricity. On one hand, such vehicles normally can cover a shorter distance on a full tank than traditional petrol-guzzling vehicles. On the other hand, the availability of alternative fuel refuelling stations is very limited.

Kuby and Lim (2005) introduce the FRLM, motivating the new model with the above concept of vehicle range. They observe that origin-destination data, rather than simple traffic count on edges, is required to model this problem properly. Multiple facilities may be required to serve individual journeys. Unlike in the FCLM, it can be shown that it is not sufficient to consider only node locations for facilities, thus making the problem harder to solve. An integer programming formulation is provided; this will be looked at in Section 3. The authors drew the following conclusions from their experimentation:

- The longer the vehicles' range the fewer facilities are needed to capture all the demand, but placing facilities only at nodes (junctions) may be unable to provide total coverage.
- There is a lack of convexity in the trade-off curve between the number of refuelling stations opened and the volume of flow they capture, unlike in the FCLM that exhibits convexity.
- Greedy solution approaches tend to give very poor results, much poorer for the same instances as they give for the FCLM.
- Unlike in maximum cover location problems, alternate optima do not often occur in the FRLM. (The authors hypothesise that in this respect the FCLM may fall between the FRLM and the max-cover problem.)

Upchurch et al. (2009) extend the above model to the case of capacitated facilities. An integer programming formulation is presented. (We do not consider the case of capacity limitations and hence this model will not be discussed in detail.) A 25-city case study ("Arizona") is used to illustrate the problem. Another important case study ("Florida", 74 cities, 302 candidate locations), returning to the assumption of uncapacitated facilities, is presented by Kuby et al. (2009). The decision support system is based on an add/swap heuristic. Lim and Kuby (2010) design some heuristic algorithms for the FRLM. One of their motivations for doing so is the complexity of the Kuby and Lim (2005) mathematical formulation. There are three heuristics but with a common subroutine to evaluate the objective function value:

- The "greedy-adding" or "add" algorithm simply adds one more facility in each iteration so as to maximise the increase in flow capture.
- The "greedy-adding with substitution" or "add-swap" algorithm also attempts in each iteration to replace an existing facility with a potential facility. Thus, each iteration consists of an "add" and a "swap" move.
- The genetic algorithm is based on the chromosome representation of a list of open facilities. (As the number of facilities is fixed in advance, this is more reasonable than a 0-1 representation.)

Unlike, say, the maximum covering problem, the evaluation of a given solution is not a straightforward task. For a given solution, i.e. a set of facilities, the evaluation subroutine must evaluate every origin-destination path to see whether it is refuelable – if so, its flow is added to the objective function value. We note that all the algorithms are capable of handling pre-existing facilities. The algorithms are used to evaluate a smaller

instance and also the large “Florida” network. The authors found that the greedy algorithms perform quite well, nearly as well as the genetic algorithms (except for the case of a very short vehicle range), and are significantly faster.

Lately, research has focused on obtaining more efficient formulations to the FRLM. The motivation for this is that the original Kuby and Lim (2005) model requires a massive preprocessing effort. All facility combinations must be checked whether they can refuel each origin-destination journey and the resulting coefficients inserted as input into the integer programming model. This takes an immense amount of time, so much so, that the authors could not even generate the integer programming model for the “Florida” instance, let alone solve it. Capar and Kuby (2012) put forward a more complex model, but without the above preprocessing requirement. This new formulation is in fact as fast as the greedy heuristics of Lim and Kuby (2010). In the model of MirHassani and Ebrazi (2013), the number of facilities is not fixed in advance, as it explicitly takes into account their establishment costs. The logic of their formulation is developed from a single-path to a multi-path formulation. Capar et al. (2013) offer a more efficient formulation than Capar and Kuby (2012). While the previous model used a “node-cover/path-cover” logic, the authors propose an “arc-cover/path-cover” model. The differences between these models will be analysed in more detail in Section 3.

Although the focus of our paper is the FRLM, and indeed many of the works on refuelling alternative-fuel vehicles follow this model, there are some notable works that adopt a different approach. There are three such works that we wish to highlight here; all three are based on the notion of set covering. Wang and Lin (2009) devise a “vehicle refuelling logic” that is more involved but also more flexible than that of the FRLM. Another important difference is that this model required only knowledge of origin-destination distances, but not of the origin-destination flow data. An integer programming formulation based on set covering is presented. A case study with 51 candidate locations is solved to optimality. Kang and Recker (2014) devise a location-routing type model for location alternative-fuel stations. The location part is based on set covering and the routing part is formed by the “household activity pattern problem.” This latter problem captures people’s day-to-day activities and is a more involved model than just considering home-to-work journeys as in the FRLM. As it is more suited to shorter journeys, the model does not need to consider the issue of multiple refuelling stops. However, time windows considerations are introduced. The authors present an integer programming formulation, based on set covering. The differences from standard location-routing problems are highlighted. Numerical experimentation shows that including tour-planning aspects in choosing locations gives a better model with more robust locational decisions; in particular, it shows that a sparser infrastructure network is sufficient than a nodal-demand based model would suggest. Wen et al. (2014) also consider set covering but investigate both maximal flow capture and total flow capture models. Their models do not require the evaluation of all feasible combinations of locations. They are tested on real-life Danish traffic flow data with encouraging results.

2.3 FRLM with deviation

The FRLM assumes that refuelling stations are located on origin-destination paths. However, just like in the FCLM (see Berman et al., 1995, and Berman, 1997), one could consider the situation where driver make some detour from their shortest path in order to refuel. This is a very reasonable assumption given that the network of alternative-fuel stations is a very sparse one at the moment.

Kim and Kuby (2012) introduced the deviation-flow refuelling location model. It is assumed that drivers are willing to make some detour from their shortest paths to visit a refuelling facility, that they would visit the facility/facilities that represent the smallest possible detour, and that the volume of flow captured by a facility decreases with the deviation distance required to reach it. Similarly to the FRLM the preprocessing is not a straightforward task. Evaluating all possible detour journeys is even more cumbersome than evaluating simple origin-destination paths. (This is further complicated by the fact that shortest paths do not contain loops but shortest deviation paths may contain loops.) Then, as facilities on detours do not capture 100% of the flow, another algorithm is required to calculate the volume of flow captured. Finally, as in the FRLM, these paths must be evaluated to see whether they can refuel an origin-destination journey. (This is further complicated by the possibility that a round-trip journey may be optimally refuelled by different facilities on the outbound and the inbound legs.) The authors present an extension of the Kuby and Lim (2005) integer programming model to account for deviations. Their discussion focuses on the effects of vehicle range, lack of convexity in the trade-off curve, effects of the deviation distance, effects of the function relating the volume of flow captured to the deviation distance, and the effects of multiple shortest paths. Kim and Kuby (2013) apply a network transformation heuristic to mitigate the preprocessing issues encountered by Kuby and Lim (2005) and then solve the deviation-flow FRLM using the “add” and “add-swap” greedy heuristics of Lim and Kuby (2010). Kang and Recker (2014) consider drivers’ routing decisions and thus explicitly allow detours in their location-routing type model. Bhatti et al. (2015) also allow drivers to make detours. In their problems, a facility provider is able to establish refuelling facilities in two tranches. It is then able to learn from the experience of establishing the first set of facilities to better place the remaining refuelling stations. Yildiz et al. (2016) introduce the additional aspect of driver route planning, as in this problem drivers do not necessarily take the shortest path. Improved models are offered and these are solved via branch-and-price.

We note that this problem is somewhat similar to the vehicle routing allocation problem, a subproblem of the family of location-routing problems. (The reader is referred to Nagy and Salhi, 2007, Prodhon and Prins, 2014, and Drexl and Schneider, 2015, for more information.) It may be interesting to develop further location-routing type models for this situation.

2.4 Multi-objective FRLM

The FRLM, like its predecessor the FCLM, is based on the concept of demand arising out of flows (journeys), rather than nodes (customer locations) as in traditional location models. However, already the proponents of the FCLM noted that both type of demand may coexist, see Hodgson and Rosing (1992) and Berman (1997). There is a very limited literature on the FRLM that takes into account also the fact that customers may prefer to refuel close to home, possibly to enable them to undertake short frequent local trips not captured by FRLM models that focus on long-distance journeys. Evidence for the use of this model is limited, with a survey by Kelley and Kuby (2013) concluding that drivers, if faced with a choice, are ten times more likely to refuel at a facility with the shortest detour rather than at the one closest to their home.

Wang and Wang (2010) extend the model of Wang and Lin (2009) to take into account both “intercity and intra-city travel” by adding set covering constraints to account for node coverage. The authors consider the dual objectives of minimising facility cost and maximising population coverage. They discuss the trade-offs between objectives and the influence of the vehicle range and the coverage distance on the solution. Badri-Koohi and Tavakkoli-Moghaddam (2012) observe that the number of stations is fixed in most previous models and state a preference for the model to determine this value. To this end, they consider the costs of establishing

alternative-fuel stations. A long-term planning approach is adopted, taking into account depreciation and inflation. Flow demand in their model is broken up into segments of routes and then approximated by node demand. Moreover, the location model used is continuous, rather than the discrete location model used in all other FRLM papers. As neither of these papers follows the standard FRLM model, there is a literature gap to extend the FRLM as defined by Kuby and Lim (2005) to account for nodal demand.

3 Models for the Alternative-fuel Station Location

3.1 Nomenclature and parameters

In the subsection, we present all notations of indices, sets, parameters and decision variables that have been used in the formulations and the proposed algorithm for the alternative-fuel station location problem. The formulations of Kuby and Lim (2005), Capar and Kuby (2012), MirHassani and Ebrazi (2013), and Capar et al. (2013), denoted by [P1], [P2], [P3] and [P4] respectively, will be described in details in the next subsection. Since some common notations are used in four of the formulations, we notice the referred formulations in the last column. In addition, to present the formulations in a uniform way we changed some notations from the original notations used in the literature. This should hopefully help the reader in appreciating the differences between the formulations.

Indices and sets:

i, j, k, n	Indices for nodes (i.e., station locations/sites)	[P1]-[P4]
q	Index for paths (i.e., origin-destination pairs)	[P1]-[P4]
h	Index for combinations of nodes	[P1]
m, r	Indices of the order of candidate nodes on a given path	[P2]
t	Index of the state of a candidate node ($t = 0$: station built, $t = 1$: station not built)	[P2]
k_m^q	Index i corresponding to the m^{th} candidate node in N_q	[P2]
N	Set of nodes, $N = \{1, 2, \dots, n\}$	[P1]-[P4]
Q	Set of paths	[P1]-[P4]
H	Set of all potential node combinations	[P1]
\bar{N}_h	Set of nodes in combination h : $\bar{N}_h = \{i \in N a_{hi} = 1\}$	[P1]
M_q	Number of candidate nodes within the distance interval $(0, D_q - R/2)$ on path q , if $(D_q - R/2 \leq 0)$, then $M_q = 0$	[P2]
N_q	Set of candidate nodes on path q sorted in sequential order from origin to destination $N_q = \{1 \text{ (origin)}, 2, 3, \dots, n_q \text{ (destination)}\}$	[P2]
A_q	Set of arcs on path q in the original network $G = (N_q, A_q)$	
N_{mt}^q	Set of candidate nodes accessible from the m^{th} candidate node on path q :	[P2]

$$N_{mt}^q = \begin{cases} \{N_q | d_{mr}^q \leq R, r > m\}, & \forall q \in Q, \quad m = 1, 2, \dots, M_q, t = 1, \\ \{N_q | d_{mr}^q < R, r > m\}, & \forall q \in Q, \quad m = 2, \dots, M_q, t = 0, \\ \{N_q | d_{mr}^q \leq R/2, r > m\}, & \forall q \in Q, \quad m = 1, t = 0. \end{cases}$$

Q_i	Subset of Q which contains all the paths passing candidate node i	[P3]
\hat{N}_q	Extended set of candidate nodes on path q (including source and sink dummy nodes)	[P3]
\hat{A}_q	Extended set of arcs in the extended network, $G_E = (\hat{N}_q, \hat{A}_q)$	[P3]
\vec{a}_{jk}	Directed arc starting from node j and ending at node k	[P4]
\vec{A}_q	Set of directed arcs on path q , sorted from origin to destination and back to origin	[P4]
K_{jk}^q	Set of candidate nodes that can refuel the directed arc $\vec{a}_{jk} \in \vec{A}_q$	[P4]

Parameters:

R	Range of vehicles	[P1]-[P4]
f_q	Volume of traffic flow on path q	[P1]-[P4]
p	Number of stations to be located	[P1]-[P4]
a_{hi}	Coefficient that equals to 1 if station i is in combination h and 0 otherwise	[P1]
b_{qh}	Coefficient that equals to 1 if combination h can refuel path q and 0 otherwise	[P1]
d_{mr}^q	Distance between the m^{th} and r^{th} candidate nodes in N_q	[P2]
D_q	Length of the shortest path for path q	[P2]

Decision variables:

$$x_i = \begin{cases} 1 & \text{if station is located at node } i, \\ 0 & \text{otherwise.} \end{cases} \quad [\text{P1}]-[\text{P4}]$$

$$y_q = \begin{cases} 1 & \text{if the flow on path } q \text{ is refueled,} \\ 0 & \text{otherwise.} \end{cases} \quad [\text{P1}], [\text{P2}], [\text{P4}]$$

$$v_h = \begin{cases} 1 & \text{if all facilities in combination } h \text{ are open,} \\ 0 & \text{otherwise.} \end{cases} \quad [\text{P1}]$$

$$c_{m0}^q = \begin{cases} 1 & \text{if } x_{k_m^q} = 0 \text{ and vehicles on path } q \text{ have enough fuel remaining} \\ & \text{at the } m^{\text{th}} \text{ candidate site to be able to reach the next} \\ & \text{opening fuel station on path } q \text{ without running out of fuel,} \\ 0 & \text{if } x_{k_m^q} = 0 \text{ and it cannot reach to an opening fuel station.} \end{cases} \quad [\text{P2}]$$

$$c_{m1}^q = \begin{cases} 1 & \text{if } x_{k_m^q} = 1 \text{ and vehicles on path } q \text{ have enough fuel after} \\ & \text{refueling at the } m^{\text{th}} \text{ candidate site to be able to reach the next} \\ & \text{opening fuel station on path } q \text{ without running out of fuel,} \\ 0 & \text{if } x_{k_m^q} = 1 \text{ and it cannot reach to an opening fuel station.} \end{cases} \quad [\text{P2}]$$

$$y_{ij}^q = \text{flow on an arc } (i, j) \in \hat{A}_q \text{ in the extended network of path } q. \quad [\text{P3}]$$

3.2 Formulations of the alternative-fuel station location

Kuby and Lim (2005) introduced the FRLM, an extension of the FCLM, which locates p refuelling facilities to maximise the number of trips refueled. In the first model, the authors proposed a mixed-integer programming formulation for the nodes-only version of the problem. The model is supported by an algorithm to determine all combinations of nodes that can refuel a path given. The combinations of nodes depend on the length of the path and the maximum vehicle range assumed. Based on the definitions of sets, parameters and decision variables described for [P1] in Section 3.1, the mathematical formulation can be presented as follows

[P1]: (Kuby and Lim, 2005)

$$\max \sum_{q \in Q} f_q y_q \quad (1)$$

$$\text{s.t.} \quad \sum_{h \in H} b_{qh} v_h \geq y_q, \quad \forall q \in Q, \quad (2)$$

$$a_{hi} x_i \geq v_h, \quad \forall h \in H, i \in \bar{N}_h, \quad (3)$$

$$\sum_{i \in N} x_i = p, \quad (4)$$

$$x_i, y_q, v_h \in \{0, 1\}, \quad \forall i \in N, q \in Q, h \in H. \quad (5)$$

The objective function (1) aims to maximise the total traffic volume which can be refueled. Constraints (2) require at least one valid combination of nodes available for path q to be refueled. Constraints (3) keep v_h to be zero unless all the facilities i in combination h are built. Constraint (4) ensure exactly p facilities to be located. Finally, constraints (5) define binary variables.

Unfortunately, generating the valid combinations of nodes H can be computationally burdensome. Some efforts have been performed to reformulate the problem by eliminating the usage of pre-generation of valid combinations. Capar and Kuby (2012) presented such a mixed-binary-integer programming formulation for solving efficiently the problem. By introducing new decision variables c_{m0}^q and c_{m1}^q , along with removing variables v_h relevant to the combinations, and modifying some corresponding constraints, a new mathematical model is formulated by

[P2]: (Capar and Kuby, 2012)

$$\max \sum_{q \in Q} f_q y_q \quad (6)$$

$$\text{s.t.} \quad c_{mt}^q + (-1)^t x_{k_m^q} \leq 1 - t, \quad \forall q \in Q, t \in \{0, 1\}, m = 1, 2, \dots, M_q, \text{ if } M_q \neq 0, \quad (7)$$

$$c_{mt}^q - \sum_{n \in N_{m_t}^q} x_n \leq 0, \quad \forall q \in Q, t \in \{0, 1\}, m = 1, 2, \dots, M_q, \text{ if } M_q \neq 0, \quad (8)$$

$$c_{m1}^q - \sum_{n \in x_{k_m^q} \cup N_{m1}^q} x_n \leq 0, \quad \forall q \in Q, m = M_q + 1, \quad (9)$$

$$c_{m0}^q = 0, \quad \forall q \in Q, m = M_q + 1, \quad (10)$$

$$\sum_{m=1}^{M_q+1} \sum_{t \in \{0,1\}} c_{mt}^q = (M_q + 1)y_q, \quad \forall q \in Q, \quad (11)$$

$$\sum_{i \in N} x_i = p, \quad (12)$$

$$x_i, y_q, c_{m0}^q, c_{m1}^q \in \{0, 1\}, \quad \forall i \in N, q \in Q, m = 1, 2, \dots, M_q. \quad (13)$$

The objective function (6) is identical to that in [P1]. Constraints (7) governs the relationship between variables c_{mt}^q for the m^{th} node on path q and variables x at the same node to ensure that c_{mt}^q equals to one in the right scenario. Constraints (8) governs the relationship between variables c_{mt}^q for the m^{th} node on path q and variables x further along the path to ensure that c_{mt}^q equals to zero unless another refuelling station can be reached from the m^{th} node on path q . Constraints (9) and (10) handle the scenario of nodes within half the range of the destination of path q . Constraints (11) allow the traffic flow on path q to be completely refueled if the sum of the refueled stations (disregard open/close stations) on the path equals to the number of the stations considered (i.e., $M_q + 1$). Constraint (12) is the same with constraint (4) in [P1] that build exactly p facilities. Finally, constraints (13) define binary variables.

Although [P2] is more efficient than [P1] as it eliminates the usage of combination pre-generation, the number of new decision variables and constraints significantly increases the size of the model. Hence, [P2] may not be efficient for real-world applications with large size. MirHassani and Ebrazi (2013) considered the case where the number of facilities is not fixed in advance but is determined from the input data of the facility establishment cost. In this aspect, the problem considered relates to the FRLM like the fixed charge location problem related to the p -median problem, and could thus be called the fixed charge FRLM. The authors proposed a flexible reformulation for a general refuelling station location problem, in which some paths may not be covered by station sites just located at the nodes of the network. In particular, the formulation allows to consider the number of candidate sites modified on the arcs of the paths to refuel the traffic flows by the additional midarc sites. In their paper, the midarc sites were uniformly generated (i.e., identical length of midarcs) along the arc in order to prevent the network from existing uncovered paths. A variation of their formulation was also presented for the case of a set value of p . This includes a new variable y_{sl}^q given by

$$y_{sl}^q = \begin{cases} 1 & \text{if the flow on path } q \text{ cannot be refueled,} \\ 0 & \text{otherwise,} \end{cases}$$

where s and l denote the dummy source and sink nodes respectively, which are added to construct the extended network. In Figure 1, $G = (N_q, A_q)$ represents the original network with the set of candidate nodes and arcs on path q (known as N_q and A_q respectively), where q is the path from node A to D . The information of distance between nodes in the network is given, along with a range of vehicles $R = 100$. In the extended network $G_E = (\hat{N}_q, \hat{A}_q)$, the dummy source and sink nodes are included into the set of candidate nodes (referred to as \hat{N}_q). In addition, other additional arcs with respect to new decision variables y_{ij}^q are built by a four-step procedure in MirHassani and Ebrazi (2013). In particular, an arc is added if the distance between two nodes (defined by the arc) does not exceed the range of vehicle (for example, the additional arcs sB, BD, Cl). Finally, an arc between two dummy nodes is modified to construct the set of arcs \hat{A}_q in the extended network. Although the extended formulation is defined well, its efficacy has not been tested for the FRLM.

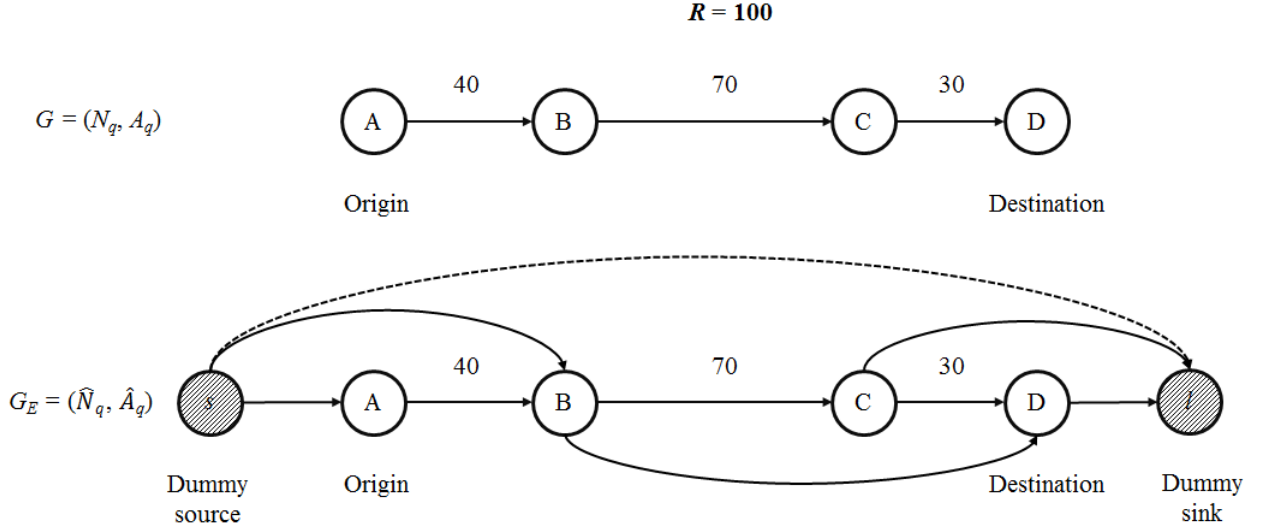


Figure 1: An example of original network vs. extended network for formulation [P3].

[P3]: (MirHassani and Ebrazi, 2013)

$$\max \sum_{q \in Q} f_q (1 - y_{sl}^q) \quad (14)$$

$$\text{s.t.} \quad \sum_{\{j|(i,j) \in \hat{A}_q\}} y_{ij}^q - \sum_{\{j|(j,i) \in \hat{A}_q\}} y_{ji}^q = \begin{cases} 1, & i = s, \\ -1, & i = l, \\ 0, & i \neq s, l, \end{cases} \quad \forall q \in Q, i \in \hat{N}_q, \quad (15)$$

$$\sum_{\{j|(j,i) \in \hat{A}_q\}} y_{ji}^q \leq x_i, \quad \forall i \in N, q \in Q_i, \quad (16)$$

$$\sum_{i \in N} x_i = p, \quad (17)$$

$$x_i \in \{0, 1\}, \quad \forall i \in N, \quad (18)$$

$$y_{ij}^q \geq 0, \quad \forall q \in Q, (i, j) \in \hat{A}_q. \quad (19)$$

The objective function (14) is to maximise the total traffic volume which can be served. Constraints (15) are flow conservation equations for the extended network G_E . Constraints (16) make sure that an inflow passes through a site only if a refuelling station is located at that site. Constraint (17) assures that only p facilities are located. Constraints (18) and (19) define location variables as binary variables and flow variables as nonnegative, respectively.

Capar et al. (2013) developed an arc cover-path-cover formulation for the alternative-fuel station location problem as replacing the usage of combination pre-generation by the concept of that a path can be refueled if all directed arcs on the round-trip path are served. For example, in Figure 2 path q (i.e., A-D pair) can

be refueled if the traffic volumes in all directed arcs (A, B), (B, C), (C, D), (D, C), (C, B), and (B, A) are served. For each directed arc \vec{a}_{jk} on path q , a set of candidate sites K_{jk}^q which can refuel the arc has to be determined based on the vehicle range R . For example, in Figure 2 (with $R = 100$) the set of candidate sites for directed arc \vec{a}_{CD} is {B, C}, since only if refuelling at station B or C we can drive to D without running out of fuel. The arc cover-path-cover mathematical formulation can be described as follows:

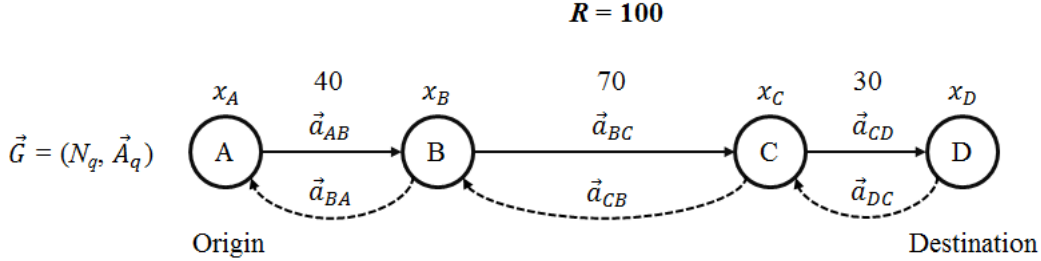


Figure 2: An example of fuel station location on a single path for formulation [P4].

[P4]: (Capar et al., 2013)

$$\max z_0 = \sum_{q \in Q} f_q y_q \quad (20)$$

$$\text{s.t.} \quad \sum_{i \in K_{jk}^q} x_i \geq y_q, \quad \forall q \in Q, \vec{a}_{jk} \in \vec{A}_q, \quad (21)$$

$$\sum_{i \in N} x_i = p, \quad (22)$$

$$x_i, y_q \in \{0, 1\}, \quad \forall i \in N, q \in Q. \quad (23)$$

The objective function (20) is still to maximise the total traffic volume which can be refueled. Constraints (21) represent the innovative formulation in which path q is refuelable if and only if every directed arc along the path is travelable after refuelling at one of the facilities built. Constraint (22) makes sure that only p facilities are located, while constraints (23) are to define binary variables.

A summary comparison among four the formulations based on the number of variables (i.e., binary and continuous) and the number of constraints is shown in Table 2. In the table, $|\cdot|$ is the number of elements in a set. Since some binary variables in the formulations can be relaxed in a mixed-integer program (solved by CPLEX solver) as continuous variables with an upper bound of 1 without impact to the optimal integer solution, to make a fair comparison we relaxed as many as possible binary variables of the formulations. Only the formulation introduced by Kuby and Lim (2005) needs a support procedure to generate valid combinations of stations H . It takes a lot of time to do and is thus known as the less efficient model. The model developed by Capar et al. (2013) is known as the most efficient formulation for the alternative-fuel station location problem so far.

Common assumptions used to formulate the problem consist of

- The traffic flow between an origin-destination pair is only through a single path (i.e., shortest path) and its volume is given.

Table 2: A comparison of the sizes of the four formulations for the alternative-fuel station location problem.

Formulation	No. of binary variables	No. of continuous variables	No. of constraints
[P1]	$ N $	$ H + Q $	$\sum_{h=1}^H \bar{N}_h + Q + 1$
[P2]	$ N + Q $	$2 \sum_{q=1}^Q M_q $	$4 \sum_{q=1}^Q M_q + 3 Q + 1$
[P3]	$ N $	$\sum_{q=1}^Q \hat{A}_q $	$\sum_{q=1}^Q \hat{N}_q + \sum_{i=1}^N Q_i + 1$
[P4]	$ N $	$ Q $	$\sum_{q=1}^Q \vec{A}_q $

- Drivers have full knowledge of refuelling station locations along their path and know when to refuel in order to complete their trip without running out of fuel.
- All vehicles have the same vehicle range R .
- Only nodes in the network are considered to be refuelling station locations.
- Refuelling stations are uncapacitated.
- Fuel consumption is proportional to traveling distance.

4 Possibilities for Future Research

The design of a heuristic algorithm for the alternative-fuel station location problem is an important issue that has not received appropriate attention in the research. The authors are currently working on a novel heuristic method to solve this problem. If it is successful, we can extend it to solve other variants, such as:

- the capacitated FRLM (Upchurch et al., 2009),
- the FRLM with deviation (Kim and Kuby, 2012, Kim and Kuby, 2013), or
- the fixed-charge FRLM (Badri-Koohi and Tavakkoli-Moghaddam, 2012, MirHassani and Ebrazi, 2013).

In practice, it is unlikely that all stations of a network would be established at once, as considered in most of the literature. It is more likely that refuelling facilities would be opened to serve an increasing demand for alternative fuels, see Bhatti et al. (2015). In this respect, an extension of the FRLM to a model similar to dynamic location-routing could be considered. Dynamic location-routing, see Nambiar et al. (1989), Salhi and Nagy (1999) and Albareda-Sambola et al. (2012), takes into account the fact that facility locations are hard to change once established, but demand and resulting routes may change frequently. We believe a dynamic FRLM could be a very interesting area for future research.

It would be very interesting to tackle practical applications, which we believe may arise in the near future, especially in the light of the recent EU directive on the establishment of a Europe-wide alternative fuel

infrastructure (European Commission, 2014). One possible application would be for the location of alternative fuel stations for the railways. While the algorithms presented in the literature could be just as applicable to rail transport as to automobiles, most papers tackle the FRLM in the context of automobile refuelling stations. Yet, as Kuby and Lim (2005) has already pointed out, there is much better origin-destination flow data available for railways, making this mode of transport an ideal field of applying FRLM models.

References

- M Albareda-Sambola, E Fernandez, and S Nickel. Multiperiod location-routing with decoupled time scales. *European Journal of Operational Research*, 217(2):248–258, 2012.
- B Badri-Koochi and R Tavakkoli-Moghaddam. Determining optimal number and locations of alternative-fuel stations with a multi-criteria approach. *Proceedings of the 8th International Industrial Engineering Conference*, 2012.
- O Berman. Deterministic flow-demand location problems. *Journal of the Operational Research Society*, 48(1):75–81, 1997.
- O Berman, MJ Hodgson, and D Krass. *Flow intercepting models*. In: Z. Drezner (Ed.), *Facility Location: A Survey of Applications and Methods*, pages 389–426. Springer, New York, 1995.
- SF Bhatti, MK Lim, and HY Mak. Alternative fuel station location model with demand learning. *Annals of Operations Research*, 230(1):105–127, 2015.
- I Capar and M Kuby. An efficient formulation of the flow refueling location model for alternative-fuel stations. *IIE Transactions*, 44(8):622–636, 2012.
- I Capar, M Kuby, VJ Leon, and YJ Tsai. An arc cover-path-cover formulation and strategic analysis of alternative-fuel station locations. *European Journal of Operational Research*, 227(1):142–151, 2013.
- European Commission. http://europa.eu/rapid/press-release_MEMO-13-24_en.htm, 2013. Accessed 10/12/2014.
- European Commission. http://europa.eu/rapid/press-release_IP-14-1053_en.htm, 2014. Accessed 10/12/2014.
- M Drexl and M Schneider. A survey of variants and extensions of the location-routing problem. *European Journal of Operational Research*, 241(2):283–308, 2015.
- MJ Hodgson. The location of public facilities intermediate to the journey to work. *European Journal of Operational Research*, 6(2):199–204, 1981.
- MJ Hodgson. A flow-capturing location-allocation model. *Geographical Analysis*, 22(3):270–279, 1990.
- MJ Hodgson and O Berman. A billboard location model. *Geographical and Environmental Modeling*, 1(1):25–45, 1997.
- MJ Hodgson and KE Rosing. A network location-allocation model trading off flow capturing and p-median objectives. *Annals of Operations Research*, 40(1):247–260, 1992.

- MJ Hodgson, KE Rosing, and J Zhang. Locating vehicle inspection stations to protect a transportation network. *Geographical Analysis*, 28(4):299–314, 1996.
- JE Kang and W Recker. Strategic hydrogen refueling station locations with scheduling and routing considerations of individual vehicles. *Transportation Science*, 49(4):767–783, 2014.
- S Kelley and M Kuby. On the way or around the corner? Observed refueling choices of alternative-fuel drivers in southern california. *Journal of Transport Geography*, 33:258–267, 2013.
- JG Kim and M Kuby. The deviation-flow refueling location model for optimizing a network of refueling stations. *International Journal of Hydrogen Energy*, 37(6):5406–5420, 2012.
- JG Kim and M Kuby. A network transformation heuristic approach for the deviation flow refueling location model. *Computers and Operations Research*, 40(4):1122–1131, 2013.
- M Kuby and S Lim. The flow-refueling location problem for alternative-fuel vehicles. *Socio-Economic Planning Sciences*, 39(2):125–145, 2005.
- M Kuby, L Lines, R Schultz, Z Xie, JG Kim, and S Lim. Optimization of hydrogen stations in Florida using the flow-refueling location model. *International Journal of Hydrogen Energy*, 34(15):6045–6064, 2009.
- S Lim and M Kuby. Heuristic algorithms for siting alternative-fuel stations using the flow-refueling location model. *European Journal of Operational Research*, 204(1):51–61, 2010.
- SA MirHassani and R Ebrazi. A flexible reformulation of the refueling station location problem. *Transportation Science*, 47(4):617–628, 2013.
- G Nagy and S Salhi. Location-routing: Issues, models and methods. *European Journal of Operational Research*, 177(2):649–672, 2007.
- JM Nambiar, LF Gelders, and LN Van Wassenhove. Plant location and vehicle routing in the malaysian rubber smallholder sector: a case study. *European Journal of Operational Research*, 38(1):14–26, 1989.
- C Prodhon and C Prins. A survey of recent research on location-routing problems. *European Journal of Operational Research*, 238(1):1–17, 2014.
- S Salhi and G Nagy. Consistency and robustness in location-routing. *Studies in Locational Analysis*, 13:3–19, 1999.
- C Upchurch, M Kuby, and S Lim. A model for location of capacitated alternative-fuel stations. *Geographical Analysis*, 41(1):85–106, 2009.
- YW Wang and CC Lin. Locating road-vehicle refueling stations. *Transportation Research E*, 45(5):821–829, 2009.
- YW Wang and CR Wang. Locating passenger vehicle refueling stations. *Transportation Research E*, 46(5):791–801, 2010.
- M Wen, G Laporte, OBG Madsen, AV Norrelund, and A Olsen. Locating replenishment stations for electric vehicles: application to danish traffic data. *Journal of the Operational Research Society*, 65(10):1555–1561, 2014.

B Yildiz, O Arslan, and OE Krasan. A branch and price approach for routing and refueling station location model. *European Journal of Operational Research*, 248(3):815–826, 2016.

University of Kent