MODELS AND HEURISTICS FOR THE FLOW-REFUELLING LOCATION PROBLEM

Gábor Nagy
University of Kent
Canterbury CT2 7FS, U.K.
G.Nagy@kent.ac.uk

Trung Hieu Tran
Nottingham University, U.K.

Thu Ba Nguyen
University of Southampton, U.K.

Niaz Wassan
University of Kent, U.K.

ABSTRACT
Purpose of this paper: Firstly, the paper serves as an overview of the emerging field of flow-refuelling location, which mainly occurs in the context of locating alternative-fuel (hydrogen, electric, liquefied natural gas and hybrid) vehicle refuelling stations. We aim to review and explain models and solution approaches, with a particular focus on mathematical programming formulations. Secondly, we propose a new heuristic for this problem and investigate its performance.

Design/methodology/approach: The subject scope of this paper is the flow-refuelling location model (FRLM). While in most location problems demand arises at customer locations, in so-called flow-capturing models it is associated with journeys (origin-destination pairs). What makes the FRLM even more challenging is that due to the limited driving range of alternative-fuel vehicles, more than one facility may be required to satisfy the demand of a journey. There are currently very few such refuelling stations, but ambitious plans exist for massive development – making this an especially ripe time for researchers to investigate this problem. There already exists a body of work on this problem; however different authors make different model assumptions, making comparison difficult. For example, in some models facilities must lie on the shortest route from origin to destination, while in others detours are allowed. We aim to highlight difference in models in our review. Our proposed methodology is built on the idea of solving the relaxation of the mixed-integer linear programming formulation of the problem, identifying promising variables, fixing their values and solving the resulting (so-called restricted) problems optimally. It is somewhat similar to Kernel Search which has recently gained popularity. We also use a parallel computing strategy to simultaneously solve a number of restricted problems with less computation effort for large-sized instances.

Findings: Our experimental results show that the proposed heuristic can find optimal solutions in a reasonable amount of time, outperforming other heuristics from the literature.

Value: We believe the paper is of value to both academics and practitioners. The review should help researchers new to this field to orient themselves in the maze of different problem versions, while helping practitioners identify models and approaches applicable to their particular problem. The heuristic proposed can be directly used by practitioners; we hope it will spark further works on this area of logistics but also on other optimisation problems where Kernel Search type methods can be applied.

Research limitations: This being the first paper applying a restricted-subproblem approach to this problem it is necessarily limited in scope. Applying a traditional Kernel Search method would be an interesting next step. The proposed heuristic should also be extended to cover for more than just one FRLM model: certainly the capacitated FRLM, the FRLM with deviation, the fixed-charge FRLM and the multi-period FRLM should be investigated.

Practical implications: Our work adds to a body of research that can inform decision-makers at governmental or international level on strategic decisions relating to the establishment or development of alternative-fuel refuelling station networks.

INTRODUCTION

The flow-refuelling location problem is a logistics problem that mainly occurs in the context of locating alternative-fuel (hydrogen, electric, liquefied natural gas and hybrid) vehicle refuelling stations. Alternative-fuel station location is a recent, but very applicable research topic within logistics. 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. In fact, alternative-fuel vehicles require a very dense refuelling infrastructure, as these vehicle typically have a short driving range. 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. 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 charging stations. Thus, this is the right time for Logistics researchers to devote their energies to finding optimal or near-optimal locations for alternative-fuel refuelling facilities.

We first review the literature of this problem, including the mathematical models proposed. Then, we present our new heuristic algorithm and show our numerical experimentation with it. Finally, a brief summary and ideas for future research are given.

LITERATURE REVIEW

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). This model is based on the concept of "locating facilities on the home-to-work journey" and the observation that in some cases it makes more sense to locate facilities near routes that customers already take. The author showed that basing locational decisions simply on arc flow volumes is not sufficient; instead, models should be based on detailed origin-destination flow data. 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.

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. Such vehicles normally can cover a shorter distance on a full tank than traditional gas-guzzling vehicles. In addition, the availability of alternative fuel refuelling stations is very limited. However, the model is also applicable to the location of conventional refuelling stations in developing countries where infrastructure is as yet lacking. Likewise, it can be applicable to territories with sparse population (and hence sparse refuelling infrastructure).

Kuby and Lim (2005) introduced 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, like most subsequent formulations, is based on binary decision variables showing whether a station is open at a node and whether a given path has its flow refuelled. However, it also contains a more cumbersome variable that shows whether every facility in a given combination is open. Unfortunately this formulation 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 their benchmark instance, let alone solve it.

While in a large part of the literature the objective is to maximise the flow captured, the model of Wang and Lin (2009) aims to minimise costs such that all flows are served. The authors devised a "vehicle refuelling logic" that is more involved but also more flexible than that of previous models. Another important difference is that this model requires only knowledge of origin-destination distances, but not of origin-destination flow data.

Lim and Kuby (2010) designed 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 authors found that the greedy algorithms perform quite well, nearly as well as the genetic algorithms, and are significantly faster.

Capar and Kuby (2012) put forward a more complex mathematical formulation, but without the preprocessing requirement of Kuby and Lim (2005). This new formulation is in fact as fast as the greedy heuristics of Lim and Kuby (2010). They replaced the decision variables relating to facility combinations with variables that show whether vehicle on a given path and refuelling (or not refuelling) at a station candidate site have enough fuel remaining to reach the next open fuel station on their path. This is a more efficient formulation in that combination pre-generation is eliminated the number of new decision variables and new constraints significantly increase the size of the model.

Capar et al. (2013) offered 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. It is based on the concept that a path can be refuelled if all directed arcs on the round-trip path are served. This eliminates both combination pre-generation and the cumbersome refuelling logic variables of Capar and Kuby (2012).

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. This version of the FRLM is known as the fixed-charge FRLM. However, their formulation is adapted also for the case of fixed number of facilities. The logic of their formulation is developed from a single-path to a multi-path formulation. This necessitates the creation of a so-called extended network. The authors provide this formulation but their computation testing was only on the fixed-charge FRLM therefore it is interesting to see how their model compares to the Capar et al. (2013) formulation.

Wen et al. (2014) investigated both maximal flow capture and total flow capture models. Their formulations are based on set covering and do not require the evaluation of all feasible combinations of locations. Ghamami et al. (2016) considered the particular case of locating refuelling stations along a travel corridor, while also allowing for congestion and delay at charging stations. Their formulation is based on the assignment problem. The authors have also designed a simulated annealing metaheuristic.

Finally we note that in this brief review it was not possible for us to describe all variants of the FRLM. Of particular note is the FRLM with deviation. While in the above models it is assumed that in order to capture a flow, a facility must lie on the origin-destination path, 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. The reader is referred to Berman et al. (1995), Kim and Kuby (2012), Yildiz et al. (2016) and Lin and Lin (2016). Another interesting problem is the multi-period FRLM, see Miralinaghi et al. (2017).

A NEW HEURISTIC FOR THE FRLM

The idea of using the optimal solution obtained by relaxing the integrality constraints of the mixed-integer linear programming (MILP) problems to generate a set of initial solutions for meta-heuristic algorithms is well-known. Recently, the idea of using the information of the optimal solution to support search process further – namely, to establish a set of promising candidate variables – was developed by Angelelli et al. (2010). This method, known as Kernel Search, identifies subsets of decision variables for the MILP problem by solving the relaxation problem and then solves the restricted problems to optimality by commercial MILP solvers. It has been successfully applied for several optimisation problems, including logistics applications. For example, Guastaroba and Speranza (2012, 2014) solved the multi-source and the single-source capacitated facility location problems, respectively. We develop here an efficient heuristic algorithm, based on the concept of restricted subproblems, for solving the alternative-fuel station location problem. Compared with Kernel Search algorithm, our algorithm has some small differences as follows:

- Although there are two sets of binary variables in the problem, we only explicitly restrict on the location variables. The number of path variables is determined based on the restricted location variables. Thus, there is implicit restriction. Restricting on one set of binary variable may help the proposed algorithm obtain the good balance of solution quality and CPU time, instead of restricting on all the sets of binary variables as in Kernel Search algorithm.
- The size of the restricted subproblems equals to the number of location variables with positive value. In Kernel Search algorithm, size of the restricted subproblems is usually a given arbitrary parameter. As a result, initial promising variable set may include

variables with zero relaxed value or remove variables with positive relaxed value. This may lead to spend additional CPU time to find the best solutions.
• A 2-exchange neighbourhood is used to generate a pool of the restricted subproblems for parallel computing strategy.

We use the formulation of Capar et al. (2013) as the basis. Solving the relaxation of this yields us the set of promising nodes – those with non-negative relaxed values. To further reduce the size of the subproblems to be solved, we remove paths that cannot be refuelled by a set of restricted nodes. Then, we generate a set of restricted subproblems by performing 2-exchange on the set of promising nodes. These can be grouped and each group allocated to a CPU core to enable parallel processing.

In more detail, the algorithm can be described as follows:
1. Solve the original problem relaxed on the constraints of binary variables. If solution is integer, stop.
2. Set upper bound and classify vertices into promising nodes N^* (those with positive value) and other nodes N^0 (those with zero value).
3. Determine restricted set of paths based on N^*. Solve restricted subproblem and update bounds. If bounds sufficiently close to each other, stop.
4. Create several sets of restricted nodes by exchanging 2 nodes between N^* and N^0. Allocate these to parallel CPU cores and solve restricted subproblems simultaneously. Update upper and lower bounds.
5. If bounds are close to each other *or* all restricted problems have been solved *or* a given number of restricted problems have been solved, stop; else return to Step 4.

COMPUTATIONAL EXPERIENCE

In this section, we investigate the computational efficacy of solving the FRLM with the heuristic algorithm proposed. We evaluate the performance of the heuristic algorithm on two well-known benchmark datasets and then compare the obtained results with the optimal solutions from CPLEX solver as well as other heuristic algorithms. The models and the proposed algorithm were implemented in Visual C++; the models were built and solved using the IBM ILOG CPLEX version 12.4.

The computational experiments were run on two well-known benchmark datasets:
• Hodgson dataset (Hodgson, 1990): This is a 25-node alternative-fuel station location network. The flow volumes in the 25 × 25 origin-destination matrix are estimated using a gravity model. The flows are then assigned to their shortest paths. The candidate sites are limited to the 25 nodes of the network. The network has 300 origin-destination pairs.
• Florida dataset (Kuby et al., 2009): This is a Florida state highway network consisting of 302 nodes (i.e. junctions) and 495 arcs. Each of the nodes serves as a candidate site. Of the 302 candidate sites, there are 74 origin-destination nodes for trips. Since the return trip is assumed to be refuelable, the network of 74 origin-destination nodes only requires 2701 unique origin-destination pairs.
For the evaluation of computational experiments, a set of scenarios are generated by changing the range of vehicles R and the number of stations to be located p. R = 4, 8, and 12 are used for Hodgson network and R = 100 is used for Florida network. Both are tested with p = 5, 10, 15, 20 and 25.

The instances were solved using the formulations of Capar and Kuby (2012) [CK], Capar et al. (2013) [CKLT], MirHassani and Ebrazi (2013) [ME], the genetic and greedy algorithms of Lim and Kuby (2010) [LKGenA and LKGreA] and our heuristic [HA]. Results are presented in Tables 1 and 2. Note Δ stands for percentage deviation from optimum and time is given in seconds. Lim and Kuby (2010) did not give computing times for the Hodgson dataset.

Table 1: Comparison for the formulations and the algorithms on Hodgson instances.

		CK		CKLT		ME		LKGenA		LKGreA		HA	
R	p	Δ	time	Δ	time	Δ	time	Δ	time	Δ	time	Δ	time
4	5	0	0.24	0	0.12	0	0.13	0	-	0	-	0	0.12
	10	0	0.20	0	0.14	0	0.12	0	-	10.75	-	0	0.13
	15	0	0.18	0	0.08	0	0.10	0	-	0	-	0	0.05
	20	0	0.18	0	0.08	0	0.10	0	-	0	-	0	0.05
	25	0	0.16	0	0.10	0	0.10	0	-	0	-	0	0.05
8	5	0	0.28	0	0.14	0	0.14	0	-	0.49	-	0	0.15
	10	0	0.32	0	0.18	0	0.18	0	-	0.93	-	0	0.26
	15	0	0.28	0	0.14	0	0.15	0	-	0	-	0	0.27
	20	0	0.20	0	0.12	0	0.11	0	-	0.15	-	0	0.05
	25	0	0.18	0	0.10	0	0.10	0	-	0	-	0	0.06
12	5	0	0.84	0	0.16	0	0.44	0	-	0	-	0	0.21
	10	0	0.30	0	0.16	0	0.16	0	-	0	-	0	0.07
	15	0	0.24	0	0.14	0	0.13	0	-	0	-	0	0.07
	20	0	0.28	0	0.14	0	0.14	0	-	0	-	0	0.07
	25	0	0.20	0	0.12	0	0.11	0	-	0	-	0	0.07
Average		0	0.27	0	0.13	0	0.15	0	-	0.82	-	0	0.11

Table 2: Comparison for the formulations and the algorithms on Florida instances.

		CK		CKLT		ME		LKGenA		LKGreA		HA	
R	p	Δ	time	Δ	time	Δ	time	Δ	time	Δ	time	Δ	time
100	5	0	148	0	34	0	76	0.71	182	0	33	0	28
	10	0	210	0	29	0	101	0.41	1352	0.43	174	0	29
	15	0	341	0	32	0	165	0.16	2412	0.50	456	0	30
	20	0	492	0	75	0	251	0.10	5800	0.31	1132	0	74
	25	0	999	0	138	0	512	0.10	10646	0.38	1800	0	129
Average		0	438	0	62	0	221	0.30	4078	0.32	719	0	58

Both tables confirm the efficiency of our heuristic algorithm. It finds the optimal solution for all the instances tested above, with computing times on most instances slightly below those of Capar et al. (2013). Our algorithm significantly outperforms both heuristics of Lim and Kuby (2010) in terms of solution quality and CPU time. For example, we only need 129 seconds to find the optimal solution for the Florida instance (R=100, p=25), while LKGenA and LKGreA take about 3 hours and about half an hour respectively, yet do not find the optimal solution. As an additional observation, we can see that the MirHassani and Ebrazi (2013) formulation is slightly less efficient than the Capar et al. (2013) formulation – these have not previously been compared to each other.

CONCLUSIONS AND SUGGESTIONS

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. In the paper, we thus develop an efficient heuristic algorithm to locate optimal refuelling stations for the maximisation of round-trip traffic volume. The algorithm is constructed on solving the sequence of restricted problems by a set of promising station candidates, and by a number of the best promising stations to be located. To determine the initial set of candidates we solve a relaxation model of the original problem with the constraints of integer variables relaxed, and then update the set in next iterations by performing 2-exchange between the set of promising candidates and the remaining station set. As solving the restricted problems, we locate the best stations in the set of promising candidates to improve the computation time of the algorithm. Besides that, we use a parallel computing strategy to simultaneously solve a number of restricted problems with less computation effort for large-sized instances. Experimental results show that the proposed algorithm can obtain the optimal solutions with less computation time (compared with CPLEX solver), and outperforms the other compared algorithms (i.e., genetic algorithm and greedy algorithm) with respect to solution quality as well as computation time.

From the successful results obtained, we can extend the heuristic algorithm to handle other interesting alternative-fuel station location problems, such as:
• The FRLM with deviation (Kim and Kuby, 2012, Lin and Lin, 2016),
• The fixed-charge FRLM (MirHassani and Ebrazi, 2013),
• The multi-period FRLM (Miralinaghi et al., 2017).
It would be very interesting to apply our algorithm to 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). In this respect, the reader is referred to the recent application study by Kuby et al. (2017), focusing on natural gas refuelling stations in the E.U. Another 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.

ACKNOWLEDGMENTS
This work has been funded by a grant from the Spanish Ministry of Economy and Competitiveness (ECO2011-24927).

REFERENCES
E Angelelli, R Mansini, and MG Speranza. Kernel search: A general heuristic for the multi-dimensional knapsack problem. *Computers and Operations Research*, 37(11):2017–2026, 2010.
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.
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_IP-14-1053_en.htm, 2014. Accessed 01/02/2017.
M Ghamami, A Zockaie, and YM Nie. A general corridor model for designing plug-in electric vehicle charging infrastructure to support intercity travel. *Transportation Research Part C: Emerging Technologies*, 68:389–402, 2016.
G Guastaroba and MG Speranza. Kernel search for the capacitated facility location problem. *Journal of Heuristics*, 18(6):877–917, 2012.
G Guastaroba and MG Speranza. A heuristic for BILP problems: The single source capacitated facility location problem. *European Journal of Operational Research*, 238(2):438–450, 2014.
MJ Hodgson. A flow-capturing location-allocation model. *Geographical Analysis*, 22(3):270–279, 1990.
JG Kim and M Kuby. The deviation-flow refueling location model for optimizing a network of refuelling stations. *International Journal of Hydrogen Energy*, 37(6):5406–5420, 2012.
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.

M Kuby, I Capar, and JG Kim. Efficient and equitable transnational infrastructure planning for natural gas trucking in the European Union. *European Journal of Operational Research*, 257(3):979–991, 2017.

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.

CC Lin and CC Lin (2016): Locating alternative fuel refueling stations to minimize drivers' deviations from their planned routes. *The Proceedings of 21st International Symposium on Logistics*, 197-204.

M Miralinaghi, BB Keskin, Y Lou, and AM Roshandeh. Capacitated refueling station location problem with traffic deviations over multiple time periods. *Networks and Spatial Economics*, 17:129–151, 2017.

SA MirHassani and R Ebrazi. A flexible reformulation of the refueling station location problem. *Transportation Science*, 47(4):617–628, 2013.

YW Wang and CC Lin. Locating road-vehicle refueling stations. *Transportation Research E*, 45(5):821–829, 2009.

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