The locomotive fleet fueling problem

Tal Raviv[1] and Mor Kaspi
Department of Industrial Engineering, Tel Aviv University, Tel Aviv, Israel

Abstract: This paper considers the problem of how to determine an optimal fueling schedule and contracting policy with fuel suppliers so as to minimize the total cost of the fueling operation. The problem is formulated as a mixed integer program and the formulation is enhanced by valid inequalities and domination rules. The enhanced model allows us to obtain near optimal solutions for large scale instances.

Keyword: Railway Application, mixed integer programming

1. Introduction

The refueling operation represents a significant share of the operational costs in the transportation industry and in particular in railway transportation, see Gray [1]. The refueling cost consists of three components: 1) the cost of the fuel; 2) indirect costs caused by the delay of trains when their locomotives are being refueled or waiting to be refueled; 3) the contracting cost paid to owners of tanker trucks that deliver the fuel to the various yards. Fuel prices and the contracting costs vary across locations due to transportation costs from the refineries, local taxes, competitiveness of the regional market and other factors. The delay cost is determined by various characteristics of the train being delayed and of the yard.

Fueling at each yard is carried out by tanker trucks contracted on a yearly or quarterly basis. A typical truck can provide a given volume of fuel per day. This figure is affected by the proximity of the yard to the nearest refinery. The railway company may contract several trucks at a yard to allow provision of a larger amount of fuel. The railway planner simultaneously decides how many tanker trucks and in which yards to contract and where to refuel each locomotive along its prescheduled itinerary.

A simplified version of this problem was introduced as a challenge for the 2010 Problem Solving Competitions of INFORMS Railway Application Section (RAS) [2]. The authors of this paper won the competition. A similar problem was studied by Nourbakhsh and Ouyang [3], where a heuristic, based on a Lagrangian relaxation was presented. Their paper contains a comprehensive review of recent literature on various refueling problems.

Previously introduced methods could not solve large real world instances of the locomotive fleet fueling problem to exact optimality. The main contribution of this paper is in presenting a tight mathematical formulation that enables the solving of such instances with negligible optimality gap.

The rest of this paper is organized as follows: In Section 2, the fleet fueling problem is formally defined, mathematically formulated, and proven to be NP-Hard. In Section 3, this formulation is enhanced by several sets of strong valid inequalities and domination rules. Possible extensions of the problem that may be applicable for various practical situations are introduced in Section 4. Section 5 demonstrates the effectiveness of the proposed method using a numerical study.

[1] Corresponding author, e-mail address: talraviv@eng.tau.ac.il

2. Notation and mathematical formulation

The fleet refueling model consists of two basic entities: yards and locomotives. Each yard is characterized by its fuel cost and by the amount of dispensable fuel per tanker, if contracted. Each locomotive is characterized by its schedule for the planning horizon. The schedule is described by a sequence of stops. Each stop is characterized by yard, date, fuel consumption (for the journey to next stop) and delay cost (to be charged if the locomotive refuels at this stop).

The delay cost is related to the stops because it corresponds to the value of the train being pulled by the locomotive at the particular time of each stop. For example, if the locomotive is scheduled to dwell at the station for some time, the delay cost may be zero. For similar considerations, the fuel consumption is stop-related as well. This representation allows us to abstract the physical topology of the railway network from the mathematical model, thus simplifying the notation.

The fleet refueling problem is to minimize the total fuel, delay and contracting costs subject to capacity constraints of the locomotives tanks and the contracted tankers. The decision variables are the number of tankers to be contracted at each yard and the amount of fuel that each locomotive has to acquire at each stop. Next, the notation required to formulate this problem as an integer programming model is presented.

Indices

i	Locomotive
j	Stop of a locomotive (the stops are indexed 1,2,…)
k	Yard
d	Days

Summation and quantification over these indices are assumed to refer to all related entities unless otherwise stated.

Parameters

Θ	Length of the planning horizon (days)
$D_{i,j}$	Delay cost of locomotive i at stop j
LT_i	Capacity (gallons) of the fuel tank of locomotive i
TT_k	Amount (gallons) of fuel that can be dispensed by truck tank each day. This parameter is referred to, in short, as the tanker truck capacity.
CC_k	Contracting cost ($) of a truck at yard k for the planning horizon
$FC_{i,j}$	Fuel consumption (gallons) of locomotive i on its journey from stop j to $j+1$
P_k	Fuel price ($/gallon) at yard k
N_i	Number of stops of locomotive i in the planning horizon
$yard(i,j)$	Yard of the j^{th} stop of locomotive i
$day(i,j)$	Day in which the j^{th} stop of locomotive i occurs.

Decision Variables

$x_{i,j}$	Binary variable that equals 1 if locomotive i refuels at its j^{th} stop
y_k	Integer variable that denotes the number of fuel tanks to be located at yard k
$f_{i,j}$	Amount of fuel acquired by locomotive i at its j^{th} stop.
$v_{i,j}$	Amount of fuel in the tank of locomotive i upon arrival at its j^{th} stop. The value of $v_{i,1}$ is

a fixed parameter.

A MILP formulation of the problem is presented next,

$$minimize \sum_k CC_k \cdot y_k + \sum_{i,j} \left(D_{i,j} \cdot x_{i,j} + P_{yard(i,j)} \cdot f_{i,j} \right) \quad (1)$$

Subject to

$$v_{i,j} = v_{i,j-1} + f_{i,j-1} - FC_{i,j-1} \qquad \forall i, j = 2, \dots, N_i \quad (2)$$

$$v_{i,j} + f_{i,j} \le LT_i \quad \forall i, j \quad (3)$$

$$f_{i,j} \le LT_i x_{i,j} \qquad \forall i, j \quad (4)$$

$$\sum_{(i,j):\, yard(i,j) = k, day(i,j) = d} f_{i,j} \le TT_k \cdot y_k \quad \forall k, d \quad (5)$$

$$y_k \text{ integer}, x_{i,j} \text{ binary}, \ f_{i,j}, v_{i,j} \ge 0 \quad (6)$$

The objective function (1) sums over the three types of costs: namely tanker truck contracting costs, delay costs and total cost of acquired fuel. Constraints (2) are inventory balance equations that update the amount of fuel in the locomotive tank at each stop based on the fuel consumption since the last stop and the amount of fuel acquired there. Constraints (3) define the capacity of the locomotives tanks. Constraints (4) stipulate that a fixed fueling cost is charged whenever a locomotive fuels at a certain stop. Constraints (5) assure that the number of tanker trucks contracted for each yard is enough to provide the amount of fuel acquired at the yard during each day of the planning horizon. Integrality and binarity of $x_{i,j}$ and y_k are determined in Constraints (6). The non-negativity constraints of $v_{i,j}$ stipulate that the locomotives never run out of fuel. The non-negativity of $f_{i,j}$ implies that the locomotive cannot discharge fuel to the tanker trucks. The computational complexity of this problem is studied next:

Proposition 1: The locomotive fleet fueling problem is strongly NP-Hard and APX-Complete.

Proof: We use a reduction from the minimum vertex cover problem. Recall that a vertex cover of a graph $G = (V, E)$ is a subset $V' \subset V$ such that for each edge $\{u, v\} \in E$ at least one of u and v belongs to V'. A minimum vertex cover of a graph is a vertex cover V' of minimum cardinality. This problem is known to be NP-Hard [5] and APX-Complete, see Papadimitriou and Yannakakis [6].

Consider an instance $G = (V, E)$ of the minimum vertex cover problem and construct an instance of the locomotive fleet fueling problem in the following manner: a yard for each vertex in V and a locomotive for each edge $\{u, v\} \in E$. Each locomotive starts at yard u, goes to v and then back to u. The initial fuel inventory of each locomotive is half of its tank capacity. The fuel consumption for the segment between u and v is half the locomotive tank for each direction. The fuel costs and contracting costs are equal for all yards. The delay costs are equal for all stops. The tanker truck capacity is set to the total capacity of the tanks of all the trains that pass via its corresponding yard, i.e., Constraint (5) is not binding.

It is easy to see that in an optimal solution each locomotive refuels once either at its first stop at yard u or at its second stop at yard v. Now, since the fuel cost and delay cost are identical at both yards, an optimal solution is one that minimizes the number of contracted tankers. ■

3. Strengthening the formulation

While the basic MILP formulation (1)-(6) is capable of obtaining feasible solutions of real life instances of the problem with an optimality gap of a few percent, closing this optimality gap may result in a substantial saving for the railway company. Next, several enhancements of the basic mathematical model that allows solving large instances of the problem with small optimality gaps are presented.

The stops subsequence inequality – consider a sequence of stops $j_1, \dots, j_a$ along the route of locomotive i such that the total amount of fuel consumed by the locomotive between stop $j_1 - 1$ to $j_a + 1$ is greater than the capacity of the locomotive's tank (that is $\sum_{j=j_1-1}^{j_a} FC_{i,j} > LT_i$), but the total amount of fuel consumed between j_1 and $J_a + 1$ is not (that is $\sum_{j=j_1}^{j_a} FC_{i,j} \leq LT_i$). In such a case the locomotive must refuel at least once in one of the stops $j_1, \dots, j_a$. Based on this observation the following inequalities can be added to the mathematical formulation:

$$\sum_{j=j_1}^{j_a} x_{i,j} \geq 1 \qquad \forall i, 1 \leq j_1 \leq j_a \leq N_i : \left(\sum_{j=j_1-1}^{j_a} FC_{i,j} > LT_i \right) \text{ and } \left(\sum_{j=j_1}^{j_a} FC_{i,j} \leq LT_i \right). \tag{7}$$

Similarly, it is possible to derive a **yard subsequence inequality** for each minimal sequence of yards along the route of each locomotive that requires at least one fueling operation and thus at least one tanker truck due to the same consideration as for the stops subsequence inequality. These inequalities can be formulated as follows:

$$\sum_{k:(\exists (j_1 \leq j \leq j_a): yard(i,j)=k)} y_k \geq 1$$
$$\forall i, 1 \leq j_1 \leq j_a \leq N_i : \left(\sum_{j=j_1-1}^{j_a} FC_{i,j} > LT_i \right) \text{ and } \left(\sum_{j=j_1}^{j_a} FC_{i,j} \leq LT_i \right). \tag{8}$$

Stops per locomotive inequality – Let $MinFuels(i)$ denote a lower bound on the number of fueling operations that each locomotive needs to complete during the planning horizon in any feasible solution. Clearly the following inequality is valid for each locomotive,

$$\sum_{j=1}^{N_i} x_{i,j} \geq MinFuels(i) \qquad \forall i. \tag{9}$$

A naïve approach to obtain $MinFuels(i)$ is by

$$MinFuels(i) = \left\lceil \frac{\sum_{j=1}^{N_i} FC_{i,j}}{LT_i} \right\rceil. \tag{10}$$

However, note that it is not necessarily a tight bound. For example, if the route consists of four yards that are $\frac{2}{3}$ tank apart, it is clear that the locomotive must be refueled at each station, i.e., three times, while the lower bound obtained from (10) is two. It is possible to cast the problem of calculating the tightest possible bound as a shortest path problem on a graph with a node for each stop of the locomotive. Thus the problem can be solved in polynomial time, see Khuller et al. [4].

A simple yet effective inequality is based on the fact that fueling at a station implies that at least one truck should be contracted there. The **x → y inequality** is formulated as follow for each stop of each locomotive

$$x_{i,j} \leq y_{yard(i,j)} \qquad \forall i,j. \tag{11}$$

These inequalities are particularly strong for stops in infrequently visited yards, since the capacity of the locomotive tank is typically much smaller than the capacity of the tanker trucks.

The **v → x inequality** is based on the observation that a locomotive must fuel at a stop if it arrives there with insufficient fuel to complete the journey to the next stop.

$$1 - \frac{v_{i,j}}{FC_{i,j}} \leq x_{i,j} \quad \forall i,j. \tag{12}$$

Next, constraints (3),(4), and (5) are further tightened. Let

$$LT_i' = \min\left\{LT_i, \sum_{j=1}^{N_i} FC_{i,j}\right\}.$$

Replace (3) and (4) by

$$v_{i,j} + f_{i,j} \leq LT_i' \quad \forall i,j \tag{13}$$

and

$$f_{i,j} \leq LT_i' \cdot x_{i,j} \quad \forall i,j. \tag{14}$$

For (5) it is possible to replace the coefficient of y_k (currently TT_k) by

$$\sum_{(i,j):\, yard(i,j)=k, day(i,j)=d} f_{i,j} \leq y_k \min\left\{TT_k, \sum_{(i,j):yard(i,j)=k, day(i,j)=d} LT_i'\right\} \quad \forall k,d. \tag{15}$$

The **locomotive decomposition inequalities** are based on a relaxation of the problems that ignores the contracting costs of the trucks and allows locomotives to be fueled at any yard. This enables the decomposition of the problem into a collection of computationally tractable sub problems of minimizing the total variable and fixed fueling cost of each locomotive. In Nourbakhsh and Ouyang [3], a similar problem is cast as a shortest path problem and it is not difficult to accomplish this for the problem presented in this paper as well. Let us denote the value of the optimal solution for each locomotive by $LCLB_i$. The following inequalities are valid,

$$\sum_j \left(DC_{i,j} \cdot x_{i,j} + P_{yard(i,j)} \cdot f_{i,j}\right) \geq LCLB_i \quad \forall i. \tag{16}$$

One can replace y_k by the sum of a vector of binary variables. That is $y_k = \sum_{q=1}^{M_k} y_{k,q}$ where M_k is an upper bound on the number of trucks that should be contracted at yard k in an optimal solution. A simple way to obtain M_k is by assuming that a full locomotive tank is dispensed in all locomotive stops. That is,

$$M_k = \max_d \sum_{(i,j):yard(i,j)=k, day(i,j)=d} \lceil LT_i'/TT_k \rceil.$$

Now the symmetry breaking constraint $y_{k,q} \geq y_{k,q+1}$ can be added for all k and $q = 1, \dots, M_k - 1$. While this modification of the model alone does not affect the lower bound obtained by the LP relaxation, it enhances the effect of the set of inequalities presented next.

Locomotive-Yard-Fuel inequality- Let $LYF(i,k)$ be an upper bound on the total amount of fuel that locomotive i can acquire in yard k throughout the planning horizon in any feasible solution. The following inequality is valid for any yard k and locomotive i that patronizes the yard.

$$\sum_{j:yard(i,j)=k} f_{i,j} \leq LYF(i,k) \cdot y_{k,1} \qquad \forall k,i \tag{17}$$

Note that in general $LYF(i,k)$ is less than the total amount of fuel consumed by the locomotive since some sections of the locomotive tour may be too far from the yard to be served by the fuel acquired there. Such a tight upper bound can be calculated in linear time

Finally, a method to exploit dominations that are likely to occur in instances of the problem is presented. To this end, let us define the notion of *non-binding dominating yards.*

A yard k is considered *non-binding* if the total amount of fuel that can be acquired from the yard during each day of the planning horizon in any feasible solution is not greater than the capacity of a single tanker truck. That is,

$$\max_{d} \sum_{(i,j):yard(i,j)=k,day(i,j)=d} LT_i' \leq TT_k.$$

A yard k is said to *dominate* locomotive i if there is an optimal fueling plan for i, i.e. with cost $LCLB_i$, in which the locomotive fuels only at yard k. For any locomotive i that admits a non-binding dominating yard k let $S_{i,k}$ be a set of stops in yard k that realizes such an optimal plan. The equalities

$$x_{i,j} = y_k \quad \forall j \in S_{i,k} \tag{18}$$

stipulate this optimal plan if a truck is contracted in the yard. These equalities are valid in the sense that they may eliminate some but not all of the optimal solutions. Removing many weakly dominated solutions from the search space is obviously desirable. Note that in many practical instances locomotives are scheduled to repeatedly cycle between few yards making it likely that the yard with the lowest fuel cost in the cycle will indeed dominate all locomotives in the cycle. If a locomotive is dominated by two or more different yards, the above equality should be added only for one of these yards.

Recall that the optimal solution for each locomotive separately is calculated for the locomotive decomposition inequality as described above. Thus, checking domination can be accomplished easily in polynomial time.

4. Extensions

In this section several practical extensions of the problem are introduced. It is shown how to incorporate these extensions in the MILP model (1)-(6). The implications of these modifications on the valid inequalities presented in Section 3 are discussed.

Some railway companies operate their locomotive fleet according to a **cyclic schedule** that repeats, for example, every two weeks. Such a schedule can be easily handled by adding the following constraints:

$$v_{i,1} = v_{i,N_i} + f_{i,N_i} - FC_{i,N_i} \quad \forall i \tag{2a}$$

where FC_{i,N_i} represents the fuel consumption of locomotive i in its journey from the last stop of a cycle to the first in the next one. In this case, there is no "initial inventory" of fuel and thus $v_{i,1}$ is a decision variable rather than a parameter. The valid inequalities presented in the previous section are readily adaptable to this cyclic formulation. Using a cyclic schedule makes sense if one wishes to use the model for the long term contracting decisions, possibly where the detailed future locomotives schedule is unknown.

Various useful **operational rules** can be expressed by constraints of the shape

$$\sum_{(i,j)\in S} x_{i,j} \leq b. \tag{19}$$

It can be used, for example, to enforce limitations on the number of refueling operations a locomotive is permitted per day. Other possible uses can be to limit the number of trains that are allowed to refuel on the same day in a particular yard, in order to avoid congestion. All the inequalities introduced in Section 3 are still valid with these new constraints. The dominations rules (18) are also applicable but consistency with the new constraints should be tested for the sub-problems as well.

When a locomotive runs out of fuel, the railway may order a **special delivery** to its location. The cost of fuel delivered by this mode is significantly higher than fuel provisioned by the contracted trucks. Hence, this procedure is typically reserved for emergencies and is not part of the railway fueling plan. Indeed, Sonami [7] confirms that "while special deliveries may be used when locomotives run out of fuel unexpectedly, this practice is not a part of the railway master plan". However, if one wishes to allow such "planned" emergency fueling it can be incorporated in the mathematical model presented in this paper, as shown in the appendix.

Some railway operators own **fixed refueling facilities** at some of the most frequented yards. The capital and operational cost of these facilities can be considered sunk costs for the time horizon that is relevant for the contracting decisions. In addition, the amount of fuel that can be provisioned by these facilities is non-binding. Clearly, it makes no sense to contract tanker trucks at these locations but the existence of fixed facilities affect the fueling plan and the optimal contracting decisions throughout the system. In order to accommodate a fixed facility in yard k, it is merely required to set its contracting cost, CC_k to zero.

5. Numerical experiment

This section presents the results of some numerical experiments conducted in order to test the effectiveness of the developed valid inequalities. The model, including all the valid inequalities described in Section 3, and the sub-problems needed to generate some of these inequalities, were implemented in Ilog-OPL and solved using Ilog-Cplex 12.1. The experiments were run on an Intel Xenon X3450® 2.67GHz workstation with 16GB of RAM.

The first tested instance was introduced as a challenge by the 2010 RAS problem solving competition and is available via the INFORMS RAS website. It is characterized by a fourteen day cyclic schedule. The schedule of each locomotive is divided into fourteen "runs" and the number of refueling operations during each run, not including the origin stops, is limited to two. This limitation is a special case of (19). The instance consists of 74 yards, 214 locomotives and some 5264 stops. The delay cost is $250 for all stops, the contracting cost is $8000, the tanker truck capacity is 25000 gallons/day, and the locomotive tank capacity is 4500 gallons.

This instance was solved using the enhanced model with an optimality gap of $0.30 out of total cost of more than 11 million dollars. While it took some 24 hours to establish this optimality gap, the best solution, as well as an optimality gap of 0.01%, was obtained in less than five minutes.

The next 48 instances are based on random networks of 75, 100 and 125 yards. Twelve day cyclic schedules of 5000 and 10000 stops were randomly generated. For each combination of these factors, two possible uniform values were selected for the contracting costs ($5000 and $7000), tanker truck capacities (25000 and 50000 gallons), and locomotive tank capacities (3500 and 5500 gallons). The delay cost in all stops was set to $250.

In order to examine the effectiveness of the valid inequalities, both the basic model (1)-(6) and the enhanced model, described in Section 3, were solved. The solver optimality tolerance was set to zero and one hour was allocated for each run. In addition, the linear relaxations of both models were solved in order to estimate the effect of the valid inequalities. Table 1 below summarizes the results of this experiment. In the first column of the table the characteristics of each instance are presented. The lower bounds obtained by the LP relaxation of the basic and enhanced models, at the root node, are presented in the next two columns. The next four columns present the best integer solutions and lower bounds obtained by both models. The next column presents the fraction of the optimality gap of the basic model closed by the enhanced one. That is

$$1-\frac{\text{Solution value}-\text{lower bound (Enahanced Model)}}{\text{Solution value}-\text{lower bound (Basic Model)}}.$$

The difference between the value of the solution obtained from the basic and enhanced models is presented in the rightmost column. This amount represents the actual saving in dollars per (twelve day) planning period as a result of employing the enhanced model.

A constant that represents a naïve lower bound was subtracted from the objective function. This bound is based on the following “back of the envelope calculation”:

$$\min_k\{P_k\}\cdot\sum_{i,j}FC_{i,j}+\sum_i\frac{\sum_j FC_{i,j}\cdot\min_j\{D_{i,j}\}}{LT_i}+\min_k\left\{\frac{CC_k}{\Theta\ \cdot TT_k}\right\}\cdot\sum_{i,j}FC_{i,j}.$$

The first term is a lower bound on the total fuel cost; the second is a lower bound on the total delay cost and last one is a lower bound on the total contracting cost. By subtracting this lower bound the value of the objective function is scaled to the amount of money that is “left on the table” and can be potentially saved.

Instance properties	LP relaxation		MILP solution after an hour					
			Basic		Enhanced			Absolute
Stops/Yards/ Contracting cost/ Tanker capacity/ Locomotive capacity	Basic	Enhanced	Solution	Rel. gap	Solution	Rel. gap	% Gap closed	improve-ment $
5000/75/5000/25000/3500	399,769.04	631,574.04	652,900.78	12.79%	652,484.86	0.06%	99.53%	415.92
5000/75/5000/25000/5500	370,257.99	506,825.10	511,701.75	11.40%	511,701.63	0.00%	100.00%	0.12
5000/75/5000/50000/3500	395,263.95	655,953.21	685,029.31	14.30%	676,612.99	0.00%	100.00%	8416.32
5000/75/5000/50000/5500	365,945.28	530,416.12	531,140.68	8.75%	531,080.80	0.00%	100.00%	59.88
5000/75/7000/25000/3500	402,878.75	676,713.77	720,950.53	18.35%	703,240.57	0.01%	99.96%	17709.96
5000/75/7000/25000/5500	373,277.10	540,871.49	545,180.94	13.21%	545,179.62	0.00%	100.00%	1.32
5000/75/7000/50000/3500	397,252.06	710,844.61	771,489.12	19.34%	736,167.40	0.02%	99.90%	35321.72
5000/75/7000/50000/5500	367,804.21	573,855.10	575,434.17	11.63%	574,143.65	0.00%	100.00%	1290.52
5000/100/5000/25000/3500	444,912.30	690,428.72	708,750.19	11.75%	702,098.75	0.00%	100.00%	6651.44
5000/100/5000/25000/5500	417,235.84	570,222.32	571,982.65	6.42%	571,010.81	0.00%	100.00%	971.84
5000/100/5000/50000/3500	439,130.37	714,518.37	728,386.82	11.15%	726,188.39	0.00%	100.00%	2198.42
5000/100/5000/50000/5500	413,197.28	594,311.97	598,005.02	7.02%	595,100.46	0.00%	100.00%	2904.56
5000/100/7000/25000/3500	448,675.85	743,992.46	767,239.31	16.12%	757,241.83	0.00%	100.00%	9997.48
5000/100/7000/25000/5500	420,181.74	611,023.29	615,159.65	9.05%	611,726.53	0.00%	100.00%	3433.12
5000/100/7000/50000/3500	441,547.22	777,717.97	818,942.46	17.29%	790,967.34	0.00%	100.00%	27975.12
5000/100/7000/50000/5500	414,881.01	644,748.80	645,789.72	8.59%	645,452.03	0.00%	100.00%	337.69
5000/120/5000/25000/3500	425,896.82	695,672.39	715,111.20	14.87%	708,287.76	0.01%	99.97%	6823.44
5000/120/5000/25000/5500	403,492.55	584,490.59	587,169.34	7.44%	585,430.46	0.00%	100.00%	1738.88
5000/120/5000/50000/3500	421,389.98	719,656.41	748,718.90	16.13%	732,271.78	0.00%	100.00%	16447.12
5000/120/5000/50000/5500	399,907.12	608,474.61	609,914.47	8.53%	609,414.47	0.00%	100.00%	500.00
5000/120/7000/25000/3500	429,271.30	761,175.59	780,255.07	16.15%	776,584.55	0.00%	99.99%	3670.52

5000/120/7000/25000/5500	406,337.60	637,950.66	638,476.68	9.21%	638,476.68	0.00%	100.00%	0.00
5000/120/7000/50000/3500	423,248.12	794,753.22	815,775.25	16.85%	810,162.17	0.00%	100.00%	5613.08
5000/120/7000/50000/5500	401,356.35	671,528.28	673,793.19	11.06%	672,054.31	0.00%	100.00%	1738.88
10000/75/5000/25000/3500	792,736.82	1,129,441.45	1,231,054.72	14.73%	1,165,331.02	1.41%	90.40%	65723.70
10000/75/5000/25000/5500	742,934.58	897,612.90	914,060.68	7.01%	913,452.60	0.03%	99.61%	608.08
10000/75/5000/50000/3500	787,895.02	1,170,348.05	1,207,180.46	11.02%	1,191,376.14	0.10%	99.05%	15804.32
10000/75/5000/50000/5500	738,376.81	929,549.31	933,753.60	6.59%	933,659.20	0.00%	100.00%	94.40
10000/75/7000/25000/3500	795,844.16	1,165,161.51	1,244,210.06	15.28%	1,201,940.00	1.20%	92.16%	42270.06
10000/75/7000/25000/5500	745,925.46	928,707.20	949,433.36	8.77%	948,313.72	0.04%	99.59%	1119.64
10000/75/7000/50000/3500	789,922.57	1,222,025.64	1,276,057.06	16.00%	1,244,024.30	0.17%	98.91%	32032.76
10000/75/7000/50000/5500	740,260.05	973,646.57	978,089.76	8.83%	978,089.76	0.00%	100.00%	0.00
10000/100/5000/25000/3500	852,860.72	1,193,190.51	1,237,057.92	11.40%	1,213,283.64	0.17%	98.47%	23774.28
10000/100/5000/25000/5500	798,553.22	984,035.98	996,245.84	8.10%	989,203.20	0.07%	99.11%	7042.64
10000/100/5000/50000/3500	848,517.12	1,237,540.28	1,366,695.81	17.02%	1,255,460.85	0.15%	99.10%	111234.96
10000/100/5000/50000/5500	795,335.65	1,026,376.25	1,030,581.13	7.96%	1,028,271.33	0.00%	100.00%	2309.80
10000/100/7000/25000/3500	855,671.98	1,239,521.63	1,310,691.36	15.45%	1,265,516.64	0.34%	97.81%	45174.72
10000/100/7000/25000/5500	800,643.35	1,020,233.17	1,035,047.12	9.95%	1,026,382.08	0.01%	99.92%	8665.04
10000/100/7000/50000/3500	850,443.54	1,301,242.49	1,478,010.27	22.26%	1,324,389.19	0.28%	98.74%	153621.08
10000/100/7000/50000/5500	796,732.37	1,078,571.88	1,081,853.47	10.19%	1,080,340.23	0.00%	99.98%	1513.24
10000/120/5000/25000/3500	850,716.01	1,229,651.83	1,314,945.40	14.95%	1,244,477.96	0.10%	99.31%	70467.44
10000/120/5000/25000/5500	807,949.40	1,000,031.31	1,008,864.02	9.25%	1,006,078.22	0.09%	98.99%	2785.80
10000/120/5000/50000/3500	847,008.85	1,276,986.99	1,323,262.89	13.06%	1,290,546.45	0.04%	99.71%	32716.44
10000/120/5000/50000/5500	805,128.35	1,046,889.75	1,049,595.75	8.68%	1,047,534.43	0.00%	100.00%	2061.32
10000/120/7000/25000/3500	853,298.04	1,291,053.96	1,350,564.26	16.02%	1,307,584.54	0.12%	99.25%	42979.72
10000/120/7000/25000/5500	810,045.08	1,043,092.51	1,055,538.36	11.00%	1,048,933.20	0.09%	99.19%	6605.16
10000/120/7000/50000/3500	848,539.91	1,356,721.52	1,411,993.29	17.22%	1,372,133.45	0.03%	99.82%	39859.84
10000/120/7000/50000/5500	806,257.36	1,107,495.57	1,114,089.95	11.72%	1,107,988.03	0.00%	100.00%	6101.92

Table 1: Results of computational experiment with up to 10000 stops and 125 yards

It is apparent from Table 1, that the enhanced formulation delivers significantly better solutions and lower bounds compared to the basic one. Indeed, on average, 99.34% of the optimality gap was closed. Moreover, out of the 48 tested instances, 90.4% of the gap was closed in the worst case. The optimality gap delivered from the enhanced formulation is in most cases negligible with an average of 0.12% and 1.41% in the worst tested case while the optimality gap of the basic model gets as high as 22.26%.

The best integer solutions that could be obtained using the enhanced model within one hour were strictly better than those obtained by the basic model in all but two cases (where both solutions were optimal). On average the difference was $18,100 per a twelve day planning period, which is equivalent to a yearly saving of more than $550,000. Given the low margins of profit in the railway industry this represents a substantial saving. The difference between the solutions obtained by the basic and the enhanced model is statistically significant (p=0.5%). It is also apparent that the lower bounds obtained from the LP relaxation of the enhanced model are much stronger compared to the basic one. This explains the superior performances of the enhanced formulation.

The effect of other parameters on the performances of the enhanced formulation was tested. The assumption that larger locomotive tank capacity (p=1.7%) and larger tanker truck capacity (p=3.9%) are associated with smaller optimality gaps was confirmed. Note that larger locomotive tank and tanker truck capacities are likely to allow more non-binding dominating yards. However, the stops and yard subsequence inequalities, as well as the locomotive tank capacity constraint (4), are likely to be stronger if the locomotive tank capacity is smaller.

We tried to compare the performances of our formulation to other solution approaches, such as the Lagrange relaxation heuristic introduced by Nourbakhsh and Ouyang [3], but unfortunately we did not have access to their benchmark problem. Note that in [3] much larger optimality gaps are reported for much smaller problem instances in similar settings. The benchmark instances and complete description of the results are available upon request from the authors.

Next, to see how this approach can be scaled up and to explore its limitations, larger instances with 196 yards and 30,000 stops were created. In this experiment, three hours per run were allocated. The results of this experiment are summarized in Table 2.

Instance properties	LP relaxation		MILP solution after an hour					
			Basic		Enhanced		% Gap closed	Absolute improve-ment $
Stops/Yards/ Contracting cost/ Tanker capacity/ Locomotive capacity	Basic	Enhanced	Solution	Rel. gap	Solution	Rel. gap		
30000/196/5000/25000/3500	1,951,307.31	2,824,408.88	2,980,394.70	11.73%	2,955,328.76	3.14%	73.21%	25,065.94
30000/196/5000/25000/5500	1,808,199.92	2,221,614.05	2,275,481.15	10.40%	2,265,728.95	1.36%	86.97%	9,752.20
30000/196/5000/50000/3500	1,944,973.06	2,931,458.74	3,294,889.74	19.00%	2,977,706.48	0.34%	98.20%	317,183.26
30000/196/5000/50000/5500	1,802,081.66	2,303,900.78	2,320,583.75	10.54%	2,314,743.63	0.07%	99.34%	5,840.12
30000/196/7000/25000/3500	1,955,263.46	2,900,277.22	3,411,248.40	22.82%	3,026,971.94	2.73%	88.05%	384,276.46
30000/196/7000/25000/5500	1,812,024.91	2,293,736.62	2,380,515.57	12.60%	2,350,832.11	1.77%	85.95%	29,683.46
30000/196/7000/50000/3500	1,947,675.72	3,049,308.29	3,471,644.78	23.27%	3,106,246.42	0.53%	97.71%	365,398.36
30000/196/7000/50000/5500	1,804,649.88	2,409,718.26	2,585,728.43	17.69%	2,422,122.11	0.12%	99.34%	163,606.32

Table 2: Results of computational experiment with up to 30000 stops and 196 yards

The relative optimality gap of the solutions delivered by the enhanced model, for the larger instances, is 1.26% on average (3.14% in the worst tested case). On average, the enhanced model is capable of closing 91.1% of the optimality gap left by the basic model. The solutions delivered by the enhanced model are all strictly better than those delivered by the basic one, with an average improvement of $162,600, which is equivalent to a yearly saving of about $4,945,000.

In order to identify which of the proposed valid inequalities are most effective and whether it is worth omitting some of them, the following experiment was conducted: the eight, 10,000 stops / 120 yards instances were solved with each of the nine valid inequalities separately. That is, the basic model supplemented by one set of inequalities at a time. The solutions obtained from these formulations were compared to the ones obtained for the same instances using the basic model. In addition, the enhanced model was solved with each of the sets of valid inequalities removed, one at a time. These solutions were compared to the solution obtained by the enhanced model.

This experiment enables the observation of the marginal contribution of each set when added first or last to the model. The results of the experiment are reported in Table 3. Each row of the table refers to a single set of valid inequalities, specified in the first column. The second to fourth columns refer to the contribution of each set when added first and the last three columns refer to the contribution when added last. For each set the following is reported: 1) the average percentage of the optimality gap closed, 2) the average absolute improvement in the objective function value, 3) the number (count) of the instances, out of the eight tested, for which the addition of the set improved the value of the objective function.

Name of inequalities	**Contribution when added first**			**Contribution when added last**		
	% Gap closed	**Absolute improv.**	**Improv. count**	**% Gap closed**	**Absolute improv.**	**Improv. count**
Stops subsequence **(7)**	-5.68%	2971.84	6	67.45%	538.60	6
Yards subsequence **(8)**	36.94%	4135.07	5	0.90%	-12.37	2
Minimal number of fuels per locomotive **(9)**	-12.52%	2385.67	5	12.73%	3.10	3
$x \rightarrow y$ **(11)**	41.10%	3599.39	5	73.94%	56654.79	6
$v \rightarrow x$ **(12)**	-263.75%	-98804.03	0	0.74%	3.72	2
Tightening constraints (3), (4) and (5), **(13) - (15)**	3.09%	316.87	5	6.37%	42.44	4
Locomotive decomposition **(16)**	6.60%	1310.98	4	91.40%	38.98	4
Locomotive-Yard-Fuel **(17)**	46.24%	2688.39	6	53.42%	373.92	6
Non-binding dominating yards **(18)**	0.14%	-171.29	5	0.63%	-2.80	1

Table 3: Marginal contribution of each set of valid inequalities

One can observe that while the effect of each valid inequality alone is not dramatic, and in some cases even negative, there is a strong positive interaction among the inequalities. Indeed, on average, each one of the nine proposed valid inequalities contribute to the reduction of the optimality gap when added to all other inequalities. Hence, it is worth using all of these inequalities. Moreover, each of the valid inequalities strictly improves the obtained solution, at least for some instances when added last. However, inequalities sets (8) and (18) slightly increase the average solution value. As a result, a slightly better solution is likely to be obtained without these inequalities but at the cost of a weaker lower bound.

All the observations above are highly sensitive to the nature of the problem instances, to the technology and setting of the solver and to the allocated solving time. However, it is clear that all of the proposed inequalities may be useful for solving instances of the locomotives refueling problems. An interesting lesson from this numerical study is that effective MILP models, coupled with a carefully crafted set of valid inequalities and a modern solver, may be competitive with heuristic methods even for large scale NP-Hard problems.

References

[1] J.T. Gray, Quarterly Rail Cost Adjustment Factor Surface First Quarter 2011, Association of American Railroads (2010), http://aar.org/StatisticsAndPublications/Rail-Cost-Indexes, Accessed on 27 December 2010.

[2] RAS (2010), Problem solving competition web page, http://www.informs.org/Community/RAS/Problem-Solving-Competition , Accessed on 27 December 2010.

[3] S.M. Nourbakhsh and Y. Ouyang, Optimal fueling strategies for locomotive fleets in railroad networks, Transportation Research Part B **44** (2010), pp. 1104–1114.

[4] S. Khuller, A.Malekian, and M. Julian, To fill or not to fill: The gas station problem, In L. Arge, M. Hoffmann, and E. Welzl (Eds.): Lecture Notes in Computer Science, Springer-Verlag Berlin Heidelberg **4698** (2007), pp. 534-545.

[5] M.R. Garey and D.S. Johnson, Computers and Intractability, A Guide to the Theory of NP-Completeness , Freeman, San Fransisco, CA, 1979.

[6] Papadimitriou C.H and Yannakakis, M., Optimization, approximation, and complexly classes, J. Comput. Syst. Sci. **43** (1991), pp. 425–440.

[7] Sonami Kamalesh, Manager Operations Research at CSX Transportation, personal communication, November 2010.

Appendix – Emergency fueling

The model presented in [3] allows emergency fueling by special deliveries to any point in the network in addition to fueling at yards, but with higher delay and fuel costs. The model, (1)-(6), can be adapted to this problem by introducing two additional decision variables:

x_{ij}^E An integer variable that represents the number of emergency stops of train i along the segment between the j^{th} and $(j+1)^{th}$ stop.

f_{ij}^E A continuous variable that represnts the amount of fuel acquired in emergency stops of train i along the segment between the j^{th} and $(j+1)^{th}$ stop.

The delay cost for emergency refueling is denoted by D^E and the price of fuel per gallon acquired in emergency refueling is denoted by P^E.

$$minimize \sum_k CC_k \cdot y_k + \sum_{i,j} \left(D_{i,j} \cdot x_{i,j} + P_{yard(i,j)} \cdot f_{ij} + D^E \cdot x_{i,j}^E + P^E \cdot f_{ij}^E \right) \qquad (1')$$

Subject to

$$v_{ij} = v_{i,j-1} + f_{i,j-1} + f_{i,j-1}^E - FC_{i,j-1} \qquad \forall i, j = 2, \dots, N_i \qquad (2')$$

$$v_{i,j} + f_{i,j} \le LT_i \qquad \forall i, j \qquad (3)$$

$$f_{i,j} \le LT_i x_{i,j}, \qquad f_{i,j}^E \le LT_i x_{i,j}^E \qquad \forall i, j \qquad (4')$$

$$\sum_{(i,j):\, yard(i,j) = k, day(i,j) = d} f_{i,j} \le TT_k \cdot y_k \qquad \forall k, d \qquad (5)$$

$$y_k,\ x_{i,j}^E \text{ integer}, x_{i,j}, \text{binary}, \ f_{i,j}, f_{i,j}^E, v_{i,j} \ge 0 \qquad (6').$$

Note the slight modification in the objective function (1), inventory balance constraint (2), and constraint (4).

Some of the valid inequalities presented in Section 3 can be used with this formulation as well, possibly with minor modifications. In particular, inequality (9) can be restated as follows:

$$\sum_{j=1}^{N_i} \left(x_{i,j} + x_{i,j}^E \right) \ge MinFuels(i) \qquad \forall i.$$

Inequalities (11)-(15) can be left unchanged and it is also possible to introduce a new inequality in the spirit of (14):

$$f_{i,j}^E \le LT_i' \cdot x_{i,j} \quad \forall i, j.$$

Inequality (16) can be reformulated as:

$$\sum_j \left(DC_{i,j} \cdot x_{i,j} + P_{yard(i,j)} \cdot f_{i,j} + D^E \cdot x_{i,j}^E + P^E \cdot f_{ij}^E \right) \ge LCLB_i \qquad \forall i.$$